CATEGORIZATION OF CLOUD IMAGE PATCHES USING AN IMPROVED TEXTON-BASED APPROACH

Soumyabrata Dev, Yee Hui Lee

School of Electrical and Electronic Engineering
Nanyang Technological University (NTU)
Singapore 639798

Stefan Winkler

Advanced Digital Sciences Center (ADSC)
University of Illinois at Urbana-Champaign
Singapore 138632

ABSTRACT

We propose a modified texton-based classification approach that integrates both color and texture information for improved classification results. We test our proposed method for the task of cloud classification on SWIMCAT, a large new database of cloud images taken with a ground-based sky imager, with very good results. We perform an extensive evaluation, comparing different color components, filter banks, and other parameters to understand their effect on classification accuracy. Finally, we release the SWIMCAT dataset that was created for the task of cloud categorization.

Index Terms— Cloud texture, classification, ground-based sky imaging

1. INTRODUCTION

The analysis of clouds and study of their features is important for many applications, including climate modeling, weather prediction, solar energy production, satellite communication, and others [1, 2]. The manual classification of cloud categories by experts is expensive and infrequent; thus, it is necessary to come up with automatic and efficient cloud classification methods.

The World Meterological Organization (WMO) recommends a genera-based classification [3], which defines how to classify cloud patches into various categories based on their shape, structure, transparency, arrangement inside the cloud, texture, color, and height of their base. For example, cumulus clouds are puffy in shape and are regularly patterned; stratus clouds are mostly featureless, and are characterized by their transparent, whitish veil appearance; nimbostratus clouds are rain clouds, having thick dark appearance; and so on.

We have proposed a method for three-level labeling of pixels in sky/cloud images [4]. Some of the earlier methods in cloud type recognition use co-occurrence and autocorrelation matrices [5] or Fourier transform [6] to classify pre-defined sky/cloud conditions. Heinle et al. [7] extract several statistical features from pre-defined sky conditions and use k-nearest neighbor classifier during the classification step. Liu et al. [8] proposed an illumination-invariant completed local ternary pattern (ICLTP) descriptor for cloud classification tasks. However, most of these approaches consist of several pre-processing stages and use various pre-defined fixed thresholds in extracting the image features. As the various cloud genera are identified and distinguished amongst themselves primarily by their distinctive color and texture, it is important to seamlessly integrate both these information in the classification framework.

In this paper, we propose a cloud categorization technique to classify various sky/cloud patches. Adapting Varma and Zisserman's textons [9], our proposed approach ably integrates both color and texture information to solve the problem of cloud categorization. Extensive experiments on a large-scale database show that it achieves very good classification accuracy.

2. PROPOSED CLASSIFICATION METHOD

The proposed classification approach presented here is based on the original texton-based texture classification introduced by Varma and Zisserman [9]. The primary differences between our proposed approach and original texton based approach are mainly twofold.

First, the conventional manner consists in generating textons from each of the individual categories, and subsequently creating the texton dictionary. This however may not be an appropriate choice in cases where differences in color and texture of images across different categories are subtle. In our proposed approach, we generate the filter responses of images across all the categories and aggregate them together. Using k-means clustering on the concatenated filter response, we generate the different k-means cluster centers. These clusters centers are the modified-textons; and they constitute the texton dictionary. This method helps exploit the redundancy of information both across and within categories.

Second, the original texton-based classification technique is based on gray-scale images. A few variants for color images have been proposed in the literature. Sun and He [10]

This research is funded by the Defence Science and Technology Agency (DSTA), Singapore.

combined the filter response and RGB color information to form a composite feature vector. Burghouts and Geusebroek [11] used color as a post-processing step on the standard gray-scale texton approach. In this paper, we rely on a methodical choice of color channel that is most suitable for cloud detection [12], and use that as input to the generation of image descriptors.

