Cloud Types Identification for Meteorological Satellite Image Using Multiple Sparse Representation Classifiers via Decision Fusion

WEI JIN[ID][1], FEI GONG[2], BIAO TANG[1], AND SHANGLI WANG[1]

[1]Faculty of Electrical Engineering and Computer Science, Ningbo University, Ningbo 315211, China
[2]Shanghai Weimob Enterprise Company, Ltd., Shanghai 201900, China

Corresponding author: Wei Jin (jinwei@nbu.edu.cn)

ABSTRACT Meteorological satellite can monitor the weather conditions in large scales effectively; some solutions and researches have been raised for cloud types identification in satellite cloud image analysis. Extracting the features of the satellite image and designing effective classifier play important roles in implementing cloud types identification system. Since different features describe the characteristics of the cloud image in different perspectives, the collaborative utilization of the different features help to improve the accuracy of cloud classification. This paper proposed a new method to identify cloud types from meteorological satellite image using multiple sparse representation classifiers via decision fusion. First, followed by different types of features extracting, multiple sparse representation-based classifiers were trained respectively. Then, the strategy of decision fusion was introduced to fuse the outputs of multiple classifiers. In order to bring about a reasonable fusion rule, the fusion weights were determined by an adaptive iterative procedure, and the iterative procedure was constructed according to the performance of each sub-classifier. Finally, an adaptive weighted fusion was implemented to determine the cloud type according to the outputs of sub-classifiers and their corresponding weights. The experimental results on FY-2G satellite data demonstrate that the proposed method gains higher recognition accuracy than each separated sub-classifier, which suggests that the strategy of decision fusion can take advantage of each sub-classifier. Moreover, the proposed method achieves competitive results when compared with the other state-of-the-art methods. The computation efficiency of the proposed method is also analyzed briefly.

INDEX TERMS Atmospheric remote sensing, cloud classification, decision fusion, satellite cloud imagery, sparse representation-based classifier.

I. INTRODUCTION

About 1/3-1/2 of the earth's surface is covered by different types cloud; meteorological satellite imagery provides cloud distribution information in wide spatiotemporal scale. Since cloud plays a crucial role in weather system and climate change, the specific weather phenomenon is always closely related to the distribution of clouds [1]–[3]. It is of great significance to improve the accuracy of satellite cloud classification so as to enhance the effectiveness of meteorological monitoring and even establish a scientific climate model. Currently, manual interpretation methods are still widely used in satellite cloud imagery analysis, these methods not only

affected by cognitive orientation and judging experience of interpreters but also face challenges to cope with massive cloud data [4], [5]. Therefore, computer aided cloud analysis has become a rapidly expanding field of research [6].

At present, the common cloud classification methods can be roughly divided into three categories [7]: 1) unsupervised methods; 2) supervised methods; 3) artificial neural networks (ANN) methods. The representative unsupervised methods are threshold method, histogram method, and cluster analysis method, etc. [7]–[11]. Threshold method was firstly proposed for cloud detection, and then developed to identify clear sky, translucent clouds/broken clouds, low

clouds, medium clouds, and high clouds, etc. However, due to the error of radiation inversion value of cloud image, together with the changed temperature of earth's surface by geographies and seasons, it is of great difficulty to determine the threshold. As a consequence, unreasonable threshold often bring about unsatisfactory classification results. In order to determine reasonable threshold, histogram-based cloud classification method has come into being. By analyzing the statistical characteristic of two or three dimensional histogram of the whole or part of the satellite image, the gray distribution of various clouds is obtained, and then the appropriate threshold can be chosen to realize cloud type identification. However, the high quality and precision of satellite image is the big prerequisite for successful application of this method, and satellite cloud image usual cannot guarantee this [12], [13]. As for cluster analysis method, it classifies samples into different cloud types according to the principle of minimum distance. Amuer [6] used K-means clustering to classify different clouds in satellite imagery of METEOSAT. By using standard deviation limited adaptive clustering (SDLAC), Berendes *et al.* [14] realized the cloud classification of MODIS satellite imagery. SDLAC adjusts the standard deviation threshold and clustering center via iterative processing to generate the new cloud type. As for supervised methods [7], it mainly includes near neighbor, maximum likelihood estimate (MLE) and support vector machine (SVM), etc. Generally, on the near neighbor method, a kind of distance measurement is used to classify different clouds. Christodoulou *et al.* [15] used an improved K-Nearest Neighbor method to classify the data of METEOSAT 7, and the classification accuracy can be significantly improved. The MLE method used the probability distribution of features to identify different types cloud, and Li *et al.* [16] used MLE to classify the clouds in MODIS imagery, which improved the overall accuracy of cloud classification, except for the medium and low cloud. Based on Structural Risk Minimization and Vapnik-Chervonenkis Dimension, the SVM is an efficient machine learning method, which is used to mine the intrinsic characteristics of training samples, and then it tries to find the best compromise between learning ability and complexity of model, thus achieving better generalization ability in cloud classification [17], [18]. The ANN is a term used for one type of machine based algorithms that can be used in cloud classification. These algorithms classify regions of interest using a methodology that performs similar functions as the human brain, such as understanding, learning, solving problems and taking decisions [19]–[22]. Liu *et al.* [23] used the ANN to realize the cloud classification of FY-2C satellite imagery, this method not only can effectively distinguish cloud region from cloudless region, but also has high accuracy for various cloud types.

The most recent research indicates that sparsely is the intrinsic properties of signal, giving an appropriate dictionary, a natural signal can be represented as a sparse linear combination of some few dictionary atoms, that is to say, sparse representation is a concise way to represent information.

According this theory, sparse representation-based classification (SRC) has been widely used in pattern recognition recently [24]–[27]. As for feature extraction, its essence is to use a small number of coefficients to describe the most information of image, so sparse representation will be conducive for feature expression. Furthermore, due to the complexity and multiplicity of spectral properties coming from various cloud types and their underlying surface, a single pixel is usually the comprehensive reflection of different clouds and surface features. We can believe that it is a linear combination of several components of cloud system in picture, which exactly coincide with the idea of sparse representation that treats the image as a linear combination of multiple atoms. Thus, introducing sparse representation into cloud image analysis is expected to improve the accuracy of cloud classification. Our team has proposed an algorithm of cloud classification in satellite imagery using over-complete dictionary via sparse representation (CCSI-ODSR) [28], which improved the accuracy in most weather systems. In CCSI-ODSR, the K-SVD is used to train an over-complete dictionary for sparse representation of cloud samples, then the sparse representation coefficient matrix is decomposed by SVD to generate a series of subspaces, finally, cloud classification can be realized by subspace projection. Since CCSI-ODSR construct the dictionary simply utilize spectral features of clouds, it doesn't make full use of the other discriminant features, and the accuracy of cloud classification is still to be improved. Because different features have their own advantages in cloud classification, the comprehensive utilizing different types of features will achieve higher cloud classification accuracy [29], and decision fusion is an effective way to realize this purpose.

At present, decision fusion has been widely used in fields such as remote sensing data processing, medical image processing, and speech signal processing and so on [30]–[32]. Mazher *et al.* [33] proposed a land cover classification method based on multi-sensor fusion of correlated probabilities (MFCP). This method uses the output probability of multiple SVM classifiers to realize land classification (cultivated-land, bare-land, wood-land etc.) via decision fusion. The experimental results show that the accuracy of MFCP is better than other methods. In order to perform the classification for urban remote sensing images, Fauvel *et al.* [34] constructed multiple posterior probability based sub-classifiers by using neural network, and the final results of the classification were obtained by fusing the output of each sub-classifier via certain fusion criteria. Prasad *et al.* [35] proposed a robust multi-classifier decision fusion framework for classification hyper-spectral, multi-temporal images by extracting features in multi-modal data space and combining linear discriminant analysis. Experiments show that this method can improve the classification accuracy of hyper-spectral, multi-temporal remote sensing images.

Inspired by the above ideas, this paper proposed a new method to identify cloud types for meteorological satellite image using multiple sparse representation classifiers via decision fusion (MSRC-DF). In MSRC-DF, the minimum

reconstructed residual of the traditional SRC is used to calculate the maximum a posteriori (MAP) of a certain sample for the specific class. Then the advantages of different sub-classifiers can be taken by decision fusion, thus higher accuracy than that of any sub-classifier can be achieved. MSRC-DF provides a new scheme for satellite cloud classification.

The rest of the paper is organized as follows: Section II briefly introduces the cloud feature extracting and cloud classification system for satellite imagery; Section III gives a detailed of proposed MSRC-DF for cloud classification; Section IV presents the experiment results; the conclusion is presented in Section V.

II. SATELLITE DATA AND CLOUD CLASSIFICATION SYSTEM

Satellite imagery provides distribution information of cloud with large spatiotemporal scale; it is an important tool for weather forecasting and climate monitoring. This paper concentrating on developing a new cloud classification method on FY-2G satellite cloud imagery. The following sub-sections will briefly present the cloud features extracting and describe the classification model for cloud type identification.

