

# Simultaneous Detection and Removal of High Altitude Clouds from an Image

Tushar Sandhan and Jin Young Choi

Department of ECE, ASRI, Seoul National University, South Korea

tushar@snu.ac.kr, jychoi@snu.ac.kr

## Abstract

Interestingly, shape of the high-altitude clouds serves as a beacon for weather forecasting, so its detection is of vital importance. Besides these clouds often cause hindrance in an endeavor of satellites to inspect our world. Even thin clouds produce the undesired superposition of visual information, whose decomposition into the clear background and cloudy layer using a single satellite image is a highly ill-posed problem. In this work, we derive sophisticated image priors by thoroughly analyzing the properties of high-altitude clouds and geological images; and formulate a non-convex optimization scheme, which simultaneously detects and removes the clouds within a few seconds. Experimental results on real world RGB images demonstrate that the proposed method outperforms the other competitive methods by retaining the comprehensive background details and producing the precise shape of the cloudy layer.

## 1. Introduction

Artificial satellites are the earth watchers! They inspect the world through electronic eyes from hundreds of kilometers above surface of the earth and beam the topological, geological and climatic information to a base station. However high altitude clouds, which cover around 70% area of the tropical regions [2], often cause hindrance in their endeavor. Clouds camouflage the vital image details by altering their color and brightness, thereby resulting in blurring and contrast reduction among the earthly objects. Thus cloud cover causes significant loss of satellite visual data [12, 29, 57, 74], which creates difficulty for cartography, water body identification [76], land-landscape ecological change detection [32], archeology, geological scene analysis [8, 19] and agricultural assessment [61].

Furthermore the shape of these clouds guides the weather forecasting [3, 43, 64, 68], for example, streak like rippling cloud patterns indicate the high-altitude climate instability and signals the approach of poorer weather [3, 43]. So the removal as well as detection of cloudy layer from an aerial image (e.g. Figure 1) has a vital importance.

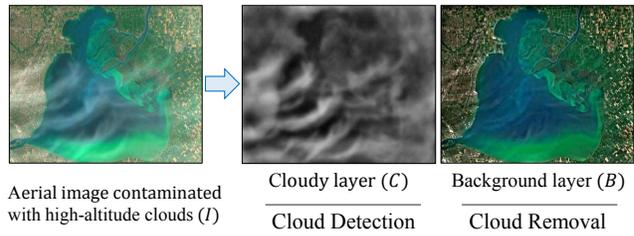


Figure 1: Cloud detection and removal using our method.

Cloud removal is essentially a background reconstruction whereas cloud detection is the background removal process, so they are not mutually exclusive; and the prior work on this affair can be grouped into two categories: (i) multimodal approaches use various modes of input data like multispectral or multitemporal images; (ii) unimodal methods rely on a single type of input data or an optical image.

**Multimodal.** Multispectral images are captured by multiple sensors, each sensitive to a selective wavelength which acquires multiple interpretations of the same scene at a given time instant using diverse spectral bands [25]. This type of methods remove clouds by fusing complementary information from different spectral images [53, 54, 62, 75]. Multitemporal approaches [11, 70, 79, 83] use both temporal and spatial coherence [38], so they have better ability to cope with thick and vast clouds. They assume that the land cover change insignificantly over a short period [25], so a long revisiting period of satellite may become a bottleneck in their performance [73]. Use of both multispectral and multitemporal data helps to improve the cloud removal performance [21, 35, 46, 59] but suffers in cloud detection [69] due to the presence of a cloud variability in temporal images. Clouds appear in almost all bands of multispectral images [69]; for example, high-altitude clouds restrict passage of the most of infrared light [39] thereby creating a void in infrared band which in turn attributed to the presence of a cloud in the infrared image; so cloud removal using multimodal input data is still a challenging issue. Moreover multimodal data are expensive, prone to a mode failure and not ubiquitous, which puts limit on their wide applicability.