2.1. Image Descriptor Generation

Assuming there are p categories, and each category has q training images $\mathbf{X}_i \in \mathbb{R}^{m \times n}$, $i = 1, 2, \ldots, q$, having a dimension $m \times n$. The individual images $\mathbf{X}_i$ are convolved with a set of filters from filter bank $\mathbf{F}_1, \mathbf{F}_2, \ldots, \mathbf{F}_k$, for every pixel location (x, y) to obtain $\mathbf{Y}_i$, where $\mathbf{Y}_i \in \mathbb{R}^{m \times n \times k}$:

$$\mathbf{Y}_i = (\mathbf{F}_1 * \mathbf{X}(x, y), \ldots, \mathbf{F}_k * \mathbf{X}(x, y))\}$$

The filter response $\mathbf{Y}_i$ is reshaped to form $\tilde{\mathbf{Y}}_i$ where $\tilde{\mathbf{Y}}_i \in \mathbb{R}^{k \times mn}$. These individual filter responses are concatenated together for all the pq training images in the dataset across all the categories to form $\mathbf{Y}_\Delta \in \mathbb{R}^{k \times mnpq}$.

We primarily use S-filters [13], which are a popular choice for texton-based classification tasks, and tune the filter set to our specific task (see Section 4.2 below). S-Filters are essentially rotation-invariant Gabor-like filters that capture frequency and scale with a sinusoidal waveform attenuated by a Gaussian.

Finally, textons are generated from $\mathbf{Y}_\Delta$ using the conventional k-means clustering algorithm. The cluster centers generated contribute to the modified-texton dictionary.

2.2. Discriminative Model

The discriminative histogram model for each category is generated by comparing the filter responses of the pixels with the generated textons in the dictionary. The particular texton from the dictionary having the smallest 2D-Euclidean norm with the pixel's filter response is mapped to the pixel. For each category $j = 1, 2, \ldots, p$, a total of q histograms are generated that describe the frequency of occurrence of particular textons in that particular category. These histograms are then averaged to determine the discriminative histogram models h_j^{model}.

2.3. Classification Metric

In the testing stage, the corresponding histogram is generated for a test image. The test histogram h^{test} is compared bin-by-bin (k) with each of the model histograms h_j^{model} using χ^2 statistics:

$$d_{\chi^2}(h^{\text{test}}, h_j^{\text{model}}) = \sum_k \frac{1}{2} \frac{(h_k^{\text{test}} - h_{j,k}^{\text{model}})^2}{h_k^{\text{test}} + h_{j,k}^{\text{model}}},$$

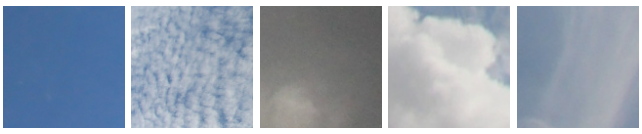
which measures how unlikely it is that one distribution is drawn from the other. The class with the lowest χ^2 distances to the test image is considered as the categorized class for the test image. As the number of histogram bins in training and test images are kept fixed, there are no demerits in using such χ^2 distance bin-by-bin dissimilarity measure.

3. SWIMCAT DATABASE

While there has been extensive work on sky/cloud images image analysis captured by ground-based cameras by a number of research groups around the world, public benchmarking databases in this domain are rare. Currently, the only available database is HYTA [14], which contains sky/cloud images along with their segmentation masks. However, there are no existing databases for cloud classification. We therefore created our own database for this purpose, which we call SWIMCAT (Singapore Whole-sky IMaging CATegories database).

SWIMCAT contains images captured using WAHRSIS, a calibrated ground-based whole sky imager, which was designed by us [15]. We selected a total of 784 patches comprising 5 cloud categories from images that were captured in Singapore over the period January 2013 to May 2014. The 5 categories of clear sky, patterned clouds, thick dark clouds, thick white clouds, and veil clouds are defined based on visual features of sky/cloud conditions, in consultation with experts from the Singapore Meteorological Services. All image patches are of dimension 125×125 pixels. A representative image from each category is shown in Fig. 1.

The SWIMCAT database is available for download from http://vintage.winklerbros.net/swimcat.html.



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

Fig. 1: (a) Clear sky (b) Patterned clouds (c) Thick dark clouds (d) Thick white clouds (e) Veil clouds

4. EXPERIMENTAL EVALUATION & RESULTS

We evaluate our modified texton-based approach described in Section 2 on the SWIMCAT database. The different images of the database are divided into two disjoint sets for training and testing. For training, a total of 40 images per category are chosen. The testing set for each category consists of 45 randomly selected images from the remaining ones. In order to understand the effects of different parameters on the final classification result, we perform extensive evaluations on our database.