A. CLOUD FEATURE EXTRACTION

FY-2G satellite is positioned over the equator 105°E, the visible and infrared spin scan radiometer carried by FY-2G has 5 imaging channels, which includes two infrared long wave channels (IR1, IR2), one water vapor channel (IR3), one infrared medium channel (IR4), and one visible spectrum channel (VIS). The imaging channels and their spatial resolution of FY-2G are shown in Table 1.

TABLE 1. The imaging channels of FY-2G satellite.

Channel name	Spectral range (μm)	Spatial resolution (km)
IR1	10.3~11.3	5
IR2	11.5~12.5	5
IR3	6.3~7.6	5
IR4	3.5~4.0	5
VIS	0.55~0.75	1.25

The 5 imaging channels of FY-2G reflect atmospheric physical information from different aspects. Split window (IR1, IR2) can discriminate underlying surface and cloud area, and also reflect the top brightness temperature accurately. The water vapor channel (IR3) can indicate the absorption property of infrared radiation by water vapor. Since the more water vapor the atmosphere contains, the more infrared radiation will be absorbed, and this makes the corresponding area whiter in water vapor imagery. Generally, clouds of different heights contain different moisture content, thus, water vapor imagery can help to classify different cloud types. The IR4 is sensitive to objects with higher temperature. It is usually used for the estimation of underlying surface temperature and the detection of fog and low-level clouds. The VIS

channel can discriminate land surface, water and various clouds, but it is unable to obtain VIS imagery at night. To ensure the generality of proposed method, the satellite data of IR1 ∼ IR4 are used for cloud classification. Moreover, to take the advantages of each channel, we extract five types of features from cloud imagery to form training samples. Based on this, five sparse representation-based classifiers were been trained to implement cloud classification via decision fusion. The details of the five types of features are described as follows:

1) Grayscale (GS) features. The grayscale features are composed by the gray values of satellite cloud images of different channels together with the gray differences between different channels. Grayscale of four channels (IR1 to IR4), are denoted as GIR1, GIR2, GIR3, and GIR4, respectively, The grayscale differences of IR1-IR2, IR1-IR3, IR1-IR4 and IR2-IR3 are denoted as GIR1-GIR2, GIR1-GIR3, GIR1-GIR4 and GIR2-GIR3, respectively. They form 8 dimensional grayscale features.

2) Brightness temperature (BT) features. The brightness temperature features are composed by brightness temperature of the four channels (IR1 to IR4) and the brightness temperature differences of them. The brightness temperatures of the four channels (IR1 to IR4) are denoted as TIR1, TIR2, TIR3 and TIR4, respectively. The brightness temperature differences of the IR1-IR2, IR1-IR3, IR1-IR4, and IR2-IR3 are denoted as TIR1-TIR2, TIR1-TIR3, TIR1-TIR4 and TIR2-TIR3, respectively. They are combined to form 8 dimensional BT features.

3) Texture (TT) features. The mean value, standard deviation, smoothness, 3rd-moment, uniformity and entropy of the four channels (IR1 to IR4) cloud images are used to form 24-dimensional TT features.

4) Time difference (TD) features. We calculate the difference of the gray value and brightness temperature of four channels (IR1 to IR4) at time t_1 and t_2 to form 8-dimensional TD features. In this paper, the interval of t_1 and t_2 is 1 hour.

5) Gabor (GB) features. Results of 2 scales Gabor transform in 3 directions of the four channels (IR1 to IR4) are extracted to form 24-dimensional GB features.

The above five types of features can reflect the essential characteristics of clouds from different perspectives: the gray value (GV) and bright temperature (BT) of each channel can directly reflect the distribution of clouds in different regions. In infrared cloud images (IR1, IR2, and IR4), the most dark areas often indicate the highest temperature areas, such as clear land, followed by clear water such as lakes, and oceans in terms of temperature. Clouds usually have a relatively low temperature, so the bright white areas in infrared imagery often indicate different types of clouds. The water vapor channel (IR3), is a special infrared band; the water vapor imagery can help to estimate the water content of different clouds. Since different cloud types contain different water content,

the water vapor imagery also helps to classify different cloud types.

At the same time, in order to improve the accuracy of cloud classification, it is necessary to block out the interference of underlying surface effectively and highlight the information of cloud layer to the greatest extent. Considering that most of the clouds are located in the lower troposphere, and the infrared split-window channels IR1 and IR2 reflect the characteristics of ground surface and cloud layer respectively; therefore, the differences of grayscale and brightness temperature between IR1 and IR2 can eliminate the interference of ground surface radiation, which help identify cirrus clouds and cumulonimbus clouds. Water vapor channel IR3 represents the characteristics of the upper troposphere, so the differences between IR1/IR2 and water vapor channel IR3 can further enhance the information of cloud layer; the experiments on cloud classification also show that the difference of IR1 and IR3 had a positive correlation with the top height of convective clouds and the difference of IR2 and IR3 had strong ability to express water content in cloud layer. As a near-infrared channel, the band of IR4 is in the energy range of solar radiation, the difference between IR1 and IR4 can alleviate the influence of solar radiation on IR1, and it is helpful to distinguish different cloud systems using the height of cloud layer. Therefore, we utilize the grayscale and brightness temperature differences of IR1-IR2, IR1-IR3, IR1-IR4 and IR2-IR3 as the features of cloud classification, other channel differences, such as IR3-IR4 were not applied due to they lack definite atmospheric physical meaning and had little contribution to cloud classification [23], [28], [29].

As for the texture features, they can describe the smoothness, sparsity, and regularity of cloud imagery [6], [29]. Satellite cloud imagery contains rich texture information, for example, the texture of stratus is smooth and uniform, and the texture of cumulus is corrugated and has some spots, the cirrus contains texture with fibroid shape. Thus extracting discriminative texture will help to improve the accuracy of the cloud classification. Since the gray distribution in image often reflects the spatial repeated cycle of local structure, and the texture with larger cycle is rougher than that with smaller cycle. Feature extracted by histogram statistical technique can show the thickness of texture. Thus, in texture features extraction, histogram statistical technique is used; the mean value, standard deviation, smoothness, 3rd-moment, uniformity and entropy that extracted to form the texture features of satellite cloud imagery. In actual weather systems, various clouds are changing all the time and may develop into each other.

While the time difference (TD) features can indicate the change regulation of various clouds in terms of time changing. TD features of satellite cloud imagery are also commonly used in cloud detection and precipitation prediction [29], [36].

In order to depict the multi-scale and multi-directional information of cloud imagery, we construct 6 Gabor filters in 2 scales and 3 directions to extract 6-dimensional

Gabor (GB) features of each pixel. We choose 2 scales and 3 directions for two reasons: (1) Extracting Gabor features at different scales and directions is helpful to depict the scale and direction characteristics of cloud images [37], [38]; (2)Gabor convolution is computationally expensive, too many scales and directions not only increase the computational complexity of the algorithm but also increase the redundancy of Gabor features. We choose 2 scales and 3 directions to achieve compromise between feature richness and algorithm complexity.

Due to different types of features have their own virtues; combining above five types of features in variety of strategies is the most intuitive way to improve the performance of cloud classification. However, features combination typically result in very high dimensional feature spaces, this adversely affects the performance of classification systems because a large feature space dimensionality necessitates a large training database to accurately model the statistics of class features. One obvious way to alleviate this problem is to train multiple classifiers using each type features, and then design an appropriate decision fusion system "fuses" these individual classifiers results into a final classification for cloud image. In this work, we focus on multi-classifier decision fusion framework rather than design combinations of the five types of features for alleviating the adversities of high dimensionality of feature vectors.

B. CLOUD CLASSIFICATION SYSTEM FOR SATELLITE IMAGERY

The satellite cloud images, recorded by infrared and visible channels, mainly reflect the spectral information, brightness temperature and albedo information of clouds. In general, each pixel of cloud image is a comprehensive reflection of different clouds and underlying surface, thus it is hard to achieve an accurate classification for satellite cloud imagery using routine methods. Now, following the international common practice of cloud classification, according to their height and vertical development, clouds are mainly divided into four families as high cloud, medium cloud, low cloud, and heap cloud [7]. The four families are subdivided into several categories: the high cloud is subdivided into cirrus, cirrostratus, and cirrocumulus; the medium cloud is subdivided into altostratus and altocumulus; the low cloud is subdivided into cumulus, stratus, stratocumulus, and nimbostratus; and the heap cloud mainly refers to cumulonimbus with exuberant vertical development.

