Support Vector Machines for Cloud Detection over Ice-Snow Areas

CHEN Gang  E Dongchen

Abstract  In polar regions, cloud and underlying ice-snow areas are difficult to distinguish in satellite images because of their high albedo in the visible band and low surface temperature of ice-snow areas in the infrared band. A cloud detection method over ice-snow covered areas in Antarctica is presented. On account of different texture features of cloud and ice-snow areas, five texture features are extracted based on GLCM. Nonlinear SVM is then used to obtain the optimal classification hyper-plane from training data. The experiment results indicate that this algorithm performs well in cloud detection in Antarctica, especially for thin cirrus detection. Furthermore, when images are resampled to a quarter or 1/16 of the full size, cloud percentages are still at the same level, while the processing time decreases exponentially.

Keywords  cloud detection; SVM; texture analysis; ice-snow covered area; polar region

CLC number  P237.3

Introduction

Clouds in Antarctica play an important role in polar climate and the Earth’s environmental system. It has a strong radioactive influence on the energy balance of ice-snow surfaces[1]. Therefore, an effective cloud and ice-snow area detection algorithm is necessary. Unlike other Earth surfaces, cloud and underlying ice-snow areas have low contrast stemming from their high albedo in the visible band in polar regions. Additional difficulty is from the temperature inversion in the infrared (IR) band because underlying ice-snow areas may have the same or lower temperature than the top cloud. With those characteristics, it is extremely difficult to distinguish cloud and underlying ice-snow areas in Antarctica. At present, some cloud and ice-snow area detection algorithms have been proposed, such as threshold[2-4], wavelet[5], neural network[6], and so on. However, the selection of threshold value is based on many factors, such as remote sensor type, sun elevation angle, atmospheric condition and season. Lots of training samples are also required to improve the detection accuracy in those methods. In this paper, texture features analysis and support vector machines (SVM) are combined for cloud detection over ice-snow covered areas in Antarctica. The experiment threshold value can be avoided in this method, while limited training samples can be used, having good results and being easily calculated in high-dimensional space.

1  Features generated from texture

In polar regions, due to some differences in atmospheric circumfluence, inside air stream and water
vapor content, cloud presents abundant appearances and characteristics. On the other side, underlying ice-snow surface usually has different modalities. Usually, texture is defined as the arrangement of elements with different sizes and shapes composing an image. Grey level co-occurrence matrix (GLCM) is the primary approach used to extract texture features of neighboring pixels in satellite images. After evaluating 14 second-order statistics of texture from GLCM, five were found with enough texture information, namely, angular second moment (ASM), contrast (CON), entropy (ENT), correlation (COR) and local homogeneity (HOM). GLCM is calculated as:

\[ P(i, j, \delta, \theta) = \{(x, y), (x + \Delta x, y + \Delta y)\} | f(x, y) = i, f(x + \Delta x, y + \Delta y) = j; \]
\[ x = 0,1,\cdots,N_x - 1; y = 0,1,\cdots,N_y - 1 \]

where \( i, j = 0,1,\cdots,L - 1 \); \( (x, y) \) is the pixel coordinate in image which has \( L \) grey levels; \( N_x \) and \( N_y \) are row and column number respectively. \( 0^\circ, 45^\circ, 90^\circ, 135^\circ \) were chosen for \( \theta \) and the value of \( \delta \) was 1. Therefore, GLCM of four directions can be achieved. In order to obtain features without directional characteristics, average values of features in \( 0^\circ, 45^\circ, 90^\circ, 135^\circ \) were used.

\section{SVM algorithm analysis}

SVM dwells on the condition of limited samples with the aim of finding out the optimal hyperplane which can minimize the actual risk. Since clouds and ice-snow surfaces are difficult to distinguish in Antarctic areas, accurate training samples are hard to get. But SVM has an advantage to solve this problem of limited samples with the hyperplane theory. Moreover, it can be extended from the linear classifier to the nonlinear classifier, with classification suitable from two classes to multiple classes. The most important thing is that dimension disaster can be avoided in the calculation. Therefore, even clouds and ice-snow surfaces are found with nonlinear relationship, they can still be linearly separated in the high-dimensional feature space and easily compared to nonlinear input feature space.

Supposing training samples are comprised of \( (x_i, y_i) \) pair-wise, where \( i = 1,\cdots,n \), and \( x \in \mathbb{R}^d \), \( y \in \{+1,-1\} \). The hyperplane can be defined as:

\[ f(x, \omega, b) = \text{sgn}(\omega \cdot \varphi(x) + b) \]

where \( \varphi(x) \) is the nonlinear mapping; \( \omega \) and \( b \) are coefficients.

Since additive normal noise may corrupt the observations in the real world, non-negative slack variables \( \xi_i \) are incorporated to handle imperfect separation problems.

\[ \text{minimise} \quad \frac{1}{2} < w \cdot w > + C \sum_{i=1}^{n} \xi_i \]
subject to
\[ y_i (<w \cdot \varphi(x_i) > + b) \geq 1 - \xi_i \], \( \xi_i \geq 0 \), \( i = 1,\cdots,n \)

where regularization parameter \( C(C > 0) \) is used to control the trade-off between the empirical risk and the model complexity.

Then they are converted to solving the dual optimization problem:

\[ \text{maximise} \quad W(\alpha) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{n} y_i y_j \alpha_i \alpha_j K(x_i, x_j) \]
subject to
\[ \sum_{i=1}^{n} y_i \alpha_i = 0, \quad C \geq \alpha_i \geq 0, \quad i = 1,\cdots,n \]

where the optimal Lagrange multipliers \( \alpha_i \) fulfill the Karush-Kuhn-Tucker (KKT) condition:

\[ \alpha_i [ <w \cdot x_i > + b > 1] = 0, \quad i = 1,\cdots,n \]

And only examples \( (x_i, y_i) \) that are support vectors (SVs) can have non-zero coefficients \( \alpha_i \).

Thus, the final decision function is:

\[ f(x, w, b) = \text{sgn}\{ (w \cdot \varphi(x)) + b \} = \text{sgn}\{ \sum_{i=1}^{n} \alpha_i y_i K(x_i, x) + b \} \]

\section{Cloud detection procedure}

Advanced spaceborne thermal emission and reflection radiometer (ASTER) multispectral images of 1 024 pixels \( \times \) 1 050 pixels in 3 different Antarctic areas received in 2001 were used in the research. In order to decrease the features correlation and the training complexity in SVM, in each research site, only three bands with discontinuous spectrum interval were selected from visible and near infrared bands.