

# Color-based Segmentation of Sky/Cloud Images From Ground-based Cameras

Soumyabrata Dev, *Student Member, IEEE*, Yee Hui Lee, *Senior Member, IEEE*,  
and Stefan Winkler, *Senior Member, IEEE*

**Abstract**—Sky/cloud images captured by ground-based cameras (a.k.a. whole sky imagers) are increasingly used nowadays because of their applications in a number of fields, including climate modeling, weather prediction, renewable energy generation, and satellite communications. Due to the wide variety of cloud types and lighting conditions in such images, accurate and robust segmentation of clouds is challenging. In this paper, we present a supervised segmentation framework for ground-based sky/cloud images based on a systematic analysis of different color spaces and components, using partial least squares (PLS) regression. Unlike other state-of-the-art methods, our proposed approach is entirely learning-based and does not require any manually-defined parameters. In addition, we release the Singapore Whole Sky IMaging SEGmentation Database (SWIMSEG), a large database of annotated sky/cloud images, to the research community.

**Index Terms**—Cloud segmentation, whole sky imager, partial least squares regression, SWIMSEG database

## I. INTRODUCTION

CLOUDS have been extensively studied in the research community over the past few decades. The analysis of clouds and their features is important for a wide variety of applications. For example, it has been used for nowcasting to deliver accurate weather forecasts [1], rainfall and satellite precipitation estimates [2], in the study of contrails [3], and various other day-to-day meteorological applications [4]. Yuan et al. have been investigating the clouds' vertical structure [5] and cloud attenuation for optimizing satellite links [6].

Sky/cloud imaging can be performed in different ways. Satellite imagery [7], [8] and aerial photographs [9] are popular in particular for large-scale surveys; airborne light detection and ranging (LiDAR) data are extensively used for aerial surveys [10]. However, these techniques rarely provide sufficient temporal and/or spatial resolution for localized and short-term cloud analysis over a particular area. This is where ground-based whole sky imagers (WSIs) offer a compelling alternative [11]. The images obtained from these devices provide high-resolution data about local cloud formation, movement, and other atmospheric phenomena.

Manuscript received 11-Dec-2015; revised 17-Mar-2016; accepted 09-Apr-2016.

S. Dev and Y. H. Lee are with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore (e-mail: soumyabr001@e.ntu.edu.sg, EYHLee@ntu.edu.sg).

S. Winkler is with the Advanced Digital Sciences Center (ADSC), University of Illinois at Urbana-Champaign, Singapore (e-mail: Stefan.Winkler@adsc.com.sg).

Send correspondence to S. Winkler, E-mail: Stefan.Winkler@adsc.com.sg.

Segmentation is one of the first steps in sky/cloud image analysis. It remains a challenging task because of the non-rigid, feature-less, and poorly-defined structure of clouds, whose shape also changes continuously over time. Thus, classical image segmentation approaches based on shape priors [12] are not suitable. Furthermore, the wide range of lighting conditions (direct sunlight to completely covered skies) adds to the difficulty.

## A. Related Work

As color is the most discriminating feature in sky/cloud images, most works in the literature use color for cloud segmentation. Long et al. [13] showed that the ratio of red and blue channels from RGB color space is a good candidate for segmentation and tuned corresponding thresholds to create binary masks. Heinle et al. [14] exploited the difference of red and blue channels for successful detection and subsequent labeling of pixels. Liu et al. [15] also used the difference of red and blue channels in their superpixel-based cloud segmentation framework. Souza et al. [16] used the Saturation (S) channel for calculating cloud coverage. Mantelli-Neto et al. [17] investigated the locus of cloud pixels in the RGB color model. Li et al. [18] proposed cloud detection using an adaptive threshold technique in the normalized blue/red channel. Yuan et al. [19] proposed a cloud detection framework using superpixel classification of image features.

In these existing methods in the literature for cloud segmentation, the selection of color models and channels has not been studied systematically. Many existing approaches [14], [15], [18], [20], [21] use combinations of red and blue channels, which is a sensible choice, because the sky is predominantly blue due to the Rayleigh scattering of light at shorter wavelengths. However, we are not aware of any experimental analysis presented regarding the efficacy of these color channels in sky/cloud image segmentation. Furthermore, all of the above methods rely on manually-defined parameters and case-based decisions for segmentation. These make the methods somewhat ad-hoc and prone to errors. Finally, most of them assign binary labels by design, which further reduces their flexibility and robustness.

## B. Outline of Our Contribution

The motivation of this paper is to propose a robust framework for color-based cloud segmentation under any illumination conditions, including a systematic analysis of color channels. The framework is based on partial least squares

(PLS) regression and provides a straightforward, parameter-free supervised segmentation method. We show that our approach is robust and offers a superior performance across two different databases as compared to current state-of-the-art algorithms. Furthermore, it allows annotating each pixel with a degree of *belongingness* to the sky or cloud category, instead of the usual binary labeling.

In our previous work [22], we presented an analysis of color channels for sky/cloud images captured by whole-sky cameras, which is an important pre-requisite for better segmentation. The fuzzy c-means clustering method we used in that work however suffers from similar shortcomings as other existing cloud segmentation methods. The main novel contributions of the present manuscript compared to our earlier work include:

- Introduction of a large public sky/cloud image database with segmentation masks;
- Extensive evaluation of color components and selection of appropriate color channels on two different sky/cloud image databases;
- Robust learning-based framework for sky/cloud segmentation that outperforms existing methods.

The rest of this paper is organized as follows. Section II introduces the color spaces under consideration and describes the statistical tools used for subsequent evaluation. Section III discusses the supervised probabilistic segmentation framework. The sky/cloud databases used for evaluation, including our new SWIMSEG database, are presented in Section IV. An exhaustive analysis of color channels is performed in Section V. Section VI presents the experimental evaluation of the segmentation framework, followed by a discussion of the results in Section VII. Section VIII concludes the paper.

## II. APPROACH & TOOLS

In this section, we describe the color models and channels we consider in this paper and present the statistical tools for evaluating their usefulness in sky/cloud image analysis. Specifically, we use Principal Component Analysis (PCA) to check the degree of correlation between the color channels and to identify those that capture the most variance. Loading factors from the primary principal component as well as the Receiver Operating Characteristic (ROC) curve for a simple thresholding applied directly to the color values of a given channel serve as indicators of a channel's suitability for cloud classification.

### A. Color Channels

We consider a set of 16 color channels and components  $c_1, c_2, \dots, c_{16}$  (see Table I). They comprise color spaces RGB, HSV, YIQ, L\*a\*b\*, different red-blue combinations ( $R/B$ ,  $R-B$ ,  $\frac{B-R}{B+R}$ ), and chroma  $C = \max(R, G, B) - \min(R, G, B)$ .

$c_1$	R	$c_4$	H	$c_7$	Y	$c_{10}$	$L^*$	$c_{13}$	$R/B$	$c_{16}$	C
$c_2$	G	$c_5$	S	$c_8$	I	$c_{11}$	$a^*$	$c_{14}$	$R-B$		
$c_3$	B	$c_6$	V	$c_9$	Q	$c_{12}$	$b^*$	$c_{15}$	$\frac{B-R}{B+R}$		

TABLE I: Color spaces and components used for analysis.

In the existing literature, mainly color channels  $c_{1-9}$  [16], [17] and  $c_{13-15}$  [13], [14], [18] have been used for sky/cloud image segmentation. In addition to these, we also consider L\*a\*b\* space ( $c_{10-12}$ ) because of its perceptual uniformity properties as well as chroma ( $c_{16}$ ), because clouds tend to be achromatic.