According to the above analysis and the specific requirements of satellite cloud image classification for meteorological services, in this paper, each satellite cloud image pixel is classified as one of the following six types: clear land, clear water, heap cloud, high cloud, medium cloud and low cloud. In this classification system, some low level clouds such as cumulus, stratus, nimbostratus, and stratocumulus are all classified as low clouds, since they are mainly made up of water drops and bring continuous rain frequently. Similarly, altostratus and altocumulus are just classified as

medium clouds. Meanwhile, cirrus, cirrostratus, and cirrocumulus are generally made up of ice crystals, with the height of cloud base usually over 5000 m, and which generally do not bring precipitation, are classified as high clouds [7]. As regards heap cloud, it usually composed by cumulonimbus, its cloud top may extend to the scope of a medium-level or even high-level cloud, which reflects a strong updraft, and they usually incur severe convection weather as thunderstorms or heavy rainfall, so cumulonimbus attracts attention in meteorological monitoring.

III. CLOUD CLASSIFICATION USING DECISION FUSION OF MULTIPLE SPARSE REPRESENTATION CLASSIFIERS

In recent years, sparse representation has been widely used in pattern recognition [24]–[28]. Research shows that, the brain visual cortex of primate consists of a large number of neurons that constitute a complex network, but when it receives external stimulus, only a small number of neurons are activated to complete the response process. Due to the neurons in different regions have their own functions, they work together to accomplish different perceptual task [28]. This mechanism is coincides with the idea of sparse representation, and provides a biological foundation of sparse representation based pattern recognition. To mimic this biological property, this paper introduces sparse representation into cloud classification. By converting the minimum reconstructed residual of sparse representation based classifier into the maximum posterior probability; a new cloud identification method is designed. In the method, the cloud samples with different features are used to train multiple sparse representation-based classifiers, which can output maximum posterior probability. Then, an adaptive iterative is performed to determine the fusion weights of sub-classifiers according to the performance of each sub-classifier. Finally, a cloud classification method that based on decision fusion is proposed.

A. SPARSE REPRESENTATION CLASSIFICATION (SRC) MODEL BASED ON POSTERIOR PROBABILITY

In traditional SRC, test sample is classified by the minimum reconstructed residual of sparse representation, while this scheme makes it difficult to realize the decision fusion for multiple sub-classifiers. In this paper, the reconstructed residual is converted into the posterior probability, thus the possibility that test sample belongs to a certain class can be judged intuitively. By adopting maximum posterior probability, the collaborative work mechanism of information perception by different neurons in the visual cortex of brain can be imitated. Then the decision fusion model of multiple sparse representation classifiers can be constructed to improve the accuracy of cloud classification.

Denote $X = [X_1, X_2, \ldots, X_K]$ as the training sample set of K cloud types, $X_i = [x_{i,1}, x_{i,2}, \ldots, x_{i,n_i}] \in \Re^{m \times n_i}$ is the subset of the i-th type(i-th class), n_i is the number of training sample of i-th class, m is the feature dimension of each sample, $x_{i,j} \in \Re^m$ is the j-th sample in X_i, $i = 1, 2, \ldots, K$, $j = 1, 2, \ldots, n_i$, the total number of training samples is

$l = \sum_{i=1}^{K} n_i$. Here, let $X \in \Re^{m \times l}$ as the dictionary for SRC, $x_{i,j}$ is an atom of the dictionary. According to sparse representation, for a test sample $y \in \Re^m$ which belonging to the i-th type, y could be expressed as a linear combination of those atoms $x_{i,j}$ from the i-th sub-dictionary X_i, $i = 1, 2, \ldots, K$, this can be expressed as,

$$y = \alpha_{i,1} x_{i,1} + \alpha_{i,2} x_{i,2} + \cdots + \alpha_{i,n_i} x_{i,n_i} \quad (1)$$

where $\alpha_{i,j}$ are the coding coefficients. If y is represented as a linear combination of the entire dictionary X, ideally, only those coefficients corresponding to sub-dictionary X_i will be nonzero, and the sparse coding coefficients can be obtained by solving the following model [24]:

$$\min_{\alpha} ||y - X\alpha||_2 s.t. ||\alpha||_0 \le E \quad (2)$$

where $\alpha = [\alpha_{1,1}, \ldots \alpha_{1,n_1}, \ldots, \alpha_{i,1}, \ldots \alpha_{i,n_i}, \ldots, \alpha_{K,1}, \ldots \alpha_{K,n_K}]^T$ is the sparse coding coefficients of y, E is a sparse threshold. The solution of (2) is a NP-hard problem, and it is usually approximated by the following ℓ^1- minimization [24],

$$\min_{\alpha} ||\alpha||_1 \quad s.t. ||y - X\alpha||_2 \le \varepsilon \quad (3)$$

where ε is an optional error tolerance. To calculate the sparse coding coefficients α, (3) can be rewritten as the following general Lagrangian model:

$$\alpha = \arg \min_{\alpha} \{||y - X\alpha||_2^2 + \lambda ||\alpha||_1\} \quad (4)$$

where λ is a positive constant, and a homotopy algorithm [39] can be used to solve the ℓ^1- minimization problem. $\delta_i(\cdot)$ is an operator that is introduced to extract the i-th sub-dictionary associated entries of α, such as $\delta_i(\alpha) = [\alpha_{i,1}, \alpha_{i,2}, \ldots, \alpha_{i,n}]^T$, then, the test sample y can be reconstructed as follows:

$$\bar{y}_i = X_i \delta_i(\alpha) \quad (5)$$

Here X_i is the i-th sub-dictionary. The reconstructed residual between y and $\bar{y}_i$ is:

$$r_i(y) = ||y - \bar{y}_i||_2 \quad (6)$$

which indicates how well the i-th sub-dictionary X_i represent y. According to the traditional sparse representation methods, the smaller the value of $r_i(y)$, the more likely y belongs to the cloud type i. So the test sample y can be assigned to the object cloud type i whose $\bar{y}_i$ has the minimal residue with y. Since it is difficult to fuse the results of multiple classifiers by this hard partition method, in this paper we try to convert the residual of y and $\bar{y}_i$ into the maximum posterior probability of y belongs to the i-th cloud type, in order to construct the decision fusion model. Specific details are as follows:

Denote the inverse of residual of y and $\bar{y}_i$ as:

$$\varphi_i = \frac{1}{||y - \bar{y}_i||_2} \quad (7)$$

Then the maximum posterior probability of y belongs to the i-th cloud type can be defined as:

$$P_i(y \in i - \text{th } type \mid \varphi_i) = \frac{\varphi_i}{\varphi_1 + \varphi_2 + \ldots + \varphi_i + \ldots + \varphi_K} \quad (8)$$

Actually, according to sparse representation theory, if y belongs to the i-th type, $\bar{y}_i$ and y are very similar, that is $y \approx X_i \delta_i(\alpha)$, and $r_i(y)$ will be smaller, then φ_i computed by (7) will be larger. Thus the possibility $P_i(y \in i - \text{th } type \mid \varphi_i)$ will be high, so the cloud type of y can be determined as:

$$\text{identity}(y) = \arg \max_i \{P_i(y \in i - \text{th } type \mid \varphi_i)\} \quad (9)$$

where $i = 1, 2, \ldots, K$, denote different cloud types. Based on this model, a soft classification scheme can be offered to design multiple sparse representation classifiers for implementing decision fusion.

B. DECISION FUSION OF MULTIPLE SPARSE REPRESENTATION CLASSIFIERS

The traditional sparse representation-based classifier uses minimum residual to classify test sample, this often leads to serious misclassification in cloud analysis. This paper training multiple sparse representation-based classifiers using different type features of cloud image, then the final classification result of test sample can be obtained via decision fusion. The decision fusion scheme mimics the cooperative mechanism of different neurons in the visual cortex of the brain. Lam and Suen [40] have proved that, if the accuracy of each sub-classifier is higher than 0.5, the accuracy of decision fusion result will be closed to 1.0 as the number of sub-classifiers increases; if the accuracy of each sub-classifier is lower than 0.5, the accuracy of decision fusion result will be worse than that of the single sub-classifier; if the accuracy of each sub-classifier is equal to 0.5, the accuracy of the fusion result will not be changed. Therefore, as long as the accuracy of each sub-classifier is greater than 0.5, it is possible to use decision fusion to improve the overall accuracy of cloud classification.

In order to achieve the best fusion effect, how to confirm every classifier's weight is very important. The tradition method is to fix weight by every classifier's classification performance. Owing to the specific classifiers constructed by different types of features have their own advantages in cloud classification; the fixed-weight is not effective in some complex conditions. Therefore, it is necessary to determine the adaptive weights for multi-classifiers decision fusion. In this paper, five types of features (GV, BT, TT, TD and GB) that described in section II.A are grouped to constructs the training sample set X,

$$X = [X_1, X_2, \ldots, X_K] \quad (10)$$

where K is the number of cloud types, $X_i \in \Re^{m \times n_i}$ is the subset of the i-th cloud type, $x_{i,j}$ is the j-th sample of X_i, $i = 1, 2, \ldots, K, j = 1, 2, \ldots, n_i, n_i$ is the number of training

sample of i-th type, and $x_{i,j}$ is defined as follows:

$$\begin{cases} x_{i,j} = [x_{i,j}^{GV}; x_{i,j}^{BT}; x_{i,j}^{TT}; x_{i,j}^{TD}; x_{i,j}^{GB}] \in \Re^d \\ x_{i,j}^{GV} \in \Re^{d_1}, \quad x_{i,j}^{BT} \in \Re^{d_2}, \quad x_{i,j}^{TT} \in \Re^{d_3}, \quad x_{i,j}^{TD} \in \Re^{d_4}, \quad x_{i,j}^{GB} \in \Re^{d_5} \end{cases} \quad (11)$$

where d_1, d_2, d_3, d_4 and d_5 are the dimensions of five types of features respectively, and $d = d_1 + d_2 + d_3 + d_4 + d_5$. Then, using these five types of features, five sub-classifiers based on sparse representation are trained respectively. Table 2 shows the corresponding feature and feature dimension of five sub-classifiers.