**Unimodal.** Missing pixel interpolation [58, 72], image inpainting [26, 42, 44, 80] and image enhancement or noise removal [10, 31, 51, 52, 84] strategies are commonly used for a single image thin cloud removal. Missing regions are synthesized by propagating the nearby texture, color and geometrical traits inside those regions. Utilizing multiple modes, the regions can be filled with the retrieved spectrogeometrical information [42]. Efficiency of these region substitution based methods relies on precise image registration, hue adjustment and presence of smooth or clutter-less background in aerial images. They yield visually plausible results but the absence of restoring information of the cloud contaminated pixels makes them inapt for inferential analysis in the remote sensing [38]. Image-based atmospheric artifacts correction methods like haze or fog removal [18, 27, 41, 82] treat all pixels equally and can effectively remove the thin clouds spread over an entire image. Thus they assume uniform distribution of haze and noise over the image plane. However high-altitude aerial scanners have large angular view. As variety of clouds may simultaneously exist in an image, it invalidates their assumptions. The clouds differ from haze, fog and smoke, mainly in texture, size, shape, unevenness and concentration of atmospheric particles [40, 78]. Alternatively model based methods [37] parametrize the clouds and ground structures which can subsume the cloud peculiarities but they require a lot of input data for refining the model.

**Detection.** Cloud mask detection is essentially an image segmentation problem [81], whereas precise cloud layer (*i.e.* including cloud feathery structure) recovery is more challenging task as aerial images have both bright non-cloud regions and semitransparent (less bright) cloud pixels. The recovery of accurate underlying cloud layer empowers automatic cloud classification methods, which assist in regional climate analysis by inspecting the cloud appearance [43, 68]. Several methods have been developed just for cloud detection using multispectral [20, 60], special apparatus [22] or a single color image [45, 81].

**Contributions.** Thin transparent cloud detection is more challenging than removal task whereas thick translucent cloud removal is more arduous than its detection. So to address the demanding problem of simultaneous cloud detection and removal from a single image; we firstly, analyze the comprehensive properties of high-altitude clouds and geological images; secondly, derive sophisticated image priors (*e.g.* sparse gradient prior with online estimation of image statistical properties, weighted map prior engrafted with cloud properties) and; thirdly, embed all this domain knowledge in a unified non-convex optimization framework. To the best of our knowledge, the proposed approach is the first optimization based scheme for simultaneous cloud detection and removal. We have constructed a new real world aerial image database to extensively verify its effectiveness.

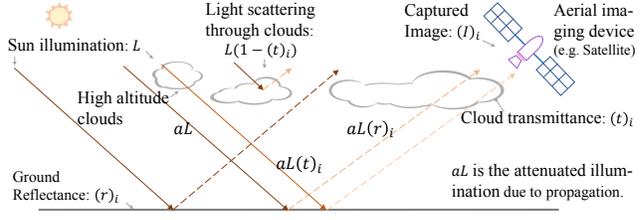


Figure 2: Aerial image formation and degradation model.

## 2. Our Approach

Considering the physical model for cloud distortion process [47] as shown in Figure 2, an aerial image  $I$  having total  $N$  pixels can be expressed at each pixel  $i$  as,

$$(I)_i = a(r)_i(t)_iL + (1 - (t)_i)L, \quad (1)$$

where  $L$  is the intensity of incident solar radiation having attenuation coefficient  $a$ , the ground object reflectance and cloud transmittance are given as  $r$  and  $t$  respectively ( $0 \leq a, r, t \leq 1$ ). The signal reflected by earthly background  $a(r)_iL$ , subsequently passes through cloud and becomes  $a(r)_i(t)_iL$ . High-altitude clouds have low thickness (sec.2.2), so  $t$  is high. Thus we can simplify the model (1) as,  $I=B+C$ , by considering first term as an earthly background layer  $B$  and second as a cloudy layer  $C$ . This layer separation problem is highly ill-posed, as two unknowns  $B$  and  $C$  each of  $N$  pixels need to be discovered from a single image  $I$  *i.e.* unknowns are twice the measurements. We address this ill-posed problem by enforcing the different priors on each latent layer in an optimization framework.