		Sky	Patterned Clouds	Thick Dark Clouds	Thick White Clouds	Veil Clouds	Average	P_{macro}	R_{macro}	P_{micro}	R_{micro}
Color channels	S	0.96	1.00	0.82	0.93	0.82	0.91	0.91	0.91	0.87	0.84
	R/B	1.00	1.00	0.98	1.00	0.69	0.93	0.94	0.93	1.00	0.75
	$R-B$	0.93	1.00	0.09	1.00	0.91	0.79	0.79	0.79	0.92	0.89
	$\frac{B-R}{B+R}$	0.76	0.98	0.38	1.00	0.60	0.74	0.78	0.74	0.97	0.65
	C	0.91	1.00	0.38	1.00	0.91	0.84	0.87	0.84	0.95	0.90
Filter sets	Set 1	1.00	1.00	0.49	0.78	0.89	0.83	0.86	0.83	0.86	0.88
	Set 2	0.98	1.00	0.60	0.78	0.89	0.85	0.87	0.85	0.88	0.88
	Set 3	1.00	1.00	0.93	1.00	0.70	0.92	0.93	0.92	0.96	0.76
	Set 4	1.00	1.00	0.53	0.78	0.89	0.84	0.86	0.84	0.88	0.88
	Set 5	0.98	1.00	0.47	0.78	0.89	0.82	0.85	0.82	0.86	0.88
Filter dimension	19×19	0.98	0.93	0.47	0.78	0.91	0.81	0.85	0.81	0.95	0.90
	29×29	1.00	0.78	0.51	0.78	0.67	0.75	0.78	0.75	0.79	0.69
	39×39	1.00	0.98	0.38	0.76	0.82	0.79	0.83	0.79	0.97	0.82
	49×49	1.00	1.00	0.97	1.00	0.68	0.93	0.94	0.93	0.97	0.75
	59×59	0.98	1.00	0.64	0.76	0.89	0.85	0.87	0.85	0.91	0.89
	69×69	0.98	0.84	0.64	0.58	0.87	0.78	0.79	0.78	0.69	0.85
	79×79	0.98	0.71	0.58	0.51	0.84	0.72	0.75	0.72	0.59	0.82
	89×89	0.98	0.71	0.69	0.47	0.87	0.74	0.76	0.74	0.57	0.85
	99×99	1.00	0.69	0.58	0.24	0.87	0.68	0.68	0.68	0.52	0.84
Number of textons	10	1.00	0.44	0.78	0.76	0.47	0.69	0.69	0.69	0.53	0.56
	20	0.98	0.98	0.42	0.8	0.82	0.80	0.84	0.80	0.95	0.82
	30	1.00	1.00	0.97	1.00	0.68	0.93	0.94	0.93	1.00	0.75
	40	1.00	1.00	0.62	0.78	0.91	0.86	0.87	0.86	0.91	0.91
	50	1.00	1.00	0.56	0.80	0.91	0.85	0.87	0.85	0.93	0.91
	60	0.98	1.00	0.58	0.82	0.89	0.85	0.88	0.85	0.91	0.89
	70	0.98	1.00	0.56	0.82	0.89	0.85	0.87	0.85	0.91	0.88
	80	0.98	1.00	0.60	0.80	0.89	0.85	0.87	0.85	0.93	0.89
	90	0.98	1.00	0.58	0.82	0.89	0.85	0.87	0.85	0.93	0.89
	100	0.98	1.00	0.60	0.82	0.91	0.86	0.88	0.86	0.93	0.91

Table 1: Performance evaluation of sky/cloud recognition rates for different color channels, filter sets, filter dimensions, and number of textons. The best performing parameters are highlighted.

For an objective evaluation of this multi-class classification task, we evaluate the individual classification accuracy for each of the 5 categories. We also report the average classification accuracy, macro- and micro-precision and -recall values [16]. Macro-precision and -recall are calculated by averaging the precision and recall values across individual classes, whereas micro-precision and -recall are calculated by aggregating the individual true positives, false positives, and false negatives of the system for different categories. In a multi-class classification task, macro-values give an indication of system performance across all the categories, whereas micro-values quantify the performance of per-image classification.

4.1. Color Channel

Based on the results of an earlier analysis of color channels we conducted for cloud detection [12], we check the performance of Saturation (S) channel of the HSV color space, various common red-blue combinations (R/B, $R-B$, $\frac{B-R}{B+R}$), and chroma $C = \max(R,G,B) - \min(R,G,B)$. All other parameters are kept fixed (Filter set 3, dimension 49×49, 30 textons). From Table 1, it is observed that most of the color channels work well for sky, patterned clouds and thick white clouds because of their distinctive texture and color. Thick dark clouds have a low saturation component and are therefore correctly classified by color channel S and R/B. The best performing color channel R/B is selected for our final classification task.