TABLE 2. Feature and Feature Dimension for Five Sub-classifiers.

Sub-classifier	Feature	Feature dimension
Sub-classifier 1	Gray value (GV)	$d_1 = 8$
Sub-classifier 2	Brightness temperature (BT)	$d_2 = 8$
Sub-classifier 3	Texture (TT)	$d_3 = 24$
Sub-classifier 4	Time difference (TD)	$d_4 = 8$
Sub-classifier 5	Gabor (GB)	$d_5 = 24$

In order to achieve classification performance more accurately, adjusting multi-classifier weight by adaptive method is becoming the key for improving recognition rate. The next sub-section will give a detailed description to determine the weights for sub-classifiers.

C. THE DETERMINATION OF WEIGHTS FOR DECISION FUSION

The basic principle in object recognition is to train classifier using labeled samples and then judge which class the new test sample belongs. In this paper, five sparse representation based sub-classifiers are designed using the five types of features respectively, because each classifier produces posteriori probabilities that the test sample belong to each cloud type, we design an adaptive iterative method to determine the fusion weights of five sub-classifiers according to the performance of respective sub-classifier.

Assuming that training samples of the i-th cloud type form a matrix $X_i = [x_{i,1}, x_{i,2}, \ldots, x_{i,n_i}] \in \Re^{m \times n_i}$, and denoting n_i as the number of the i-th type samples. We define a matrix X for the entire training set of all the K cloud types as $X = [X_1, X_2, \ldots, X_K] \in \Re^{m \times n}$, where $n = n_1 + n_2 +, \ldots, + n_K$ is the total number of samples in X. To determine the weight of respective sub-classifier, some labeled cloud samples are used to form the validation set $V = [V_1, V_2, \ldots, V_K] \in \Re^{m \times N}$, and the structure of V is similar to X, m is the feature dimension of each sample, N is the total number of validation samples in V.

Before using V to reasonably determine the fusion weights of sub-classifiers, it is necessary to eliminate the undesirable consequences of noises or outliers in sample set. According to the outputs of respective sub-classifier, a methodology that used to eliminate noises and outliers is summarized as follows: Firstly, X is used to construct dictionaries for different

sub-classifiers according to the method in Section III.A and Section III.B; then, for a labeled validation sample $v_k \in V$, $k = 1, 2, \ldots, N$, its label is Label(v_k). We use the five sub-classifiers to classify v_k, and obtain the posterior probability P_{si}, here, P_{si} is the posterior probability that output by s-th sub-classifier, and it represents the probability that v_k belongs to the i-th type, $s = 1, 2, \ldots, M, i = 1, 2, \ldots, K$. Then the classification result of v_k respect to s-th sub-classifier will be acquired as follows:

$$\text{identity}(v_k)_s = \arg \max_i \{P_{si}\} \tag{12}$$

If identity(v_k)$_s \neq$ Label(v_k) holds for all the sub-classifiers, $s = 1, 2, \ldots, M$, it can be considered that v_k is noise or outlier, in this case, we remove v_k from V. After all the samples are exam using the above method, the filtered validation sample set $\bar{V} \in \Re^{m \times \bar{N}}$ can be obtained, $\bar{N}$ is the total number of validation sample. Then, the validation sample set $\bar{V}$ will be used to determine the fusion weight of respective sub-classifier. The weight for decision fusion of respective sub-classifier is initialized as follows:

$$\omega_s = \frac{1}{M} \tag{13}$$

where $s = 1, 2, \ldots, M, M$ is the number of sub-classifiers, in this paper, $M = 5$. Then, an adaptive iterative algorithm is designed to optimize the ω_s for decision fusion.

Denotes T_m as the maximum number of iterations, and the iteration count T is set to 1 for the first iteration. For a labeled validation sample $\bar{v}_c \in \bar{V}, c = 1, 2, \ldots, \bar{N}$, its label can be indicate as Label($\bar{v}_c$). X is used to construct dictionaries for different sub-classifiers according to the method described in Section III.A and Section III.B. As for the validation sample, the s-th sub-classifier output the posterior probability $\bar{P}_{si}$; here $\bar{P}_{si}$ represents the probability that $\bar{v}_c$ belongs to the i-th cloud type, $s = 1, 2, \ldots, M, i = 1, 2, \ldots, K$. The classification result of $\bar{v}_c$ by s-th sub-classifier can be obtained as,

$$\text{identity}(\bar{v}_c)_s = \arg \max_i \{\bar{P}_{si}\} \tag{14}$$

In this work, we propose that posteriori probabilities across the five sub-classifiers corresponding to each class label are combined using the fusion weights ω_s. The cumulative probability can be computed as follows,

$$u_i = \sum_{s=1}^{M} \omega_s \bar{P}_{si} \tag{15}$$

The cumulative probability corresponding to each class are compared to decide the label of $\bar{v}_c$ by decision fusion using maximum posterior probability as follows,

$$\text{identity}(\bar{v}_c) = \arg \max_i \{u_i\} \tag{16}$$

If the result of decision fusion is consistent with the actual label of the sample, that is to say identity($\bar{v}_c$) = Label($\bar{v}_c$), we examining the classification result of each sub-classifier with the actual label and processing the next sample in case

identity($\bar{v}_c$)$_s$ = Label($\bar{v}_c$) holds for all sub-classifiers. Otherwise, if there are l ($0 < l < M$) sub-classifiers with classification results that differ from Label($\bar{v}_c$), then the weights of these l sub-classifiers will be decreased by a constant δ; in this case, we rank all the sub-classifiers in the descending order according to the posterior probability $\bar{P}_{si}$ (here i refers to the actual label of the sample, that is $i = $ Label($\bar{v}_c$)), then, the weights of the top l sub-classifiers will be increased by δ. For example, if the results of 2 sub-classifiers are not consistent with the actual label of the sample, the weights of these 2 sub-classifiers will be decreased by δ, meanwhile, the weights of 2 sub-classifiers with the largest and second largest posterior probability will be increased by δ. In this way, the sum of the weights can be kept to 1. The updated weights ω_s^* can be formulated as,

$$\omega_s^* = \begin{cases} \omega_s - \delta, & \text{identity } (\bar{v}_c)_s \neq \text{Label}(\bar{v}_c) \\ \omega_s + \delta, & \text{classifier with larger } \bar{P}_{si} \end{cases} \tag{17}$$

Here ω_s^* satisfies the following condition,

$$\sum_{s=1}^{M} \omega_s^* = 1 \tag{18}$$

After that, we do the same for next validation sample in $\bar{V}$, and go on updating the fusion weights. When all the samples in $\bar{V}$ are traversed, the current iteration is completed and set $T = T + 1$, then, go for the next iteration and go on updating the fusion weights until the iteration count T reaches its maximum T_m.
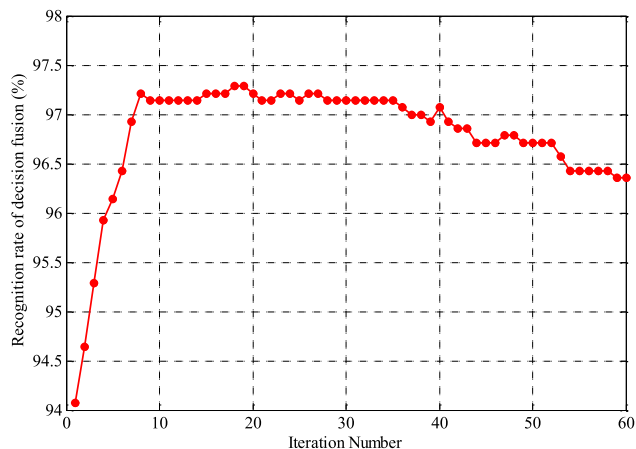

FIGURE 1. Recognition accuracy versus iteration number of decision fusion.

Fig.1 shows the recognition accuracy versus iteration number, it can be seen that, as the iteration goes on, accuracy of decision fusion can be improved at beginning, which means the fusion weights of respective sub-classifier is optimized to some extent.