### B. Principal Component Analysis

We use Principal Component Analysis (PCA) to determine the underlying structure of an image represented by the 16 color channels from Table I and analyze the inherent correlations amongst these components. Consider a sample image  $\mathbf{X}_i$  of dimension  $m \times n$  from a dataset consisting of  $N$  images ( $i = 1, 2, \dots, N$ ). The individual color channels  $c_{1-16}$  are extracted for  $\mathbf{X}_i$  and reshaped into column vectors  $\tilde{\mathbf{c}}_j \in \mathbb{R}^{mn \times 1}$  where  $j = 1, 2, \dots, 16$ . The  $\tilde{\mathbf{c}}_j$  obtained from the different color channels of the sample image  $\mathbf{X}_i$  are stacked alongside, to form the matrix  $\hat{\mathbf{X}}_i \in \mathbb{R}^{mn \times 16}$ .

$$\hat{\mathbf{X}}_i = [\tilde{\mathbf{c}}_1, \tilde{\mathbf{c}}_2, \dots, \tilde{\mathbf{c}}_j, \dots, \tilde{\mathbf{c}}_{16}], \text{ where } i = 1, \dots, N \quad (1)$$

We normalize  $\hat{\mathbf{X}}_i$  with the mean  $\bar{c}_j$  and standard deviation  $\sigma_{c_j}$  of the individual color channels of the image  $\mathbf{X}_i$ , and represent it by  $\ddot{\mathbf{X}}_i$ :

$$\ddot{\mathbf{X}}_i = \left[ \frac{\tilde{\mathbf{c}}_1 - \bar{c}_1}{\sigma_{c_1}}, \frac{\tilde{\mathbf{c}}_2 - \bar{c}_2}{\sigma_{c_2}}, \dots, \frac{\tilde{\mathbf{c}}_j - \bar{c}_j}{\sigma_{c_j}}, \dots, \frac{\tilde{\mathbf{c}}_{16} - \bar{c}_{16}}{\sigma_{c_{16}}} \right] \quad (2)$$

Subsequently, we compute the covariance matrix  $\mathbf{M}_i$  for each of  $\ddot{\mathbf{X}}_i$ . The covariance matrices  $\mathbf{M}_i$  for each of the images provide interesting insights on how correlated the different color channels are, and on the inherent relationships between the various color channels.

The eigenvalue decomposition of  $\mathbf{M}_i$  yields

$$\mathbf{M}_i \mathbf{e}^{st} = \lambda^t \mathbf{e}^{st}, \quad (3)$$

where  $\mathbf{e}^{st}$  represents the  $s^{th}$  eigenvector for  $t^{th}$  eigenvalue, and  $\lambda^t$  represents the  $t^{th}$  eigenvalue. These eigenvalues and eigenvectors are used for computing the relative contributions of different color channels to the new orthogonal axes.

As an indication of which color channel may be most suitable for sky/cloud segmentation, we analyze the relative loading factors of the different color channels on the primary principal component. We calculate the loading factors by considering the relative contribution of the different color channels on the primary principal component. We consider the eigenvector  $\mathbf{e}^{s1}$  corresponding to the primary principal component, and compute the absolute values of the 16 components of this eigenvector as the respective loading factors of 16 color channels.

### C. Discrimination Threshold

In addition to principal component analysis, we also use the Receiver Operating Characteristics (ROC) curve [23], [24] to identify the important color components in sky/cloud image segmentation. The ROC curve represents the fraction of true positives and false positives samples in the sky/cloud binary classification problem for varying values of discrimination

threshold. The area between this ROC curve and the random classifier slope is widely used as a measure for the discriminatory performance in a classification framework.

We consider the set of 16 color components and compute the area under the ROC curve for each, which is subsequently used to identify the best color components for sky/cloud segmentation.

### III. PROBABILISTIC SEGMENTATION

Using statistical tools, we identify the most discriminative color channels for cloud segmentation. Out of 16 color channels, the best performing color channels may be subsequently used as discriminatory feature  $\mathbf{X}_i^f$  in the segmentation framework.

Our objective is to predict the label of a pixel as either sky or cloud, given a set of training images with manually annotated ground truths. We formulate this task of sky/cloud image segmentation in a probabilistic manner using Partial Least Squares (PLS) regression. Both PCA and PLS are based on the same underlying motivation, namely to extract components that accounts for maximum variation in input data.

We start with the assumption that the sky/cloud segmentation problem is dependent on a large number of explanatory independent variables (color channels in this case). Using dimensionality reduction techniques like PCA, we observe the degree of correlation amongst the variables and the amount of variance captured. PCA provides this information based only on the independent input color channels and project the data into new sets of orthogonal axes. In PLS however, both the input color channels and the output labels (sky/cloud) are considered. Using a set of manually defined ground truth images, we model the relationship between the input features and the output label. In the testing stage, we use this estimated model to predict the output labels of the pixels of a given test image.

Suppose, for a sample  $\mathbf{X}_i \in \mathbb{R}^{m \times n}$ , the feature vector is denoted by  $\mathbf{X}_i^f \in \mathbb{R}^{mn \times k}$  and the corresponding binary ground-truth image as  $\mathbf{Y}_i \in \mathbb{R}^{mn \times 1}$ . The feature vector  $\mathbf{X}_i^f$  consists of  $k$  color channels from the set of 16 color channels. We decompose the feature vector  $\mathbf{X}_i^f$  and labels  $\mathbf{Y}_i$  as

$$\mathbf{X}_i^f = \mathbf{T}_i \mathbf{P}_i^T + \mathbf{E}_i, \quad (4)$$

$$\mathbf{Y}_i = \mathbf{U}_i \mathbf{Q}_i^T + \mathbf{F}_i, \quad (5)$$

where  $\mathbf{T}_i \in \mathbb{R}^{mn \times p}$  and  $\mathbf{U}_i \in \mathbb{R}^{mn \times p}$  are  $p$  extracted latent matrices.  $\mathbf{P}_i \in \mathbb{R}^{k \times p}$  and  $\mathbf{Q}_i \in \mathbb{R}^{1 \times p}$  are the loading matrices, and  $\mathbf{E}_i \in \mathbb{R}^{mn \times k}$  and  $\mathbf{F}_i \in \mathbb{R}^{mn \times 1}$  are the residual matrices. Henceforth, for the sake of brevity and without loss of generality we drop the suffix  $i$  in the notations.

In the generalized partial least squares regression sense, we decompose  $\mathbf{X}^f$  and  $\mathbf{Y}$  so as to maximize the covariance between  $\mathbf{T}$  and  $\mathbf{U}$  [25]. Under the assumption that the input matrix  $\mathbf{X}^f$  can be expressed as a linear combination of a few latent variables, we can assume the existence of a linear relationship between  $\mathbf{T}$  and  $\mathbf{U}$  such that

$$\mathbf{U} = \mathbf{T}\mathbf{D} + \mathbf{H}, \quad (6)$$

where  $\mathbf{D} \in \mathbb{R}^{p \times p}$  is a diagonal matrix and  $\mathbf{H}$  is the residual matrix. Combining Eqs. (5) and (6), we get

$$\mathbf{Y} = \mathbf{T}\mathbf{D}\mathbf{Q}^T + (\mathbf{H}\mathbf{Q}^T + \mathbf{F}), \quad (7)$$

$$\mathbf{Y} = \mathbf{T}\mathbf{C}^T + \mathbf{F}^*, \quad (8)$$

where  $\mathbf{C}^T$  is the matrix of regression coefficients and  $\mathbf{F}^*$  is the residual matrix. We extract score vectors  $\mathbf{t}$  and  $\mathbf{u}$  from predictors and response variables respectively, and express as:

$$\mathbf{t} = \mathbf{X}\mathbf{w}, \quad (9)$$

$$\mathbf{u} = \mathbf{Y}\mathbf{c}, \quad (10)$$

where  $\mathbf{w}$  and  $\mathbf{c}$  are the corresponding weight vectors of  $\mathbf{X}$  and  $\mathbf{Y}$ . These score vectors are necessary to define the response and predictor variables in the PLS setting.  $\mathbf{W}$  represents a set of orthogonal basis vectors, such that  $\mathbf{W} = (\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_p)$ . From [26], we can express the regression coefficient matrix as

$$\mathbf{T} = \mathbf{X}\mathbf{W}(\mathbf{P}^T\mathbf{W})^{-1}. \quad (11)$$

Using Eqs. (7) and (11), we get

$$\mathbf{Y} = \mathbf{X}\mathbf{W}(\mathbf{P}^T\mathbf{W})^{-1}\mathbf{C}^T + \mathbf{F}^*, \quad (12)$$

$$\mathbf{Y} = \mathbf{X}\mathbf{B} + \mathbf{F}^*, \quad (13)$$

where  $\mathbf{B} = \mathbf{W}(\mathbf{P}^T\mathbf{W})^{-1}\mathbf{C}^T$  represents the matrix of regression coefficients.