4.2. Filter Set

For our experiments, we use different sets of S-filters that aim to capture the majority of frequency components, and subsequently choose the most suitable filter set. Table 2 shows the (σ,τ) pairs for various scenarios. The individual filters are L_1-normalized before convolution with the input images.

Table 1 shows the effect of different filter sets (σ,τ) on the final results. The other parameters are kept constant (Color channel R/B, Filter dimension 49×49, 30 textons). The DC component in the filter set is required to capture the low frequencies of the image. An absence of the DC component particularly affects the performance for thick dark clouds, as such images contain pre-dominantly low frequency components. The optimal filter Set 3 corresponds to a scenario that contains DC component and higher values of both σ and τ.

(σ,τ)	Set 1	Set 2	Set 3	Set 4	Set 5
(0.5,0)		✓	✓	✓	✓
(2,1)	✓	✓	✓	✓	✓
(4,1)	✓	✓	✓	✓	✓
(4,2)	✓	✓	✓	✓	✓
(6,1)	✓	✓	✓	✓	✓
(6,2)	✓	✓	✓	✓	✓
(6,3)	✓	✓	✓	✓	✓
(8,1)	✓	✓	✓	✓	✓
(8,2)	✓	✓	✓	✓	✓
(8,3)	✓	✓	✓		✓
(10,1)	✓	✓	✓	✓	
(10,2)	✓	✓	✓	✓	
(10,3)	✓	✓	✓	✓	
(10,4)	✓	✓	✓		
(12,1)			✓		
(12,2)			✓		
(12,3)			✓		
(12,4)			✓		

Table 2: (σ,τ) values for different sets. Set 1 represents the original S-Filter set.

4.3. Filter Dimension

As the *SWIMCAT* database consists of images with fixed dimension 125×125, the optimum filter size needs to be tuned accordingly. As before, we keep the other parameters fixed (Color channel R/B, Filter set 3, 30 textons). From Table 1, we see that the average classification accuracy across different categories is poor for small filter sizes, as they fail to capture the predominant essence of the cloud textures. Also, at large filter size, the boundary effects reduce accuracy. The best performance is achieved with an intermediate filter size of 49×49 pixels, which is thus a good choice for the final classification task.

4.4. Number of Textons

Table 1 shows the effect of the number of textons on the classification performance. The other parameters are kept fixed (Color channel R/B, Filter Set 3, dimension 49×49). The results are poor when the number of textons is small. However, beyond 30 textons, there is no significant performance gain in the classification accuracy. Thus, we select the number of textons as 30 to avoid overfitting.

4.5. Classification Results

Using the parameters selected in the above Sections, we conduct an evaluation of cloud classification on the SWIMCAT database: we use R/B color channel, filter set 3, with filter dimension 49×49 and 30 textons.

The confusion matrix obtained is shown in Fig. 2; each column of the matrix represents the instances in a predicted class, and each row represents the instances in the actual class. Our proposed classification approach achieves near-perfect classification accuracy for most categories. Veil clouds present a challenge because of their resemblance to clear sky. Furthermore, in R/B color representation, veil clouds can be similar in appearance to thick dark clouds and are sometimes mis-classified as such.

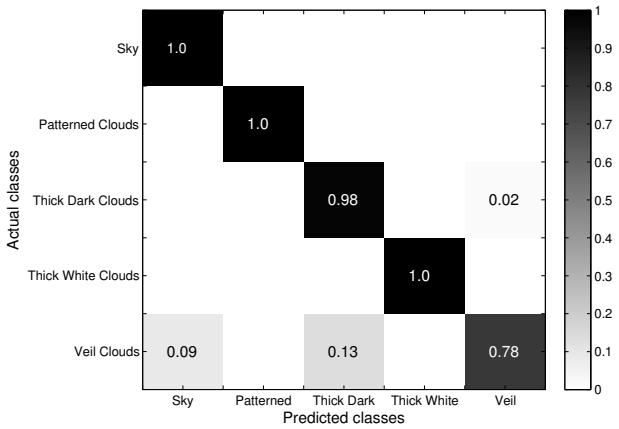


Fig. 2: Confusion matrix using our proposed approach.