However, when the number of iterations reaches about 20, the recognition accuracy reaches its peak. If we increase iterations continually, the accuracy of decision fusion decreases
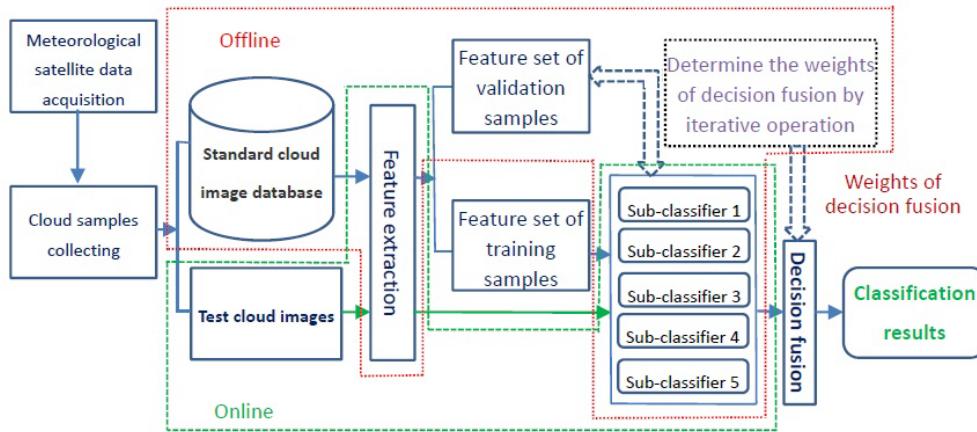
FIGURE 2. The working flow of MSRC-DF.

TABLE 3. Classification accuracy for different cloud types of 5 sub-classifiers.

Cloud Type	Clear Water	Clear Land	Heap Cloud	High Cloud	Medium Cloud	Low Cloud	Overall accuracy
SUB-CLASSIFIER 1	88.50%	94.00%	88.00%	92.00%	91.00%	87.00%	90.08%
SUB-CLASSIFIER 2	91.00%	93.00%	91.50%	94.00%	93.00%	89.50%	92.00%
SUB-CLASSIFIER 3	90.00%	87.50%	79.00%	70.00%	87.00%	95.00%	84.75%
SUB-CLASSIFIER 4	89.50%	60.50%	64.50%	66.50%	58.00%	61.50%	66.75%
SUB-CLASSIFIER 5	74.00%	92.00%	67.50%	65.50%	88.50%	74.50%	77.00%

gradually, this indicates that the fusion weights are excessively regulated. Therefore, in order to determine reasonable fusion weights, the maximum number of iteration is set as $T_m = 20$ in this paper. When the number of iterations reached the set value, the optimized weights can be obtained. The pseudo-code to obtain fusion weights for decision fusion is summarized as follows,

Once the fusion weights are determined, multiple sub-classifiers can be used to realize a cloud classification via decision fusion. Fig.2 shows the working flow of the proposed MSRC-DF.

IV. SIMULATION RESULTS AND ANALYSIS

This section presents the simulations conducted on real satellite data. Experiments were performed on a desktop computer with 2.9 GHz CPU and 8 GB RAM. 9 daytime FY2G satellite data containing IR1, IR2, IR3, IR4, VIS channel imagery with all predefined cloud types are used in our simulations. Three meteorologists examine the selected cloud images and identified all the possible cloud types as well as the background for each pixel based on visual inspection and other relevant knowledge. We select 600 samples for each defined classes among the labeled cloud images. 200 of those labeled samples are randomly drawn for training, and 200 labeled samples of the rest are randomly drawn for validation, while the rest are used for testing. All experiments in sub-sections IV.A ∼ IV.D use the same training samples, and feature vectors for each selected sample are extracted and normalized by ℓ_2-norm. For the sake of experimental comparison, we examined the results generated

by five sub-classifiers, and then we implemented the proposed MSRC-DF by determining the fusion weight for each sub-classifier and obtain fusion result. In our experiments, according to the correlative references, we set the regularization parameter λ as 0.001 for each sub-classifier and other SRC based comparison methods. As to the parameter δ in Eq. (17), it is an empirical parameter, we set it as 0.0002 for adjusting fusion weights adaptively. Finally, a comparison with state-of-the-art methods such as ANN [23], SRC [24], CCSI-ODSR [28], MFCP [33] is presented. The computational efficiency of our method and the compared methods is also evaluated in terms of running time.

A. THE FEASIBILITY OF DECISION FUSION FOR THE OUTPUTS OF FIVE SUB-CLASSIFIERS

In this sub-section, we will verify the feasibility of decision fusion for the outputs of five sub-classifiers. To test the accuracy of each sub-classifier, 200 samples for each cloud type were used for testing. Table 3 shows the classification accuracy of each sub-classifier.

It can be seen from Table 3 that, the classification accuracy of all the five sub-classifiers exceeds 50.00%, which meets the conditions that Lam and Suen [40] proved for effective decision fusion. Thus, by fusing the outputs of the above five sub-classifiers via decision fusion, the proposed MSRC-DF will improve the accuracy of cloud classification. The performance of the proposed MSRC-DF and the five sub-classifiers will be compared in the next sub-section.

TABLE 4. Confusion matrix of the classification results of different cloud types by MSRC-DF.

Cloud Type	Classified as					
	Clear Water	Clear Land	Heap Cloud	High Cloud	Medium Cloud	Low Cloud
Clear water	196	4	0	0	0	0
Clear land	3	197	0	0	0	0
Heap cloud	0	0	194	0	5	1
High cloud	0	0	3	195	1	1
Medium cloud	0	0	0	4	194	2
Low cloud	0	0	2	1	4	193

Algorithm 1 Iteration to obtain fusion weights for decision fusions

1: Input:
Training samples $X = [X_1, X_2, \ldots, X_K]$ and validation samples $V = [V_1, V_2, \ldots, V_K]$ for K types of cloud;
2: Normalize the columns of X and V to have unit ℓ^2 norm. Constructing the dictionaries of five sub-classifiers using X according to different types of cloud feature; eliminating the noises and outliers from V and forming validation set $\bar{V}$;
3: Initialize the fusion weight of each sub-classifier by Eq.(13): $\omega_s = 1/M$ (M is the number of sub-classifiers, in this paper, $M = 5$). Set the maximum iteration number T_m; set iteration count $T = 1$;
4: While $T \leq T_m$
    For $c = 1$ to $\bar{N}$ ($\bar{N}$ is the sample sizes of $\bar{V}$)
        Classify $\bar{v}_c$ via respective sub-classifier
        by Eq.(14);
        The outputs of each sub-classifier will be fused by Eq.(15) and Eq.(16) to identify the identity of $\bar{v}_c$;
        If identity($\bar{v}_c$) = Label($\bar{v}_c$)
            Validating the output of each sub-classifier identity($\bar{v}_c$)$_s$($s = 1$ to M) with the actual label of the sample;
        If identity($\bar{v}_c$)$_s$ = Label($\bar{v}_c$) is true for all sub-classifiers
            Processing the next sample;
            $c = c + 1$;
        Else
            Ranking all the sub-classifiers in the descending order according to the posterior probability $\bar{P}_{si}$(i refers to the actual label of the sample);
            Updating weights ω_s^* for each sub-classifiers according to Eq. (17)
            Processing the next sample;
            $c = c + 1$;
        End if
        End if
    End for
    $T = T + 1$;
End while
5: Output The updated weights ω_s^*.

B. ACCURACY EVALUATION OF MSRC-DF AND SUB-CLASSIFIERS

To test the accuracy of the proposed classification scheme, 200 samples for each cloud type were used for testing.

Table 4 gives a confusion matrix of the classification results of MSRC-DF.

As seen from Table 4, for total 1200 test samples from 6 different cloud types, 1169 samples are correctly classified by MSRC-DF, which means an overall classification accuracy of 97.4%.
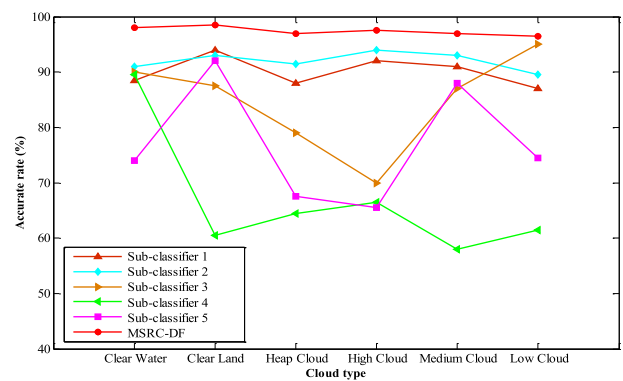

FIGURE 3. Comparison of the classification accuracy (%) of MSRC-DF and five sub-classifiers.