In our proposed approach, we select a set of training images from the database and obtain the regression coefficient matrix  $\mathbf{B}$  as per Eq. (13). In the testing stage, we obtain the corresponding segmentation result as

$$\mathbf{Y}_{\text{test}} = \mathbf{X}_{\text{test}}\mathbf{B} \quad (14)$$

by using the obtained regression coefficient matrix  $\mathbf{B}$ . However, because of the residual matrices and the inherent imperfections in modeling, the values of  $\mathbf{Y}_{\text{test}}$  are not discrete in nature. We do not employ any thresholding at this stage, but rather normalize these values to the range  $[0, 1]$  ( $0 = \text{sky}$ ,  $1 = \text{cloud}$ ). These normalized values provide a probabilistic indication of the “belongingness” of a particular pixel to a specific class (i.e. cloud or sky) instead of a binary decision.

### IV. SKY/CLOUD IMAGE DATABASES

One of the most important pre-requisites for rigorous evaluation of any segmentation algorithm is an annotated dataset of suitable images. Most research groups working on sky/cloud segmentation use their own proprietary images that are not publicly available. In this section, we mention the currently available database containing ground truth images and describe the need for a more comprehensive and consistent database.

#### A. HYTA Database

Till date, the only publicly available sky/cloud image dataset annotated with segmentation masks is the HYTA database by Li et al. [18]. It consists of 32 distinct images with different sky/cloud conditions along with their binary ground-truth segmentation masks. These images are collected from two main sources, namely ground-based sky cameras located

at Beijing and Conghua, China. Additionally, several images found on the Internet from other cameras and locations are included. The ground truth segmentation masks of each of the images are generated using a Voronoi polygonal region generator. The images in HYTA are of various dimensions; the average is  $682 \times 512$  pixels. A few sample images of HYTA along with their ground truth are shown in Fig. 1.



Fig. 1: Sample images from the HYTA database (top row) along with corresponding sky/cloud segmentation ground truth (bottom row).

### B. SWIMSEG Database

Because of the dearth of public databases for sky/cloud images, we created SWIMSEG, the Singapore Whole Sky **I**Maging **S**EGmentation Database.<sup>1</sup> It consists of 1013 images that were captured by a custom-designed ground-based sky imager called WAHRSIS (Wide Angle High Resolution Sky Imaging System), developed and deployed at Nanyang Technological University in Singapore [11]. The imaging system of our ground-based sky camera consists of a Canon EOS Rebel T3i (a.k.a. EOS 600D) camera body and a Sigma 4.5mm F2.8 EX DC HSM Circular Fish-eye Lens with a field of view of 180 degrees.

The images captured by this system suffer from distortions in illumination and color, which need to be corrected. An example is shown in Fig. 2. First, because of a phenomenon called vignetting, the images are comparatively darker at the edges as compared to the center. Vignetting occurs because the incident light rays reach the camera sensor with varying angles and is particularly noticeable for fish-eye lenses. We analyze vignetting in our imaging system using an integrating sphere [11] and perform the necessary correction in the image, the result of which is shown in Fig. 2(b).

Second, the colors in the captured images vary due to the different weather conditions and capturing hours of the images and often do not correspond to the actual colors in the scene. We use a color checkerboard with 18 different color patches as a reference, whose exact  $L^*a^*b^*$  values are specified by the manufacturer. Using the mapping between captured and actual color, we thereby obtain the color-corrected image, which is shown in Fig. 2(c).

Finally, the images from the database are undistorted using the geometric calibration function, which models the effect of the lens. This function relates each pixel of the captured

<sup>1</sup> The SWIMSEG database can be downloaded from <http://vintage.winklerbros.net/swimseg.html>.

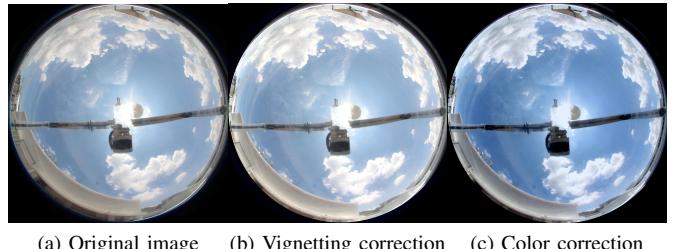


Fig. 2: Image correction w.r.t. illumination and color.

image to the azimuth and elevation angle of the corresponding incident light ray. Using that information, we can project the image onto a hemisphere whose center is the focal point. This is shown in Fig. 3(a). We generate undistorted pictures using a ray tracing approach by considering an imaginary camera (with a standard lens) at center of the hemisphere, which points towards a user defined direction with a given azimuth and an elevation angle. In order to simulate this camera, we consider an output image with dimension  $600 \times 600$ . Each output image corresponds to a viewing angle of  $68^\circ$ . The rays passing through each pixel will intersect the hemisphere and then converge towards the center. The value of a pixel is then equal to the color of the hemisphere at its intersection point. We perform this process at varying azimuth angles for a particular angle of elevation to generate several undistorted images. Figure 3 shows the output image as well as the lines corresponding to its borders and diagonals on the original image.

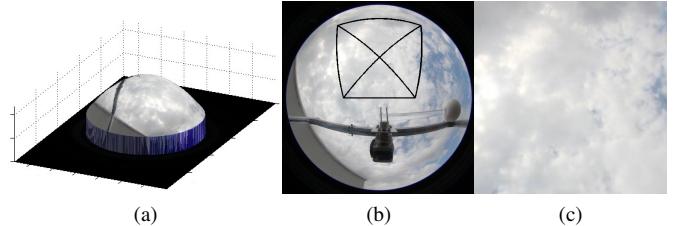


Fig. 3: Generation of undistorted images using ray-tracing. (a) Projection of the image on a hemisphere; (b) Input image with borders and diagonals of the output image; (c) Undistorted output image.

The corresponding sky/cloud segmentation masks of the images were created in consultation with cloud experts from the Singapore Meteorological Services. A few representative sample images from this database are shown in Fig. 4.

SWIMSEG consists of 1013 patches that were handpicked carefully from images taken over the period from October 2013 till July 2015 in Singapore. In the selection process, different parameters (viz. azimuth and elevation angle of sun, time of day, cloud coverage percentage, camera parameters) were considered. The camera elevation angle is set to three distinct angles ( $36^\circ$ ,  $45^\circ$ , and  $60^\circ$ ), whereas the azimuth angle is sampled at various angles around the unit hemisphere. The camera is set to auto mode while capturing pictures, and

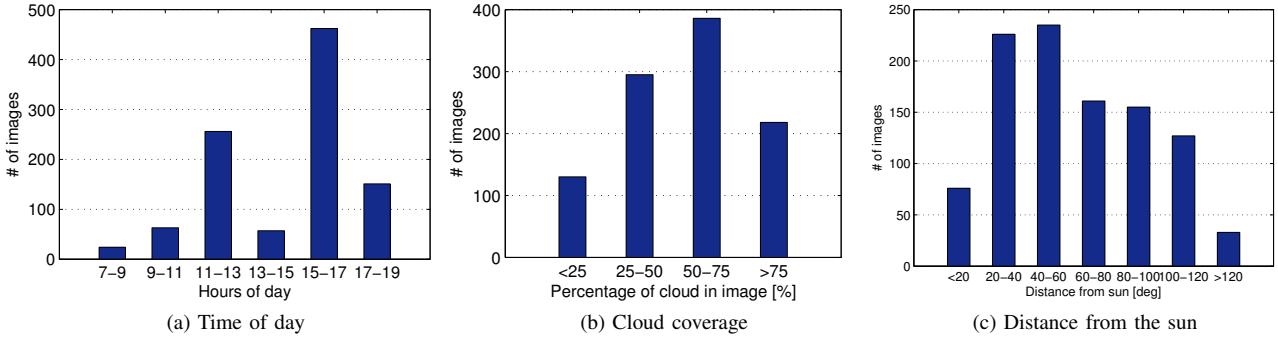


Fig. 5: Distribution of images in SWIMSEG according to various characteristics.