Table 3 benchmarks our proposed approach against the original texton approach [9] with Leung-Malik (LM), Schmid (S), or Maximum Response (MR) filters, again using the SWIMCAT database. Amongst the three sets, S-filters are still the best, but the performance of the original texton approach remains far below our proposed method.

Method	Average	P_{macro}	R_{macro}	P_{micro}	R_{micro}
Proposed approach	**0.95**	**0.95**	**0.95**	**0.98**	**0.81**
Original texton [9]					
- with LM filter	0.72	0.67	0.72	0.62	0.64
- with S filter	0.79	0.78	0.79	0.83	0.73
- with MR filter	0.71	0.67	0.71	0.66	0.57

Table 3: Benchmarking results for cloud categorization.

5. CONCLUSION

This paper describes a new texton-based classification method that effectively combines color and texture information for multi-class classification of sky/cloud image patches. The inherent structure of the input data is exploited to determine the appropriate color representation for the task. Furthermore, we introduce SWIMCAT, a large database of ground-based sky/cloud images annotated with cloud categories, which we release to the research community. Extensive experiments are performed using this database; the results show the efficacy of our proposed cloud classification approach.

6. REFERENCES

[1] J. X. Yeo, Y. H. Lee, and J. T. Ong, "Performance of site diversity investigated through RADAR derived results," *IEEE Transactions on Antennas and Propagation*, vol. 59, no. 10, pp. 3890–3898, October 2011.

[2] F. Yuan, Y. H. Lee, and Y. S. Meng, "Comparison of radio-sounding profiles for cloud attenuation analysis in the tropical region," in *Proc. IEEE International Symposium on Antennas and Propagation*, 2014.

[3] World Meteorological Organization, *International Cloud Atlas*, vol. 1, 1975.

[4] S. Dev, Y. H. Lee, and S. Winkler, "Multi-level semantic labeling of sky/cloud images," in *Proc. International Conference on Image Processing (ICIP)*, 2015.

[5] M. Singh and M. Glennen, "Automated ground-based cloud recognition," *Pattern Analysis and Applications*, vol. 8, no. 3, pp. 258–271, October 2005.

[6] J. Calbó and J. Sabburg, "Feature extraction from whole-sky ground-based images for cloud-type recognition," *Journal of Atmospheric and Oceanic Technology*, vol. 25, no. 1, pp. 3–14, January 2008.

[7] A. Heinle, A. Macke, and A. Srivastav, "Automatic cloud classification of whole sky images," *Atmospheric Measurement Techniques*, vol. 3, no. 3, pp. 557–567, 2010.

[8] S. Liu, C. Wang, B. Xiao, Z. Zhang, and Y. Shao, "Illumination-invariant completed LTP descriptor for cloud classification," in *Proc. 5th International Congress on Image and Signal Processing (CISP)*, 2012.

[9] M. Varma and A. Zisserman, "A statistical approach to texture classification from single images," *International Journal of Computer Vision*, vol. 62, no. 1-2, pp. 61–81, April-May 2005.

[10] F. Sun and J. He, "The remote-sensing image segmentation using textons in the normalized cuts framework," in *Proc. International Conference on Mechatronics and Automation (ICMA)*, 2009.

[11] G. Burghouts and J. Geusebroek, "Color textons for texture recognition," in *Proc. British Machine Vision Conference (BMVC)*, 2006.

[12] S. Dev, Y. H. Lee, and S. Winkler, "Systematic study of color spaces and components for the segmentation of sky/cloud images," in *Proc. International Conference on Image Processing (ICIP)*, 2014.

[13] C. Schmid, "Constructing models for content-based image retrieval," in *Proc. Conference on Computer Vision and Pattern Recognition (CVPR)*, 2001.

[14] Q. Li, W. Lu, and J. Yang, "A hybrid thresholding algorithm for cloud detection on ground-based color images," *Journal of Atmospheric and Oceanic Technology*, vol. 28, no. 10, pp. 1286–1296, October 2011.

[15] S. Dev, F. Savoy, Y. H. Lee, and S. Winkler, "WAHRSIS: A low-cost, high-resolution whole sky imager with near-infrared capabilities," in *Proc. IS&T/SPIE Infrared Imaging Systems*, 2014.

[16] V. Van Asch, "Macro- and micro-averaged evaluation measures," Tech. Rep., University of Antwerp, 2013.