Master Thesis Outline TQET33

The Institute of Technology at Linköping University

Maja Gasslander
maja.gasslander@gmail.com

November 18, 2015

1 Introduction

This master thesis will be carried out at Vricon. Supervisor at Vricon is Anton Nordmark, supervisor at ISY is Martin Danelljan and the examiner is Fahad Khan at ISY.

1.1 Contact Information

Anton Nordmark: anton.nordmark@vricon.com
Martin Danelljan: martin.danelljan@liu.se
Fahad Khan: fahad.khan@liu.se

2 Preliminary Title

A preliminary title for the master thesis is *Segmentation of Clouds in Satellite Images.*

3 Problem Formulation

The aim of this thesis is to segment satellite images into cloud and non-cloud pixels. Different combinations of existing methods for feature extraction and classification will be evaluated. The goal is to investigate machine learning based approaches for segmenting the pixels.

4 Approach

A preliminary approach for solving the task is divided into several subproblems. Each subproblem is described in more detail in the subsections below. The proposed pipeline can be seen in figure 1.

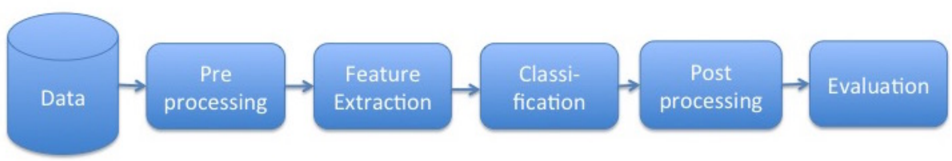

Figure 1: Pipeline

4.1 Data

All images used for cloud classification will be provided by Vricon. Each image contains both panchromatic, i.e. black and white, high resolution data, and also an eight band multispectral imagery. From these images the training data are created and represented as a feature vector, each with an assigned class label. Images with various types of landscape and cloud amount will be used for the classification.

4.2 Preparatory work

Before it is possible to start working on the actual problem, some minor administration has to be done. For example creating a working directory and a user in Vricons system. Also, since a lot of their own functions will be used for handling the images, some small amount of time will be used to understand their code.

4.3 Pre processing

To make the classification easier the images are pre processed. This will be done by tiling the images into smaller tiles. If necessary, a histogram equalization will be applied. This will flatten the greylevel histogram of the image so that all intensities are as equal as possible and thereby the image contrast is increased.

4.4 Feature Extraction

After the pre processing is done it is time to perform the feature extraction. A good feature clearly seperates the two classes from each other. Texture and color are the two features that will recieve the biggest focus in this master thesis.

4.4.1 Texture

- Gabor Filters:
  A set of Gabor filters with different frequencies and orientation might be good to extract texture features from the clouds [5]. The equation for a complex Gabor filter looks like follows:
  $$g(x, y; \lambda, \theta, \psi, \sigma, \gamma) = e^{-\frac{x'^2+\gamma^2 y'^2}{2\sigma^2}} e^{i(2\pi\frac{x'}{\lambda}+\psi)} \tag{1}$$
  where
  $x' = x\cos\theta + y\sin\theta$
  and
  $y' = -x\sin\theta + y\cos\theta$

- Homogeneous Texture Descriptor:
  Clouds have a very homogeneous texture, so the Homogeneous Texture Descriptor (HTD) might be good to extract texture features in the clouds. It is based on computing the local spatial-frequency statistics of the texture, [6].

- Grey level co-occurrence matrix:
  A grey level co-occurrence matrix (GLCM) is a histogram of co-occurring greyscale values at a given offset over an image. The GLCM functions characterize the texture of an image by calculating how often pairs of pixels with specific values and in a specified spatial relationship occur in an image.

- Local Binary Pattern [9]:
  In Local Binary Pattern (LBP), each pixel in an image tile is compared to its 8 neighbouring pixels. For each tile a histogram is computed.

- Histogram of Oriented Gradients [11]:
  The Histogram of Oriented Gradients (HOG) is a feature descriptor that counts the occurences of gradient orientation in a local area of an image.

4.4.2 Color

Since the clouds often differ in color compared to other areas with similar texture (for example water), color might be a good feature for the classification. Two methods for color extraction that will be tested is RGB histogram, and Color Names. Color Naming is the action of assigning linguistic color labels to image pixels [8].