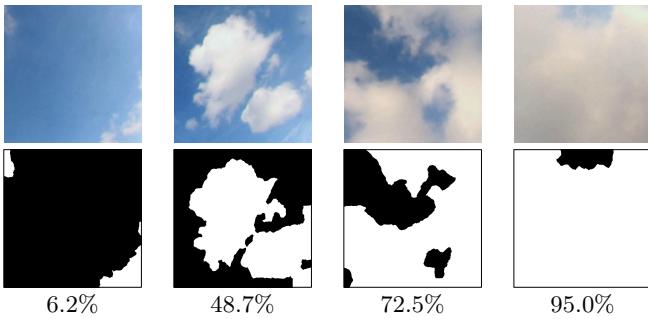


Fig. 4: Sample images from the SWIMSEG database (top row), along with corresponding sky/cloud segmentation ground truth (bottom row) and percentage of cloud coverage.

we observe that the settings tend to remain similar across all images. This is primarily because weather conditions in Singapore remain fairly similar throughout the year.

Fig. 5 characterizes the images in the dataset according to the time of day they were taken, cloud coverage, and distance from the sun. The sun rises and sets at around the same time throughout the year in Singapore, namely at 7am and 7pm local time. The number of clear sky patches is relatively low as compared to moderately-cloudy or overcast conditions. This is because Singapore experiences a tropical climate in which clear-sky conditions are rare. Although most of the image patches are sampled at a certain distance from the sun to avoid saturation effects, we still include a substantial number of images close to sun.

## V. EXPERIMENTAL EVALUATION

### A. Principal Component Analysis (PCA)

We use PCA as described in Section II-B to check the degree of correlation between the 16 color channels, and also to identify the ones that capture most of the variance of the input data.

Figure 6 shows the amount of variance captured by the most important principal components for HYTA and SWIMSEG databases. Table II shows the variance captured by each principal component for the concatenated distributions. In both databases, the first two principal components capture a large

Database	1 <sup>st</sup>	2 <sup>nd</sup>	3 <sup>rd</sup>	4 <sup>th</sup>	5 <sup>th</sup>	6 <sup>th</sup>
HYTA	60.8%	24.8%	7.86%	5.11%	0.84%	0.36%
SWIMSEG	68.44%	17.01%	6.82%	5.82%	1.64%	0.18%

TABLE II: Percentage of variance captured across different principal components for HYTA and SWIMSEG databases.

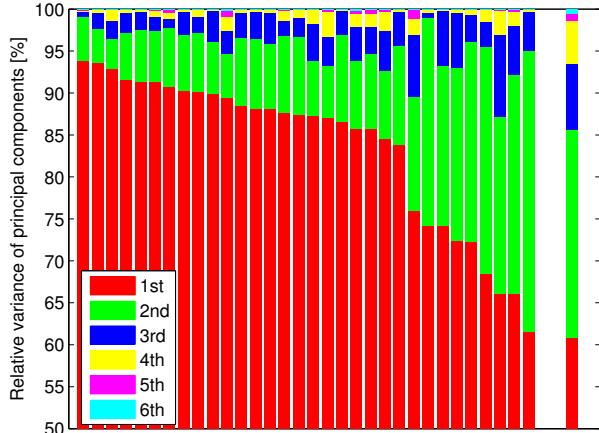
part of the variance across all the images, namely 85.6% and 85.4% of the total variance, respectively.

Fig. 7 shows the bi-plots of 16 color channels for HYTA and SWIMSEG. The bi-plots are consistent across both databases, with only small differences. We can also infer that certain color channel pairs such as  $c_5$  and  $c_{15}$ ,  $c_8$  and  $c_{14}$ , etc. are highly correlated with each other. The most noticeably different channel is  $c_{11}$  ( $a^*$ ), largely because the red/green axis is not very useful for distinguishing sky from cloud and contains mostly noise.

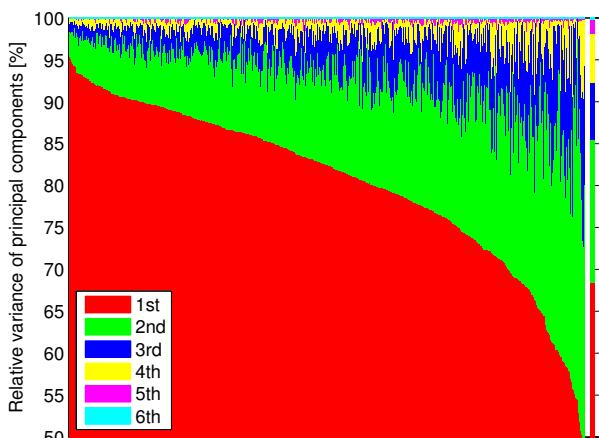
### B. Loading Factors and ROC Curves

In order to determine the most appropriate color channel(s) for effective segmentation of sky/cloud images, we calculate the loading factors, i.e. the projection of the input vectors in the bi-plot onto the 1<sup>st</sup> principal component axis, as described in Section II-B. The channels with the highest loading factors are most discriminatory and thus favorable for segmentation. Furthermore, we use the area between the ROC curve and the random classifier slope to identify the important color channels in the sky/cloud binary image classification task, as described in Section II-C.

Figure 8 shows the results obtained for the loading factors and the area under ROC curve for all color channels of the concatenated distribution in HYTA and SWIMSEG databases. We see a clear trend amongst these values across the different color channels. Red/blue ratio channels  $c_{13,15}$  and saturation channel  $c_5$  rank high. On the other hand, color channels like  $c_4$  and  $c_{11}$  are not discriminatory and thereby rank lower. Also, as could be expected from the bi-plots, the results are again perfectly consistent across both databases (except for  $c_{11}$ ). The color channels with the maximum contribution to the 1<sup>st</sup> principal component are  $c_{13}$ ,  $c_{15}$  and  $c_5$ . Therefore we consider these color channels good candidates for our segmentation problem.



(a) HYTA



(b) SWIMSEG

Fig. 6: Distribution of variance across the principal components for all images in the HYTA and SWIMSEG databases. Each bar represents a single image. The rightmost bar in each plot shows the variance for the concatenation of all images (cf. Table II; best viewed in color).

## VI. SEGMENTATION RESULTS

In the previous Section, we identified suitable candidate color components for segmentation, such as  $c_5$  (saturation) and various blue-red ratios ( $c_{13}$ ,  $c_{14}$ ,  $c_{15}$ ). These color channels are expected to perform better than others. We now evaluate the segmentation results for all 16 color channels, by individually considering them as the discriminatory feature vector  $\mathbf{X}^f \in \mathbb{R}^{mn \times 1}$  for segmentation. We showed previously [22] that there is hardly any performance improvement when using more than 1 color channel in the feature matrix  $\mathbf{X}_f$ .

We perform extensive experiments of our proposed approach on the HYTA and SWIMSEG databases. Each of the images in the two datasets are expressed in the form of  $\hat{\mathbf{X}}_i$ , as shown in Eq. (1). The corresponding feature vector  $\mathbf{X}_i^f$  is computed for both training and testing stages. The

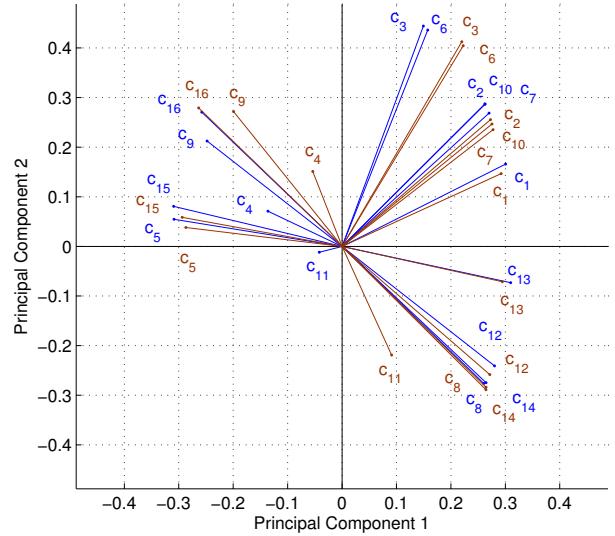


Fig. 7: Bi-plot representation of 16 color channels for HYTA (in blue) and SWIMSEG (in red) databases.

training images are used to calculate the matrix of regression coefficients  $\mathbf{B}$  as per Eq. (13). In our experiments, we use `plsregress`, the MATLAB implementation of the SIMPLS algorithm [25] for this. We set  $p = k = 1$ , as we use a single color channel for evaluation. During the testing phase,  $\mathbf{X}_{\text{test}}$  of the testing image is computed. Using the matrix of regression coefficients  $\mathbf{B}$ , the corresponding  $\mathbf{Y}_{\text{test}}$  is computed.