### 2.1. Aerial Image Gradient Statistics

The research in natural image statistics indicates that typical real-world images obey sparse spatial gradient distributions [34, 65]. Similarly, we design our prior on background layer  $B$  as sparse gradients, but improve further by tuning it to the detailed statistics of aerial image gradients. Figure 3 shows typical cloud free aerial images (layer  $B$ ), a high altitude (Cirrus [56]) cloud image (layer  $C$ ) and their respective gradient distributions. The distribution for  $B$  can be modeled as a generalized Laplace distribution like,

$$P(B) = \prod_{i \in \mathbb{I}} \frac{1}{w_1} \exp \left( -\frac{1}{\sigma_1} \sum_{\partial_j \in \mathbb{J}_B} |(\partial_j * B)_i|^\alpha \right), \quad (2)$$

where  $i$  is the pixel index  $\in \mathbb{I} = \{1, \dots, N\}$ ,  $*$  is the convolution operator,  $w_1$  is a normalizing weight,  $\sigma_1$  is a scale parameter, and  $\partial_j$  denotes the derivative filters belonging to the set  $\mathbb{J}_B$ . We used  $\mathbb{J}_B = \{[1, -1], [1, -1]^T\}$  in our implementation. The power  $\alpha$  is a parameter within  $[0.0, 1.0]$ . If we let  $\alpha=1$  then the energy term in (2) will reduce to the

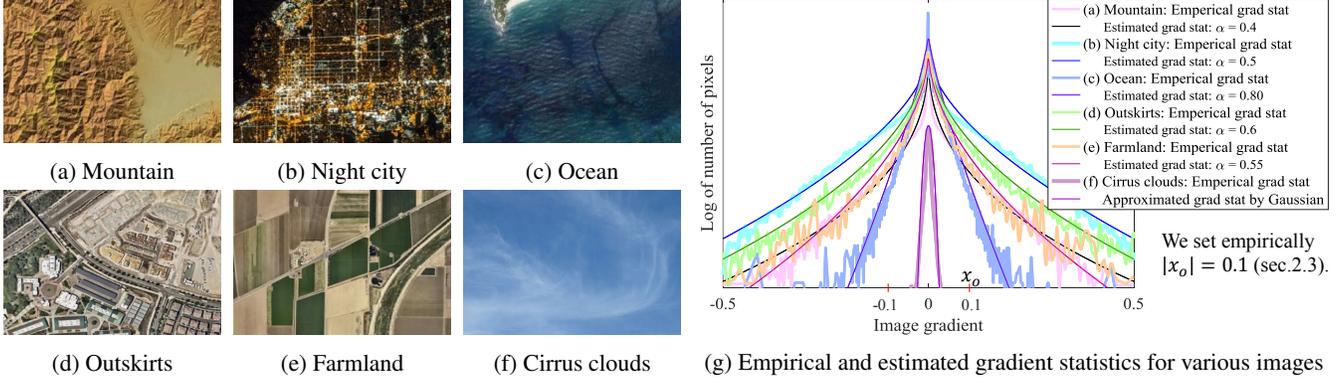


Figure 3: Aerial image statistics: (a) to (e) show typical geological images scanned by satellites; (f) portrays high-altitude (Cirrus [56]) clouds captured from the ground level; their empirical and estimated (via (7)) gradient statistics are given in (g).

$\ell_1$ -norm of gradients (*a.k.a.* total variation (TV) regularization [9, 14]), which is convex function and convenient to optimize using off-the-shelf solvers [4, 7]. However empirically it is evident that the gradient statistics for various kinds of aerial images (few are shown in Figure 3a to 3e) obeys strictly hyper-Laplacian ( $\alpha < 1$ ) distribution and there is no unique  $\alpha$  for all aerial images. Flat terrains (*e.g.* desert, sea, grass land *etc.*) demand high  $\alpha \in ([0.8, 0.9])$ , where as cluttered regions (*e.g.* metropolis, mountain *etc.*) urge low  $\alpha \in ([0.5, 0.6])$ . So instead of specifying  $\alpha$  a priori, it is necessary to infer it from a given data. But initially the layer  $B$  is latent and we only have access to input image  $I$ . How can we infer gradient statistical properties ( $\alpha$ ) for  $B$  from  $I$ ? (see the section 2.3).