4.5 Classification

For the classification Support Vector Machines (SVM) will be uesd. SVM is a supervised learning method. Consider a set of *training* vectors $\mathbf{x}_i, i = 1, ..., k$, which are the feature vectors, and the corresponding vector $\mathbf{y}$ of length k containing the class values;

$$y_i = \begin{cases} 1, & \text{if } \mathbf{x}_i \text{ in class 1} \\ -1, & \text{if } \mathbf{x}_i \text{ in class 2} \end{cases} \tag{2}$$

Then we have the separating hyperplane as $\mathbf{w}^T\mathbf{x} + b = 0$

$$\begin{cases} (\mathbf{w}^T\mathbf{x}_i) + b > 0 & \text{if } y_i = 1 \\ (\mathbf{w}^T\mathbf{x}_i) + b < 0 & \text{if } y_i = -1 \end{cases} \tag{3}$$

The decision function then becomes $f(\mathbf{x}) = sgn(\mathbf{w}^T\mathbf{x} + b)$, where $\mathbf{x}$ is the testdata. There are many possible choices for $\mathbf{w}$ and b. Equation 3 can be rewritten as

$$y_i(\mathbf{w}^T\mathbf{x}_i + b) \geq 1, \text{for all } 1 \leq i \leq k \tag{4}$$

which gives an optimization problem:

$$\min_{\mathbf{w},b} \frac{1}{2}\mathbf{w}^T\mathbf{w} \tag{5}$$

subject to (for all $i = 1, ..., k$)

$$y_i(\mathbf{w}^T\mathbf{x}_i + b) \geq 1$$

If the data is not linearly seperable it can be mapped to a higher dimensional feature space using the Kernel trick: $K(\mathbf{x}_i, \mathbf{x}_j) = \phi(\mathbf{x}_i)^T\phi(\mathbf{x}_j)$
where

$$\phi(\mathbf{x}_i) = (\phi_1(\mathbf{x}_i), \phi_2(\mathbf{x}_i), ...)$$

Also a slack variable is introduced to allow mislabeled examples. The optimization becomes a trade off between a large margin and a small error penalty. The new equation is:

$$\min_{\mathbf{w},b,\xi} \frac{1}{2}\mathbf{w}^T\mathbf{w} + C\sum_{1=1}^{k}\xi_i \tag{6}$$

$$\xi_i \geq 0, i = 1, ..., k$$

subject to (for all $i = 1, ..., k$)

$$y_i(\mathbf{w}^T\phi(\mathbf{x}_i) + b) \geq 1 - \xi_i$$

Some common kernels that will be tested during this thesis are:

- Linear Kernel: $K(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_i \cdot \mathbf{x}_j$
- Polynomial Kernel: $K(\mathbf{x}_i, \mathbf{x}_j) = (\gamma \mathbf{x}_i \cdot \mathbf{x}_j + r)^d, \gamma > 0$
- Gaussian Radial Basis Function: $K(\mathbf{x}_i, \mathbf{x}_j) = e^{(-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|)^2}, \gamma > 0$
- Sigmoid Kernel: $K(\mathbf{x}_i, \mathbf{x}_j) = \tanh(\gamma \mathbf{x}_i \cdot \mathbf{x}_j + r)$

4.6 Postprocessing

After the classification is done some postprocessing might need to be implemented to clean up some of the wrongly classified objects. Markov Random Fields (MRFs) is an effective classification postprocessing strategy, that smoothes the classification and provides more homogenous results [10]. The MRF model can be written

$$E(X, C) = -\sum_{x \in X} \ln(p_{x,i}) + \beta \sum_{y \in N_x} [1 - \delta(C(x), C(y)]$$

where X and C denotes the image and its labeling space, N_x represents the neighbourhood centered by pixel x. $C(x)$ and $C(y)$ are the labels of pixel x and y. $\delta(x, y)$ is the Kronecker function, i.e. $\delta(x, y) = a$ for $x = y$ and $\delta(x, y) = 0$ for $x \neq y$, used to penalize the change of labels in the neighbourhood N_x.

4.7 Evaluation

When the classification is done the result needs to be evaluated to know how the algorithm performed. This will be done by comparing the classified images to a hand labeled ground truth. Some ground truth already exists, and some might need to be hand labeled during the thesis. Two measures for evaluation will be used:

- Confusion Matrix:
  A Confusion Matrix is a matrix containing the counts of correctly and incorrectly labeled data, see table 1.