In the ideal scenario, the value of  $\mathbf{Y}_{\text{test}}$  is either 0 or 1. However, in the actual implementation, the values are continuous in nature. Suppose we denote the obtained values of  $\mathbf{Y}_{\text{test}}$  as  $\{v_1, v_2, \dots, v_{mn}\}$ . Let  $\theta$  be the function that normalizes the obtained values of  $\mathbf{Y}_{\text{test}}$  to the interval  $[0, 1]$ :

$$\theta(v_i) = \frac{v_i - \min v_i}{\max v_i - \min v_i}.$$

We perform *soft* thresholding by normalizing these values in the range  $[0, 1]$ . Therefore, this normalized value indicates the belongingness of a pixel to belong to the cloud (or sky) category.

The objective evaluation is based on the true positive ( $TP$ ), true negative ( $TN$ ), false positive ( $FP$ ), and false negative ( $FN$ ) samples in a binary image. We report Precision, Recall, F-score, and Misclassification rate for the different benchmarking algorithms, which are defined as follows:

$$\begin{aligned} \text{Precision} &= TP / (TP + FP), \\ \text{Recall} &= TP / (TP + FN), \\ \text{F-score} &= \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}, \\ \text{Misclassification rate} &= \frac{FP + FN}{TP + TN + FP + FN}. \end{aligned}$$

### A. Choice of Color Channels

As described earlier, the segmentation performance is dependent on the choice of color channels. In order to understand

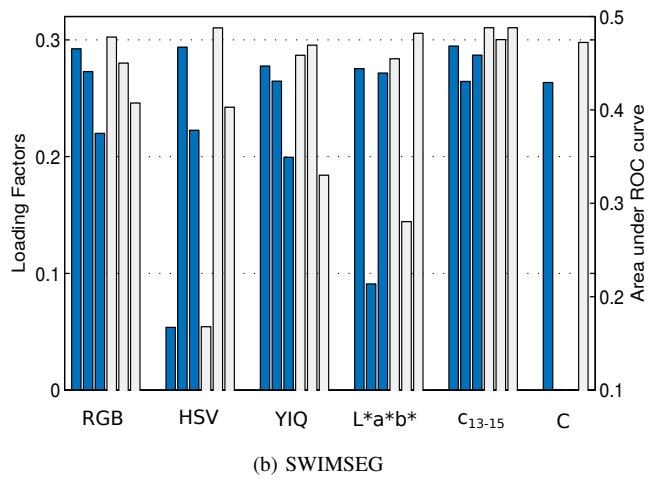
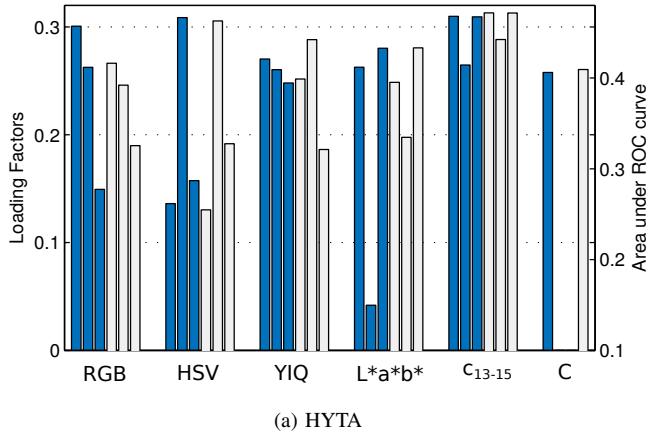


Fig. 8: Selection of color channels for effective sky/cloud image segmentation. We represent the loading factors in blue bars (primary y-axis), and area under the ROC curve in gray bars (secondary y-axis), for HYTA and SWIMSEG databases. The 16 color channels are grouped by color space. Each individual bar represents a single color channel.

the effect of color channels, we use the individual color channels to constitute the feature vector  $\mathbf{X}^f$  in this experiment. We choose training and test set sizes as per Table III.

Database	Training Set	Test Set	Total
HYTA	17	15	32
SWIMSEG	500	513	1013

TABLE III: Number of training and test images for HYTA and SWIMSEG databases.

We compute F-scores (the harmonic mean of precision and recall) to characterize the segmentation performance of each color channel. We perform a 5-fold and 10-fold cross validation in HYTA and SWIMSEG datasets, respectively. In each of these experiments, 1 fold is used for testing and the remaining folds are used in the training stage. Figure 9 shows the results for HYTA and SWIMSEG datasets.

The classification performance for individual color channels follows the same trend in both HYTA and SWIMSEG dataset.

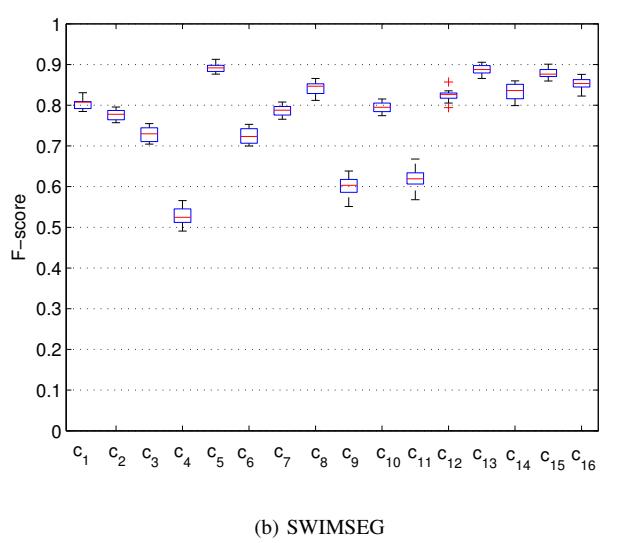
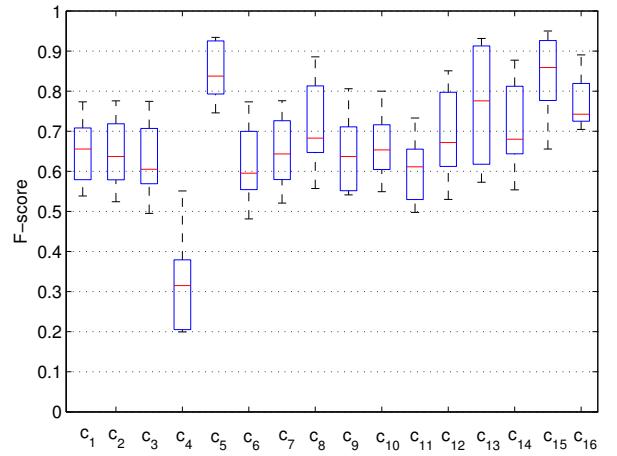


Fig. 9: Binary classification performance for different color channels.

Color channels  $c_5$  and the red-blue ratio channels  $c_{13,15}$  perform best, which is consistent with what we found in Section V. Based on these results, we conduct the subsequent evaluations with  $c_{15}$  as the feature vector  $\mathbf{X}^f$ .

#### B. Effect of Database Characteristics

We study the effect of the different characteristics of the database on the obtained F-score values of our proposed segmentation algorithm by checking whether the obtained segmentation results are biased towards a particular set of image parameters, specifically time of day, cloud coverage, and distance from the sun. We present the distribution of F-score values across these parameters in Fig. 10.

The average F-score is slightly high in the morning and late afternoon, while it drops during mid-day. This is because the images captured during mid-day have a higher tendency to get over-exposed because of the presence of direct sun. On the other hand, the images captured at the morning and

evening hours of the day receive slant sun-rays, and contain more clouds. The variation of the F-score values w.r.t. the percentage of cloud in an image is shown in Fig. 10(b). We observe that the performance is marginally better for higher percentage of cloud coverage in an image. Finally, we note that there is no clear relationship of the distance of the image patch from the sun on the obtained F-score values, which is illustrated in Fig. 10(c).

### C. Benchmarking

We benchmark our proposed approach with current state-of-the-art cloud detection algorithms: Li et al. [18], Souza et al. [16], Long et al. [13], and Mantelli-Neto et al. [17]. We also compare against another popular image segmentation algorithm viz. SLIC superpixels [27].