Fig.3 gives a comparison of the classification accuracy of MSRC-DF and the five sub-classifiers. It is clear that MSRC-DF gets better results than each individual sub-classifier. Specifically, sub-classifier 1 and sub-classifier 2 are based on GV and TB features respectively, for any type of cloud, their classification accuracy are relatively stable, and the overall classification accuracy is around 90.00%; sub-classifier 3 is constructed on TT features, it achieve high classification accuracy for low cloud (including cumulus and stratus, etc.) due to the TT features better characterize the texture of cumulus and stratus; meanwhile, GB features can reflect the characteristics of different types cloud in the frequency domain, which makes sub-classifier 5 achieves high classification accuracy for clear land and medium cloud; however, for sub-classifier 4, which trained by time difference features, it achieves acceptable accuracy just for clear water, the accuracy for other types of cloud are unsatisfactory. On the whole, MSRC-DF achieves the highest accuracy for all types of clouds; these indicate that decision fusion is effective in cloud classification.

Table 5 lists the fusion weight of each sub-classifier that used in MSRC-DF. The weight of each sub-classifier can be initialized by Eq.(13), which is a constant of 0.2. From Table 5, it can be seen that, after iterative training, the initial weights of sub-classifiers have been adjusted adaptively. In particular, the adjusted fusion weight of the

**TABLE 5.** Fusion Weight of Each Sub-classifier.

Sub-classifier	Feature	Fusion Weight
Sub-classifier 1	Gray value (GV)	0.3285
Sub-classifier 2	Brightness temperature (BT)	0.2598
Sub-classifier 3	Texture (TT)	0.1925
Sub-classifier 4	Time difference (TD)	0.1190
Sub-classifier 5	Gabor (GB)	0.1002

sub-classifier 1 is the largest and the adjusted fusion weight of the sub-classifier 5 is the smallest, these are almost consistent with the ability of cloud type identification of each sub-classifier.

C. COMPARISONS WITH OTHER METHODS

Many methods for clouds classification have been put forward nowadays and certain success has been achieved. Here, the proposed MSRC-DF was compared with four existing cloud classification methods such as ANN [23], SRC [24], CCSI-ODSR [28] and MFCP [33]. In the experiment, five types of features of cloud samples are simply concatenated to form feature vectors for SRC, ANN and CCSI-ODSR; the ANN classifier was composed by 2 hidden layers, and the neurons of the first and the second layer are 9 and 4 respectively. As a typical decision fusion related method, in the experiment, the MFCP was implemented by fusing the outputs of five SVM based sub-classifiers, and the feature vectors of each sub-classifier is similar to those of MSRC-DF. As in sub-section IV.A and IV.B, randomly selected 200 samples from each cloud type were used for training and 200 samples for testing were selected from the remaining samples. The confusion matrixes of classification results of the comparison methods are shown in Table 6, Table 7, Table 8, and Table 9 respectively, and the confusion matrix of classification results of MSRC-DF is already shown in Table 4 (in sub-section IV.B). Table 10 tabulates the recognition rates for different cloud types using SRC, ANN, CCSI-ODSR, MFCP and MSRC-DF. The overall accuracy for cloud classification of SRC, ANN, CCSI-ODSR, MFCP and MSRC-DF are shown in Fig.4.

From the above experiments, it is clear that MSRC-DF achieves better results than the other methods for almost all cloud types, except MSRC-DF is slightly worse than CCSI-ODSR and MFCP for clear water. For SRC, there are serious misclassification of different types cloud, the overall accuracy of SRC is less than 90%, this indicate that SRC with concatenated different types features cannot be used for satellite cloud image recognition effectively. ANN has a certain ability of self-learning and pattern learning, it has reasonable classification accuracy for clear water and clear land, however, because of its poor generalization ability, the accuracy for heap cloud and low cloud are slightly lower. In case of CCSI-ODSR, for clear water, clear land and medium cloud, the classifcation accuracies are almost the same as the corresponding one of MSRC-DF, but for heap cloud and low
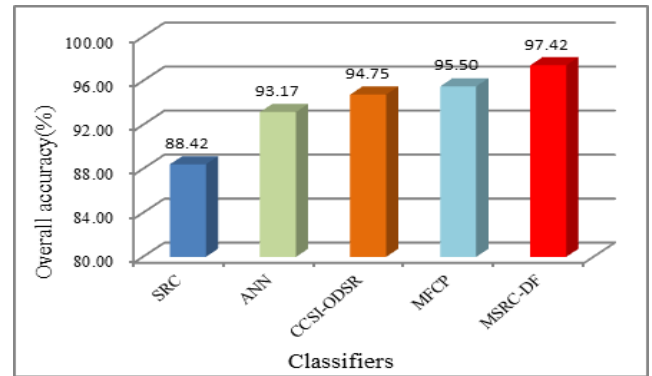
FIGURE 4. The accuracy for cloud classification of SRC, ANN, CCSI-ODSR, MFCP and MSRC-DF.

cloud the classification accuracies are still not satisfactory. As for MFCP, due to it is fuse the outputs of five SVM based sub-classifiers via decision fusion, its performance of cloud identification is superior to SRC, ANN and CCSI-ODSR, especially for clear water, clear land and low cloud. In terms of the overall accuracy, among the five methods, the proposed MSRC-DF provides the best classification results.

D. BENCHMARKS ON FY-2G SATELLITE DATA

To demonstrate the performance of different methods intuitively, in this section, the color-coded cloud classification results by ANN, CCSI-ODSR, MFCP, and MSRC-DF, are visualized. In this experiment, four methods are benchmarked on the FY-2G satellite data obtained at 2 p.m. Beijing time, on 13 September 2016. From the satellite image of each channel, we select specific sub-image with spatial resolution of 512 × 512 pixels cover the main system of typhoon "Meranti" and most of the southeast coast of China, etc.

The IR1 channel image and cloud types labeled image whose cloud types are identified by meteorologist are shown in Fig. 5(a) and (b), respectively. In Fig. 5(b), triangle (▲) indicates clear water, inverted triangle (▼) indicates clear land, star (★) indicates heap cloud, circle (●) indicates low cloud, square (■) indicates medium cloud and cross (+) indicates high cloud.

The color-coded classification results using ANN, CCSI-ODSR, MFCP, and MSRC-DF were shown in Fig.6, whereas the result of SRC was not used for this comparison due to its poor performance.

It can be seen from Fig.6 that, in most of the classification areas, the classification results of ANN, CCSI-ODSR, MFCP, and MSRC-DF are the same as that of the meteorologist-marking image. For example, for clear water and land, the four methods all give relatively reasonable classification results. As for the spiral cloud band of super typhoon "Meranti" and cumulonimbus near the typhoon center, the classification results of these methods can still be acceptable. Considering to the details of the classification results, MSRC-DF gives a better result than that of ANN, CCSI-ODSR and MFCP. For example, in Fig. 6,

TABLE 6. Confusion matrix of the classification results of different cloud types by SRC.

Cloud Type	Classified as					
	Clear Water	Clear Land	Heap Cloud	High Cloud	Medium Cloud	Low Cloud
Clear water	172	20	0	2	0	6
Clear land	14	177	0	0	3	6
Heap cloud	0	0	176	10	4	10
High cloud	0	0	11	174	6	9
Medium cloud	0	0	4	15	179	2
Low cloud	3	4	2	0	8	183

TABLE 7. Confusion matrix of the classification results of different cloud types by ANN.

Cloud Type	Classified as					
	Clear Water	Clear Land	Heap Cloud	High Cloud	Medium Cloud	Low Cloud
Clear water	194	4	0	0	0	2
Clear land	4	196	0	0	0	0
Heap cloud	0	0	177	11	3	9
High cloud	0	0	8	186	4	2
Medium cloud	0	0	1	10	184	5
Low cloud	1	2	9	6	1	181

TABLE 8. Confusion matrix of the classification results of different cloud types by CCSI-ODSR.

Cloud Type	Classified as					
	Clear Water	Clear Land	Heap Cloud	High Cloud	Medium Cloud	Low Cloud
Clear water	197	2	0	0	0	1
Clear land	4	195	0	0	0	1
Heap cloud	0	0	186	3	2	9
High cloud	0	0	7	185	5	3
Medium cloud	0	0	1	5	193	1
Low cloud	1	6	2	0	10	181

TABLE 9. Confusion matrix of the classification results of different cloud types by MFCP.

Cloud Type	Classified as					
	Clear Water	Clear Land	Heap Cloud	High Cloud	Medium Cloud	Low Cloud
Clear water	198	2	0	0	0	0
Clear land	2	196	0	0	0	2
Heap cloud	0	0	186	5	1	8
High cloud	0	2	5	188	3	2
Medium cloud	0	1	2	3	190	4
Low cloud	0	0	2	4	6	188

TABLE 10. Recognition rates of the comparative cloud classification methods.