Predicted Class

Actual Class	Class 1	Class 2
Class 1	f_{11}	f_{12}
Class 2	f_{21}	f_{22}

Table 1: A Confusion Matrix between 2 classes

- Receiver operating characteristic:
  A receiver operating characteristic (ROC), or ROC curve is a graphical plot that illustrates the performance of a binary classifier. The plot is generated by plotting the true positiv rate against the false positiv rate at different threshold settings.

5 Time Plan

The given task has been divided into several sub tasks to easier plan the work. The estimated time for each sub task is listed in the time plan in figure 2 below. Note that week 53-2 and week 11-12 are marked red because of Christmas break and exams.

5.1 Preliminary dates

- Half-time Control: February 3, 2016.
- Presentation: April 20, 2016.

5.2 Goals for Half-time Control

At the half-time control, preliminary dated to February 3, the main goal is to have the major part of the theory written in the master thesis report. Some preliminary results from the first classifications hopefully will be shown aswell.

ACTIVITIES		Time Plan, week number																									
Nr	Description	45	46	47	48	49	50	51	52	53	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	>17	sum (h)
1	**Writing Planning Report**	7	9	9																								25
2	**Literatur study**	33	30	20	10																							93
3	**Preparatory work**			5	9	10																						24
4	**Pre- and postprocessing**			5	10	8	10	8	5				4	3	3	2	2	2	2	2								66
5	**Feature extraction**				5	5	10	8	5				8	8	8													57
6	**Classification**					4	5	8	5				8	9	9	8	8	5	6									75
7	**Code Documentation**												5	8	7	9	5	3	2	2			5	5				51
8	**Evaluation**					5	5	6	4				5	7	8	5	8	9	9	6			15	15				107
9	**Writing Thesis Report**				5	7	9	9	4				9	4	4	15	16	20	20	29			19	10		4		184
10	**Presentation**																							9	31	20		60
11	**Opposition**																										40	40
12	**Meeting Supervisor**		1	1	1	1	1	1	1				1	1	1	1	1	1	1	1			1	1	1			18
13	**Sum (h)**	40	40	40	40	40	40	40	24	0	0	0	40	40	40	40	40	40	40	40	0	0	40	40	32	24	40	800
14	**Thesis Outline**			18/11																								
15	**Half Time Control**														3/2													
16	**Thesis Report first draft**																			9/3								
17	**Thesis Report**																							6/4				
18	**Presentation**																									20/4		

Figure 2: Time Plan

6 Literature

[1] Emre Başeski and Çağlar Şenaras, *Texture and Color Based Cloud Detection.* HAVELSAN A.S., 2015.

[2] H.K. Chethan, R. Raghavendra and G. Kumar Hemantha, *Texture based Approach for Cloud Classification using SVM.* International Conference on Advances in Recent Technologies in Communication and Computing, 2009.

[3] J. Canny, *A computational Approach to Edge Detection.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 1986.

[4] Rafael C. Gonzales and Richard E. Woods, *Digital Image Processing.* Pearson Education, Inc., 2008.

[5] D. Zheng, Y. Zhao and j. Wang, *Feature Extraction using a Gabor Filter Family.* Proceedings of the 6th IASTED International Conference, Signal and Image Processing, 2004.

[6] Y. Man Ro, M. Kim, H.K. Kang, B.S. Manjunath and J. Kim, *MPEG-7 Homogeneous Texture Descriptor.* ETRI Journal, volume 23, Number 2, 2001.

[7] A. Ben-Hur and J. Weston, *A User's Guide to Support Vector Machines.* ETRI Journal, volume 23, Number 2, 2001.

[8] J. van de Weijer, C. Schmid, J. Verbeek and D. Larlus, *Learning Color Names for Real-World Applications* IEEE Transactions in Image Processing,2009.

[9] Timo Ojala, Matti Pietikäinen and Topi Mäenpää, *Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns* IEEE Trans. Pattern Anal. Mach. Intell., volume 24, Number 7, 2009.

[10] Xin Huang, Qikai Lu, Liangpei Zhang and Antonio Plaza, *New Postprocesssing Methods for Remote Sensing Image Classification: A Systematic Study* IEEE Transactions on Geoscience and Remote Sensing, november 2014.

[11] Navneet Dalal and Bill Triggs, *Histograms of Oriented Gradients for Human Detection* International Conference on Computer Vision & Pattern Recognition, volume 2, 2009. `http://lear.inrialpes.fr/pubs/2005/DT05`