Li et al. [18] use a Hybrid Thresholding Algorithm (HYTA) that combines both adaptive and fixed threshold approach. Souza et al. [16] use the saturation component of the intensity, hue, and saturation (IHS) color model in defining appropriate thresholds to obtain the final segmentation mask. Long et al. [13] use the ratio of red and blue channels to model the molecular scattering in the atmosphere. Mantelli-Neto et al. [17] use multi-dimensional Euclidean geometric distance and Bayesian approaches to identify sky and cloud pixels.

The popular superpixel segmentation Simple Linear Iterative Clustering (SLIC) divides the input image into a number of superpixels. We combine it with the DBSCAN algorithm [28] to cluster the superpixels and subsequently categorize them into sky/cloud based on the average ratio of red and blue of a particular superpixel.

In addition to these state-of-the-art cloud segmentation algorithms, we also compare our proposed algorithm with other typical machine-learning based methods. We extract discriminative features from image, and use them in a classifier model. In our case, we employ SVM classifier along with various image features.

We use several feature descriptors viz. Local Binary Patterns (LBP) [29], Color Histogram (colorHIST) [30], and Dense Scale Invariant Feature Transform (dSIFT) [31] for extracting discriminatory features from sky/cloud images. Quite recently, Yuan et al. [19] used Bag-Of-Words (BOW) model to model cloud- and non-cloud- satellite images from dSIFT local features. Yuan et al. [19] over segments a satellite image into sub-images using SLIC superpixel generation method. Image descriptors are thereby extracted from these sub-patches using dense SIFT extractor. Subsequently, the BOW model is used to model the cloud- and non-cloud images and the corresponding codebook is generated. A Support Vector Machine (SVM) classifier is finally used for the super-pixel classification into cloud- and non-cloud- parts [19]. In addition to these features, we also use texture descriptors (using Schmid filter [32]) along with BOW and SVM as a benchmarking cloud segmentation algorithm. Furthermore, we also use the raw gray-scale (GRAY) color channel values as discriminatory feature in feature extraction stage. These features are used in training a SVM classifier. In the testing stage, these descriptors are again computed for the test images, and the trained SVM classifier

classifies them to either sky or cloud pixels. Therefore, these different image features together with SVM classifier form a collection of machine-learning based sky/cloud classification algorithms.

In our proposed approach, we assign a membership value for the cloud category to each pixel. However, as the available ground-truth masks are binary, we convert our probabilistic map into a binary map by simple thresholding in order to perform the objective evaluations. The threshold value can be any value in the closed interval [0,1]; a lower value indicates a higher tendency for a pixel to classify as *cloud* and a higher value indicates a higher tendency to classify as *cloud* pixel. Figure 11 shows the ROC curve for our proposed algorithm by changing the threshold value in the closed interval [0,1] for both HYTA and SWIMSEG databases. 20 trials were performed for each threshold value, with training and testing sets of images chosen randomly for each trial. The subsequent results were obtained with a threshold of 0.5.

Figure 12 shows a few sample images from the HYTA and SWIMSEG databases together with the segmentation ground truth and the corresponding output images obtained by our proposed approach and other existing methods. Most of the approaches work quite well when there is a clear contrast between sky and cloud regions, as is often the case for the HYTA database. The method of Mantelli-Neto et al. [17] fails on images that contain non-uniform illumination of cloud mass from sunlight. The detection of thin clouds is the most problematic: Long et al. [13] and SLIC+DBSCAN methods classify the entire image patch as cloud in such cases. The method of Li et al. [18] fails in cases where the boundary between sky and cloud is ill-defined. We observe from the last two examples in Fig. 12 that Li et al. over-segments the corresponding input images. Except for Souza et al. [16], most of the approaches fail for images of the SWIMSEG dataset with generally less well-defined sky/cloud separation. Similarly, the other machine-learning based methods viz. GRAY+SVM and dSIFT+BOW+SVM [19] perform poorly for the wide range of images. This is because the Bag-of-words model obtained for cloud- and non-cloud- superpixels fails to discriminate between the superpixels in the test images. The results obtained by our PLS-based approach show that it works well across diverse illumination conditions and for both databases.

The simulations were conducted in Matlab on a 64-bit Ubuntu 14.04 LTS workstation, Intel i5 CPU @2.67GHz. In terms of computation time, our proposed method takes an average of 1.31s and 1.89s for a single image from the HYTA and SWIMSEG databases, respectively. In the training stage, HYTA requires 21.5s and SWIMSEG required 1018.6s for computing the regression coefficient matrix  $\mathbf{B}$ . The computation times of other benchmarking algorithm are provided in Table IV.

Table IV provides an objective evaluation of our proposed approach with other state-of-the-art algorithms for all images in the testing set of the two databases. Existing algorithms require extensive tuning using various manually-defined thresholds, and achieve either high precision or high recall in our evaluation, but not both. Long et al. and Mantelli-Neto et

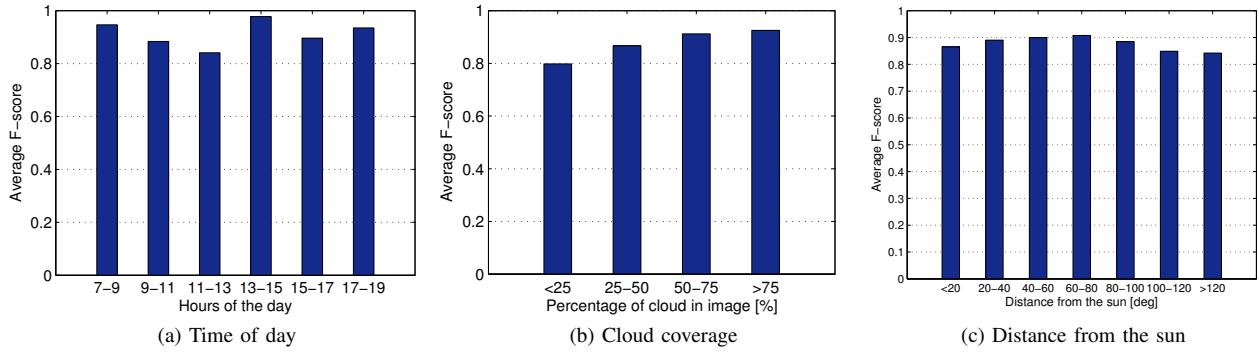


Fig. 10: Distribution of average F-score value as a function of various characteristics of the images in SWIMSEG.

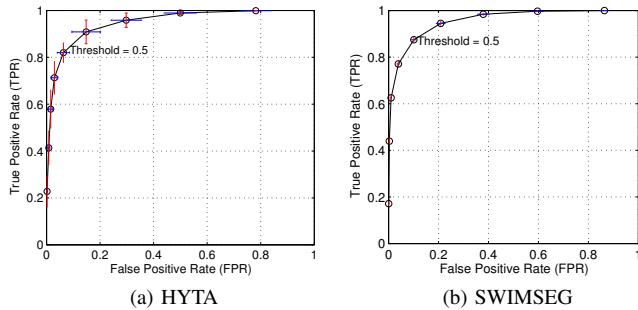


Fig. 11: ROC curves of our proposed algorithm, including error bars in X- and Y-directions.

al. have high recall values. However, they fare poorly in the corresponding precision values. Souza et al. on the other hand has a high precision value, but low recall. These existing methods often suffer from under- or over-identification of cloud pixels as they are threshold-based, and a single threshold may not work well for all types of sky/cloud images. Our proposed approach performs best for both HYTA and SWIMSEG databases on the basis of F-scores, highlighting the effectiveness of PLS-based learning method.

## VII. DISCUSSION

There are three primary advantages of our proposed approach compared to other cloud detection algorithms.

First, our cloud segmentation framework is not based on any pre-defined assumptions about color spaces and does not place any restrictions on the type of input images. We systematically compare different color channels and identify the most suitable ones. We also explain the reason for their better performance based on rigorous statistical evaluations in two datasets.

Second, many existing cloud segmentation algorithms rely on a set of thresholds, conditions, and/or parameters that are manually defined for a particular sky/cloud image database. Our proposed cloud segmentation approach is entirely learning-based and thus provides a systematic solution to training for a given database.