Cloud Type	Clear Water	Clear Land	Heap Cloud	High Cloud	Medium Cloud	Low Cloud
SRC	86.00%	88.50%	88.00%	87.00%	89.50%	91.50%
ANN	97.00%	98.00%	88.50%	93.00%	92.00%	90.50%
CCSI-ODSR	98.50%	97.50%	93.00%	92.50%	96.50%	90.50%
MFCP	99.00%	98.00%	93.00%	94.00%	95.00%	94.00%
MSRC-DF	98.00%	98.50%	97.00%	97.50%	97.00%	96.50%

compared with the classification results of CCSI-ODSR and MSRC-DF, ANN misclassify some heap clouds as high clouds near the region of spiral cloud band. Meanwhile, some heap clouds around the typhoon center are classified wrongly as low cloud by CCSI-ODSR. Although the overall performance of MFCP is acceptable, it is still difficult to distinguish heap clouds from exterior spiral cloud band. According to the discussion above, in terms of classification accuracy and visual inspection, MSRC-DF achieves the best performance.

E. COMPUTATIONAL EFFICIENCY ANALYSIS

This sub-section considers the computational efficiency of SRC, ANN, CCSI-ODSR, MFCP and MSRC-DF. The training and testing time of the comparison methods are recorded,
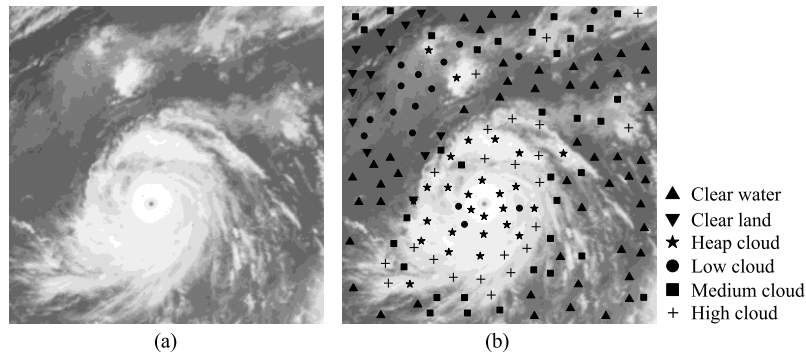
FIGURE 5. Original IR1 image and cloud types labeled image. (a) IR1 image (b) Cloud types labeled by a meteorologist.
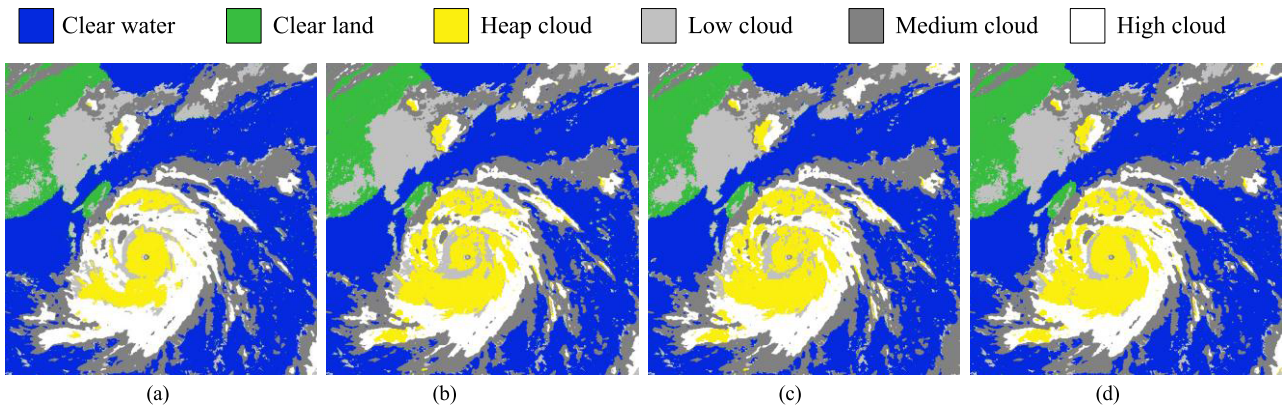

FIGURE 6. Cloud classification results using ANN, CCSI-ODSR, MFCP and MSRC-DF. (a) ANN (b) CCSI-ODSR (c) MFCP (d) MSRC-DF.

TABLE 11. Training time and testing time of different method.

Method	Training time (s)	Testing time (ms)
SRC	Null	3.37
ANN	38.92	3.53
CCSI-ODSR	45.78	11.06
MFCP	33.27	2.93
MSRC-DF	29.51	3.79

as shown in Table 11. For SRC, no training is needed; its training time is ignored. For each method, the training time is the total training cost for all the 6×200 training samples, the testing time is an average for all the 6×200 testing samples.

It is clear that, compared with ANN, CCSI-ODSR and MFCP, the training time of the proposed MSRC-DF is the shortest. The reason lies in that CCSI-ODSR consumes a lot of time to train the dictionary; ANN needs to continuously adjust the network parameters for training network model, which leads to a longer training time; MFCP takes some time to find the optimal classification hyper-planes for each SVM sub-classifier, so its overall training time is still longer. For MSRC-DF, since the dimension of the feature of each sub-classifier is relatively small, and the computation

complexity is relatively lower, this leads to shorter training time. As for testing, several subspaces need to be generated for CCSI-ODSR in testing stage, and then cloud classification is realized by subspace projection, the high computation complexity leads to longest testing time. Meanwhile, the testing cost of MSRC-DF, ANN, SRC and MFCP are in the same order. According to the above experiments, the proposed MSRC-DF achieves high classification accuracy with acceptable time efficiency; MSRC-DF is expected to be applied in the analysis of actual satellite cloud images.

V. CONCLUSION

Clouds play an important role in the balance of atmospheric radiation. The distribution of different clouds depicts the characteristics of atmospheric circulation, and then affects the effectiveness of weather forecasting and climate monitoring. At present, cloud types identification for meteorological satellite image are still not mature under complex conditions. Due to the different types of features can describe the clouds from different aspects, in order to improve the accuracy of cloud classification, it is very important to develop a decision fusion method utilizing the advantages of the different types of features. In this paper, based on FY-2G satellite images, five types of features are extracted for constructing five sparse

representation-based classifiers, and then cloud identification were realized by fusing outputs of every sub-classifiers via decision fusion. This fusion method is effective because the fusion weights were determined by an adaptive iterative scheme. Experimental results show that the proposed method is a better choice than the traditional methods in terms of accuracy and efficiency.

REFERENCES

[1] S. Liu and Z. Zhang, "Learning group patterns for ground-based cloud classification in wireless sensor networks," *EURASIP J. Wireless Commun.*, vol. 13, p. 69, Dec. 2016, doi: 10.1186/s13638-016-0564-x.

[2] J. Gan *et al.*, "Cloud type classification of total-sky images using duplex norm-bounded sparse coding," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 7, pp. 3360–3372, Jul. 2017, doi: 10.1109/JSTARS.2017.2669206.

[3] A. Taravat, F. D. Frate, C. Cornaro, and S. Vergari, "Neural networks and support vector machine algorithms for automatic cloud classification of whole-sky ground-based images," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 3, pp. 666–670, Mar. 2015, doi: 10.1109/LGRS.2014.2356616.

[4] Y. Liu, D.-G. Xi, Z.-L. Li, and C.-X. Shi, "Automatic tracking and characterization of cumulonimbus clouds from FY-2C geostationary meteorological satellite images," *Adv. Meteorol.*, vol. 2014, Aug. 2014, Art. no. 478419, doi: 10.1155/2014/478419.

[5] J. Roman, R. Knuteson, T. August, T. Hultberg, S. Ackerman, and H. Revercomb, "A global assessment of NASA AIRS v6 and EUMETSAT IASI v6 precipitable water vapor using ground-based GPS SuomiNet stations," *J. Geophys. Res., Atmos.*, vol. 121, no. 15, pp. 8925–8948, Aug. 2016, doi: 10.1002/2016JD024806.

[6] Z. Ameur, S. Ameur, A. Adane, H. Sauvageot, and K. Bara, "Cloud classification using the textural features of meteosat images," *Int. J. Remote Sens.*, vol. 25, no. 21, pp. 4491–4503, Nov. 2010.

[7] R. Tapakis and A. G. Charalambides, "Equipment and methodologies for cloud detection and classification: A review," *Sol. Energy*, vol. 95, no. 5, pp. 392–430, Sep. 2013.

[8] L. Liu, X. Sun, F. Chen, S. Zhao, and T. Gao, "Cloud classification based on structure features of infrared images," *J. Atmos. Ocean. Technol.*, vol. 28, no. 3, pp. 410–417, Mar. 2011, doi: 10.1175/2010JTECHA1385.1.

[9] J. Yang, W. Lu, Y. Ma, and W. Yao, "An automated cirrus cloud detection method for a ground-based cloud image," *J. Atmos. Ocean. Technol.*, vol. 29, no. 4, pp. 527–537, Apr. 2012, doi: 10.1175/JTECH-D-11-00002.1.

[10] H. Ishida *et al.*, "Scheme for detection of low clouds from geostationary weather satellite imagery," *Atmos. Res.*, vol. 143, pp. 250–264, Jun. 2014, doi: 10.1016/j.atmosres.2014.02.015.