Third, conventional algorithms provide a binary output image from the input sky/cloud image. Although these binary

images are informative in most cases, they lack flexibility and robustness. We have no indication of the effectiveness of the thresholding of the input image. In reality, because of the nature of clouds and cloud images, it is undoubtedly better to employ a *soft* thresholding approach. Our proposed approach achieves a probabilistic classification of cloud pixels. This is very informative as it provides a general sense of *belongingness* of pixels to the cloud class. However, as only binary ground-truth images are available, we convert these probability maps into binary images for performing a quantitative evaluation of our algorithm. In our future work, we plan to extend this analysis by creating *probabilistic ground truth* images, where the ground truth of the input images is generated by aggregating annotations from multiple experts.

Finally, from the extensive experiments on the two datasets (cf. Table IV), we observe that the performance with images from the SWIMSEG database is generally better than for HYTA, even though the behavior of color channels is similar in both (i.e. the same color channels do well for both SWIMSEG and HYTA). We believe this is because the images in SWIMSEG were captured with a camera that has been calibrated for color, illumination, and geometry [11].

Naturally, many challenges remain, for example: How do different weather conditions affect the classification performance? The weather in Singapore is relatively constant in terms of temperature and humidity, with little variation throughout the year. Usually it is either partly cloudy, or rainy (in which case the sky will be completely overcast, making segmentation unnecessary). As a result, our database is not suitable for investigating this question. Completely overcast conditions can be dealt with by simple pre-processing of the image before segmentation, e.g. by calculating the number of clusters, as we have done in our previous work [33].

## VIII. CONCLUSIONS

We have presented a systematic analysis of color spaces and components, and proposed a probabilistic approach using PLS-based regression for the segmentation of ground-based sky/cloud images. Our approach is entirely learning-based and does not require any manually-defined thresholds, conditions, or parameters at any stage of the algorithm. We also release an extensive sky/cloud image database captured with a calibrated

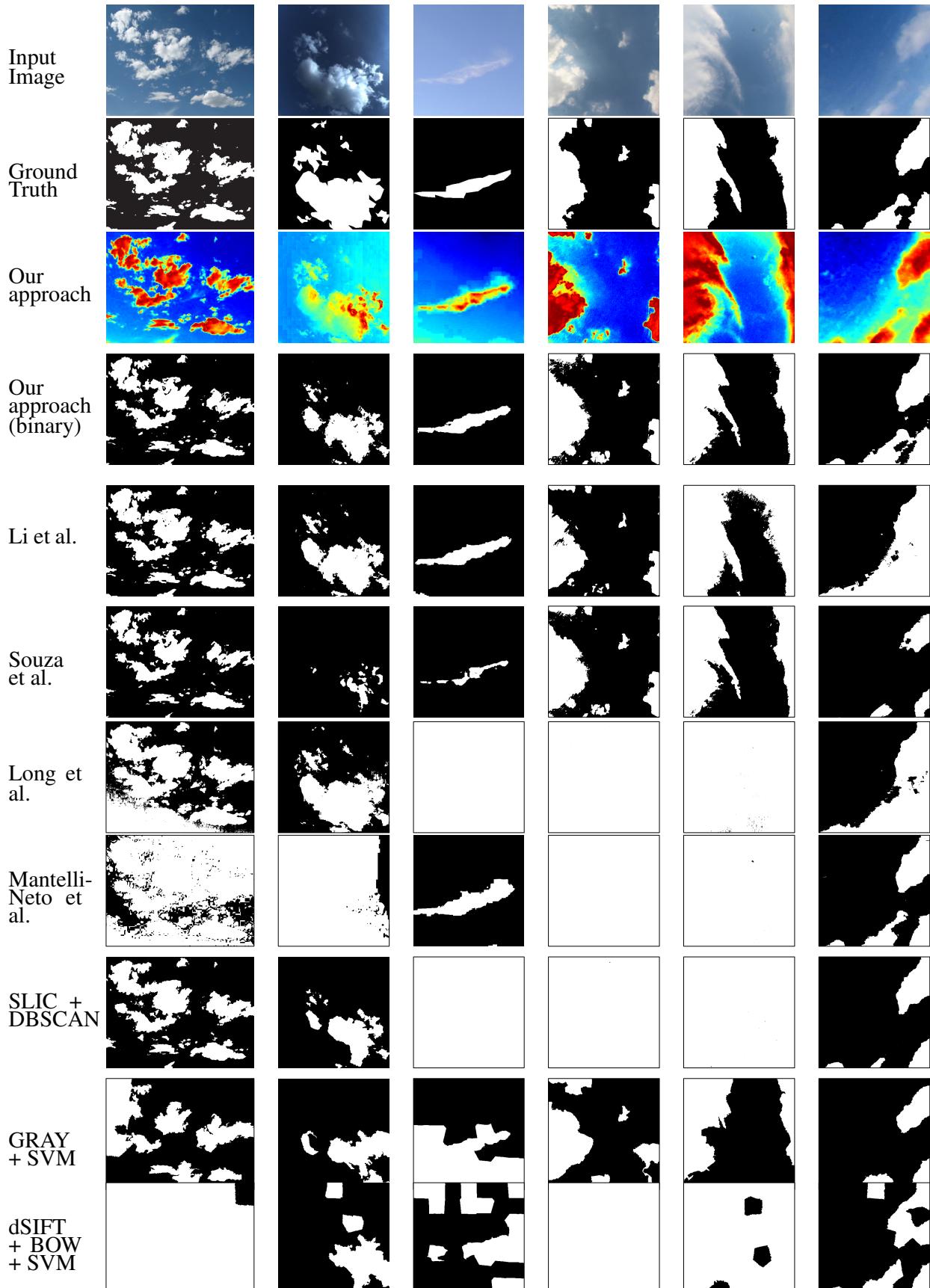


Fig. 12: Sample images from HYTA (left 3 columns) and SWIMSEG (right 3 columns) databases, along with the corresponding ground truth and the sky/cloud segmentation obtained using various methods.

	Methods	Precision	Recall	F-score	Misclassification rate	Time [s]
<b>HYTA</b>	Li et al.	0.89	0.82	0.81	0.12	1.34
	Souza et al.	0.88	0.67	0.65	0.15	1.33
	Long et al.	0.64	<b>0.97</b>	0.71	0.26	1.22
	Mantelli-Neto et al.	0.54	<b>0.97</b>	0.63	0.44	1.43
	SLIC + DBSCAN	0.65	0.83	0.60	0.32	5.10
	GRAY + SVM	0.89	0.58	0.67	0.31	2.44
	LBP + SVM	0.80	0.65	0.72	0.22	4.29
	ColorHIST + SVM	0.75	0.74	0.75	0.21	2.47
	dSIFT + BOW + SVM	0.61	0.66	0.62	0.38	4.91
	Texture + BOW + SVM	0.82	0.62	0.69	0.25	2.60
	Proposed Method	<b>0.94</b>	0.80	<b>0.85</b>	<b>0.10</b>	1.31
<b>SWIMSEG</b>	Li et al.	0.90	0.86	0.88	0.11	2.06
	Souza et al.	0.95	0.76	0.81	0.14	2.04
	Long et al.	0.71	<b>0.98</b>	0.80	0.23	1.83
	Mantelli-Neto et al.	0.70	0.97	0.79	0.24	2.16
	SLIC + DBSCAN	0.72	0.79	0.65	0.33	5.39
	GRAY + SVM	0.87	0.56	0.64	0.33	2.61
	LBP + SVM	0.62	0.65	0.63	0.36	4.73
	ColorHIST + SVM	0.81	0.64	0.66	0.31	2.63
	dSIFT + BOW + SVM	0.65	0.88	0.72	0.28	5.04
	Texture + BOW + SVM	0.82	0.71	0.70	0.31	2.74
	Proposed Method	<b>0.92</b>	0.90	<b>0.90</b>	<b>0.09</b>	1.89

TABLE IV: Performance evaluation using binary ground truth of HYTA and SWIMSEG databases. The best performance according to each criterion is indicated in bold. All experiments are evaluated on the same set of testing images from the HYTA and SWIMSEG databases. We also list the computation time for a single image for all algorithms.

ground-based camera that has been annotated with ground-truth segmentation masks.

Our future work will include the annotation of a database with probabilistic ground-truth segmentation maps as well as the extension of this method to High-Dynamic-Range (HDR) images. Going beyond segmentation, it is also important to classify clouds into different types [33] or estimate cloud altitude and movement [34], which are both part of our current research.

#### ACKNOWLEDGMENT

This work is supported by a grant from Singapore's Defence Science & Technology Agency (DSTA).