[11] D. P. Y. Suseno and T. J. Yamada, "Two-dimensional, threshold-based cloud type classification using MTSAT data," *Remote Sens. Lett.*, vol. 3, no. 8, pp. 737–746, Jun. 2012, doi: 10.1080/2150704X.2012.698320.

[12] O. Kärner, "A multi-dimensional histogram technique for cloud classification," *Int. J. Remote Sens.*, vol. 21, no. 12, pp. 2463–2478, Aug. 2000, doi: 10.1080/01431160050030565.

[13] W. Zhuo, Z. Cao, and Y. Xiao, "Cloud classification of ground-based images using texture–structure features," *J. Atmos. Ocean. Technol.*, vol. 31, no. 1, pp. 79–92, Jan. 2014, doi: 10.1175/JTECH-D-13-00048.1.

[14] T. A. Berendes, J. R. Mecikalski, W. M. MacKenzie, Jr., K. M. Bedka, and U. S. Nair, "Convective cloud identification and classification in daytime satellite imagery using standard deviation limited adaptive clustering," *J. Geophys. Res., Atmos.*, vol. 113, p. D20, Oct. 2008, doi: 10.1029/2008JD010287.

[15] C. I. Christodoulou, S. C. Michaelides, and C. S. Pattichis, "Multifeature texture analysis for the classification of clouds in satellite imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 11, pp. 2662–2668, Nov. 2003, doi: 10.1109/TGRS.2003.815404.

[16] J. Li, W. P. Menzel, Z. Yang, R. A. Frey, and S. A. Ackerman, "High-spatial-resolution surface and cloud-type classification from MODIS multispectral band measurements," *J. Appl. Meteorol.*, vol. 42, no. 2, pp. 204–226, Feb. 2003.

[17] Y. Lee, G. Wahba, and S. A. Ackerman, "Cloud classification of satellite radiance data by multicategory support vector machines," *J. Atmos. Ocean. Technol.*, vol. 21, no. 2, pp. 159–169, Feb. 2004.

[18] P. Mills, "Efficient statistical classification of satellite measurements," *Int. J. Remote Sens.*, vol. 32, no. 21, pp. 6109–6132, Jul. 2011, doi: 10.1080/01431161.2010.507795.

[19] A. Mellit and S. A. Kalogirou, "Artificial intelligence techniques for photovoltaic applications: A review," *Prog. Energy Combustion Sci.*, vol. 34, no. 5, pp. 574–632, Oct. 2008.

[20] M. Paliwal and U. A. Kumar, "Neural networks and statistical techniques: A review of applications," *Expert Syst. Appl.*, vol. 36, no. 1, pp. 2–17, Jan. 2009.

[21] M. R. Azimi-Sadjadi, J. Wang, K. Saitwal, and D. Reinke, "A multi-channel temporally adaptable system for continuous cloud classification from satellite imagery," in *Proc. IJCNN*, Washington, DC, USA, 2001, pp. 1625–1630.

[22] J. Alonso-Montesinos, M. Martínez-Durbán, J. D. Sagrado, I. M. D. Águila, and F. J. Batlles, "The application of Bayesian network classifiers to cloud classification in satellite images," *Renew. Energy*, vol. 97, pp. 155–161, Nov. 2016.

[23] Y. Liu, J. Xia, C.-X. Shi, and Y. Hong, "An improved cloud classification algorithm for china's FY-2C multi-channel images using artificial neural network," *Sensors*, vol. 9, no. 7, pp. 5558–5579, Jul. 2009, doi: 10.3390/s90705558.

[24] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 2, pp. 210–227, Feb. 2009, doi: 10.1109/TPAMI.2008.79.

[25] Q. Liu, "Kernel local sparse representation based classifier," *Neural Process. Lett.*, vol. 43, no. 1, pp. 85–95, Feb. 2016.

[26] M. Cui and S. Prasad, "Class-dependent sparse representation classifier for robust hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2683–2695, May 2015, doi: 10.1109/TGRS.2014.2363582.

[27] H. Du, X. Zhang, Q. Hu, and Y. Hou, "Sparse representation-based robust face recognition by graph regularized low-rank sparse representation recovery," *Neurocomputing*, vol. 164, no. 21, pp. 220–229, Sep. 2015.

[28] W. Jin, L. Wang, X. Zeng, Z. Liu, and R. Fu, "Classification of clouds in satellite imagery using over-complete dictionary via sparse representation," *Pattern Recognit. Lett.*, vol. 49, pp. 193–200, Nov. 2014, doi: 10.1016/j.patrec.2014.07.015.

[29] Y. Liu, D.-G. Xi, Z.-L. Li, and C.-X. Shi, "Analysis and application of the relationship between cumulonimbus (Cb) cloud features and precipitation based on FY-2C image," *Atmosphere*, vol. 5, no. 2, pp. 211–229, 2014.

[30] G. Thoonen, Z. Mahmood, S. Peeters, and P. Scheunders, "Multisource classification of color and hyperspectral images using color attribute profiles and composite decision fusion," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 2, pp. 510–521, Apr. 2012, doi: 10.1109/JSTARS.2011.2168317.

[31] M. Takruri, M. W. Rashad, and H. Attia, "Multi-classifier decision fusion for enhancing melanoma recognition accuracy," in *Proc. 5th Int. Conf. Electron. Devices, Syst. Appl.*, 2016, pp. 1–5, doi: 10.1016/j.dsp.2015.11.004.

[32] S. M. Mirhassani, H. N. Ting, and A. A. Gharahbagh, "Fuzzy decision fusion of complementary experts based on evolutionary cepstral coefficients for phoneme recognition," *Digit. Signal Process.*, vol. 49, no. 2, pp. 116–125, Feb. 2016.

[33] A. Mazher and P. Li, "A decision fusion method for land cover classification using multi-sensor data," in *Proc. 4th Int. Workshop Earth Observ. Remote Sens. Appl.*, Guangzhou, China, Jul. 2016, pp. 145–149.

[34] M. Fauvel, J. Chanussot, and J. A. Benediktsson, "Decision fusion for the classification of urban remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 10, pp. 2828–2838, Oct. 2006, doi: 10.1109/TGRS.2006.876708.

[35] S. Prasad, L. M. Bruce, and H. Kalluri, "A robust multi-classifier decision fusion framework for hyperspectral, multi-temporal classification," in *Proc. IGARSS*, Jul. 2008, pp. II273–II276, doi: 10.1109/IGARSS.2008.4778980.

[36] J. S. Bojanowski, R. Stöckli, A. Tetzlaff, and H. Kunz, "The impact of time difference between satellite overpass and ground observation on cloud cover performance statistics," *Remote Sens.*, vol. 6, no. 12, pp. 12866–12884, Dec. 2014, doi: 10.3390/rs61212866.

[37] D. Fu and L. Xu, "Satellite cloud image texture feature extraction based on Gabor wavelet," in *Proc. 4th Int. Congr. Image Signal Process.*, Oct. 2011, pp. 248–251, doi: 10.1109/CISP.2011.6100004.

[38] M. Yang, L. Zhang, S. C. K. Shiu, and D. Zhang, "Gabor feature based robust representation and classification for face recognition with Gabor occlusion dictionary," *Pattern Recognit.*, vol. 46, no. 7, pp. 1865–1878, Jul. 2013.

[39] D. L. Donoho and Y. Tsaig, "Fast solution of ℓ_1-norm minimization problems when the solution may be sparse," *IEEE Trans. Inf. Theory*, vol. 54, no. 11, pp. 4789–4812, Nov. 2008.

[40] L. Lam and C. Y. Suen, "A theoretical analysis of the application of majority voting to pattern recognition," in *Proc. 12nd IAPR Int. Conf. Pattern Recognit., Conf. C, Signal Process.*, vol. 3, Oct. 2002, pp. 418–420, doi: 10.1109/ICPR.1994.576970.

BIAO TANG received the bachelor's degree from Hefei University, China, in 2016. He is currently pursuing the master's degree with the Faculty of Electrical Engineering and Computer Science, Ningbo University. His current research interests include novel algorithm development and image processing.
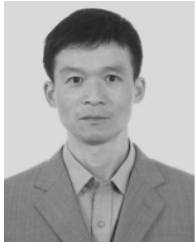
WEI JIN received the Ph.D. degree in optical engineering from Chongqing University, China, in 2006. He is currently a Professor with the Faculty of Electrical Engineering and Computer Science, Ningbo University, China. His current research interests include sparse representation, compressed sensing, computer vision, and image processing.

FEI GONG received the bachelor's degree from Ludong University, China, in 2015, and the master's degree in signal and information processing from Ningbo University, China, in 2018. He then joined the Image Processing Group, WEIMOB, Shanghai. His current research interests include image segmentation, computer vision, and deep learning.

SHANGLI WANG received the bachelor's degree from Xi'an Petroleum University, China, in 2017. She is currently pursuing the master's degree with the Faculty of Electrical Engineering and Computer Science, Ningbo University. Her current research interests include image processing and deep learning.

• • •