#### REFERENCES

- [1] C. Papin, P. Bouthemy, and G. Rochard, "Unsupervised segmentation of low clouds from infrared METEOSAT images based on a contextual spatio-temporal labeling approach," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, no. 1, pp. 104–114, Jan. 2002.
- [2] M. Mahrooghy, N. H. Younan, V. G. Anantharaj, J. Aanstoos, and S. Yarahmadian, "On the use of a cluster ensemble cloud classification technique in satellite precipitation estimation," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 5, no. 5, pp. 1356–1363, Oct. 2012.
- [3] J. M. Weiss, S. A. Christopher, and R. M. Welch, "Automatic contrail detection and segmentation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 36, no. 5, pp. 1609–1619, Sept. 1998.
- [4] C. I. Christodoulou, S. C. Michaelides, and C. S. Pattichis, "Multifeature texture analysis for the classification of clouds in satellite imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, no. 11, pp. 2662–2668, Nov. 2003.
- [5] F. Yuan, Y. H. Lee, and Y. S. Meng, "Comparison of cloud models for propagation studies in Ka-band satellite applications," in *Proc. IEEE International Symposium on Antennas and Propagation*, 2014, pp. 383–384.
- [6] 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, pp. 259–260.
- [7] T. Shiraishi, T. Motohka, R. B. Thapa, M. Watanabe, and M. Shimada, "Comparative assessment of supervised classifiers for land use-land cover classification in a tropical region using time-series PALSAR mosaic data," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 4, pp. 1186–1199, April 2014.
- [8] I.-H. Lee and M. T. Mahmood, "Robust registration of cloudy satellite images using two-step segmentation," *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 5, pp. 1121–1125, May 2015.
- [9] Q. Zhang and C. Xiao, "Cloud detection of RGB color aerial photographs by progressive refinement scheme," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 11, pp. 7264–7275, Nov. 2014.
- [10] Y. Chen, L. Cheng, M. Li, J. Wang, L. Tong, and K. Yang, "Multiscale grid method for detection and reconstruction of building roofs from airborne LiDAR data," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 10, pp. 4081–4094, Oct. 2014.
- [11] S. Dev, F. Savoy, Y. H. Lee, and S. Winkler, "WAHRIS: A low-cost, high-resolution whole sky imager with near-infrared capabilities," in *Proc. IS&T/SPIE Infrared Imaging Systems*, 2014.
- [12] M. E. Farmer and A. K. Jain, "A wrapper-based approach to image segmentation and classification," *IEEE Transactions on Image Processing*, vol. 14, no. 12, pp. 2060–2072, Dec. 2005.
- [13] C. N. Long, J. M. Sabburg, J. Calbó, and D. Pages, "Retrieving cloud characteristics from ground-based daytime color all-sky images," *Journal of Atmospheric and Oceanic Technology*, vol. 23, no. 5, pp. 633–652, May 2006.
- [14] A. Heinle, A. Macke, and A. Srivastav, "Automatic cloud classification of whole sky images," *Atmospheric Measurement Techniques*, vol. 3, no. 3, pp. 557–567, May 2010.
- [15] S. Liu, L. Zhang, Z. Zhang, C. Wang, and B. Xiao, "Automatic cloud detection for all-sky images using superpixel segmentation," *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 2, pp. 354–358, Feb. 2015.
- [16] M. P. Souza-Echer, E. B. Pereira, L. S. Bins, and M. A. R. Andrade, "A simple method for the assessment of the cloud cover state in high-latitude regions by a ground-based digital camera," *Journal of Atmospheric and Oceanic Technology*, vol. 3, no. 5, pp. 437–447, March 2006.
- [17] S. L. Mantelli-Neto, A. von Wangenheim, E. B. Pereira, and E. Coimunello, "The use of Euclidean Geometric Distance on RGB color space for the classification of sky and cloud patterns," *Journal of Atmospheric and Oceanic Technology*, vol. 27, no. 9, pp. 1504–1517, Sept. 2010.

- [18] 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, Oct. 2011.
- [19] Y. Yuan and X. Hu, "Bag-of-words and object-based classification for cloud extraction from satellite imagery," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, no. 8, pp. 4197–4205, Aug. 2015.
- [20] 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, Jan. 2008.
- [21] S. Liu, Z. Zhang, B. Xiao, and X. Cao, "Ground-based cloud detection using automatic graph cut," *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 6, pp. 1342–1346, June 2015.
- [22] S. Dev, Y. H. Lee, and S. Winkler, "Systematic study of color spaces and components for the segmentation of sky/cloud images," in *Proc. IEEE International Conference on Image Processing (ICIP)*, 2014, pp. 5102–5106.
- [23] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861–874, June 2006.
- [24] H. Mamitsuka, "Selecting features in microarray classification using ROC curves," *Pattern Recognition*, vol. 39, no. 12, pp. 2393–2404, Dec. 2006.
- [25] S. de Jong, "SIMPLS: An alternative approach to partial least squares regression," *Chemometrics and Intelligent Laboratory Systems*, vol. 18, no. 3, pp. 251–263, March 1993.
- [26] R. Manne, "Analysis of two partial-least-squares algorithms for multivariate calibration," *Chemometrics and Intelligent Laboratory Systems*, vol. 2, no. 1-3, pp. 187–197, Aug. 1987.
- [27] R. Achanta, A. Shajii, K. Smith, A. Lucchi, P. Fua, and S. Süstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 11, pp. 2274–2282, Nov. 2012.
- [28] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Proc. 2nd International Conference on Knowledge Discovery and Data Mining*, 1996, pp. 226–231.
- [29] T. Ojala, M. Pietikainen, and T. Maenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971–987, July 2002.
- [30] C. L. Novak and S. A. Shafer, "Anatomy of a color histogram," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1992, pp. 599–605.
- [31] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, Nov. 2004.
- [32] C. Schmid, "Constructing models for content-based image retrieval," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2001, pp. 39–45.
- [33] S. Dev, Y. H. Lee, and S. Winkler, "Multi-level semantic labeling of sky/cloud images," in *Proc. IEEE International Conference on Image Processing (ICIP)*, 2015.
- [34] F. M. Savoy, J. Lemaitre, S. Dev, Y. H. Lee, and S. Winkler, "Cloud-base height estimation using high-resolution whole sky imagers," in *Proc. IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2015, pp. 1622–1625.



**Yee Hui Lee** (S'96-M'02-SM'11) received the B.Eng. (Hons.) and M.Eng. degrees from the School of Electrical and Electronics Engineering at Nanyang Technological University, Singapore, in 1996 and 1998, respectively, and the Ph.D. degree from the University of York, UK, in 2002.

Dr. Lee is currently Associate Professor and Assistant Chair (Students) at the School of Electrical and Electronic Engineering, Nanyang Technological University, where she has been a faculty member since 2002. Her interests are channel characterization, rain propagation, antenna design, electromagnetic bandgap structures, and evolutionary techniques.



**Stefan Winkler** is Distinguished Scientist and Director of the Video & Analytics Program at the Advanced Digital Sciences Center (ADSC), a joint research center between A\*STAR and the University of Illinois. Prior to that, he co-founded Genista, worked for Silicon Valley companies, and held faculty positions at the National University of Singapore and the University of Lausanne, Switzerland.

Dr. Winkler has a Ph.D. degree from the École Polytechnique Fédérale de Lausanne (EPFL), Switzerland, and an M.Eng./B.Eng. degree from the University of Technology Vienna, Austria. He has published over 100 papers and the book "Digital Video Quality" (Wiley). He is an Associate Editor of the IEEE Transactions on Image Processing, a member of the IVMSP Technical Committee of the IEEE Signal Processing Society, and Chair of the IEEE Singapore Signal Processing Chapter. His research interests include video processing, computer vision, perception, and human-computer interaction.



**Soumyabrata Dev** (S'09) graduated summa cum laude from National Institute of Technology Silchar, India with a B.Tech. in Electronics and Communication Engineering in 2010. Subsequently, he worked in Ericsson as a network engineer from 2010 to 2012.

Currently, he is pursuing a Ph.D. degree in the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. From Aug-Dec 2015, he was a visiting student at Audio-visual Communication Laboratory (LCAV), École Polytechnique Fédérale de Lausanne (EPFL), Switzerland. His research interests include remote sensing, statistical image processing and machine learning.