## 2.2. High Altitude Cloud Properties

High altitude clouds (having three main genera: Cirrus, Cirrocumulus and Cirrostratus) are composed of tiny ice crystals [56]. So they reflect around 9% of the incoming sunlight and prevent almost 50% of the outgoing infrared radiation from earth [39], which creates obstruction in an aerial imaging. They often produce hair-like filaments or wispy strands stretched across several thousand kilometers in the sky, but only a few kilometers deep [6, 17]. So these clouds appear translucent, pale, thin and are furnished with milky sheen or striated sheet [15, 56].

The cloud layer  $C$  does not have sharp or high magnitude gradients. Its gradient distribution falls quickly and thereby rendering very short tail, which can be modeled by narrow Gaussian distribution as in Figure 3g, *i.e.* with a small standard deviation  $\sigma_2$  the Gaussian falls quickly and given as,

$$P(C) = \prod_{i \in \mathbb{1}} \frac{1}{w_2} \exp \left( -\frac{1}{\sigma_2^2} \sum_{\partial_j \in \mathbb{J}_C} \|(\partial_j * C)_i\|^2 \right), \quad (3)$$

where  $w_2$  is a normalizing weight and we have used second order derivative filters in set  $\mathbb{J}_C = \{[1, -2, 1], [1, -2, 1]^T\}$ .

Above properties also indicate that the cloud layer  $C$  is not uniform and may not span the entire image plane. So the construction of spatial cloud or non-cloud confidence map will help us to boost the layer separation process by imposing additional prior on one of the latent layers. Cloud regions show low saturation component in HSV color space [63]. It is statistically observed that the cloud pixels distribute linearly and closely parallel to the cross diagonal of RGB cube [36, 49]. Let  $(I_{rgb})_i$  be the RGB color triplet (3 dimensional vector) at pixel  $i$ ,  $(\overline{I_{rgb}})_i$  be the mean of its RGB components; and  $S$  be saturation of  $I$  then we define the cloud confidence map for each of its  $N$  pixels as,

$$(M_C)_i = \frac{1}{w_3} \exp \left( -\sigma_3 (S)_i - \sigma_4 \left( \| (I_{rgb})_i \|^2 - (\overline{I_{rgb}})_i^2 \right) \right), \quad (4)$$

where  $\sigma_3, \sigma_4$  are positive scalars and  $w_3$  is normalizing weight such that  $(M_C)_i \leq 1$ . Whereas non-cloud or background confidence map is  $(M_B)_i = 1 - (M_C)_i$ . So while separating the layers, we impose the soft penalty for intrusion of  $C$  in non-cloud regions as,

$$\frac{\gamma}{2} \|M_B \odot C\|^2, \quad (5)$$

where  $\gamma$  is a scalar parameter controlling the importance of penalty term and  $\odot$  denotes element-wise multiplication.

## 2.3. Estimating Gradient Statistics for Latent Layer

As  $C$  lacks high magnitude gradients and has very short tail gradient distribution (sec. 2.2), we can infer statistics for  $B$  (*i.e.*  $\alpha$ ) from  $I$  by considering only high magnitude gradients or outlying tail regions ( $> |x_o|$ ) as shown in Figure 3g. Let  $|\mathbb{J}_B|$  be the cardinality of set  $\mathbb{J}_B$ ,

$$x = \frac{1}{|\mathbb{J}_B|} \sum_{\partial_j \in \mathbb{J}_B} (\partial_j * I), \quad (6)$$

$P(x)$  be the histogram of (6) and considering model (shifting) parameter as  $b$ , we estimate  $\alpha = \hat{\alpha}$  by solving the con-

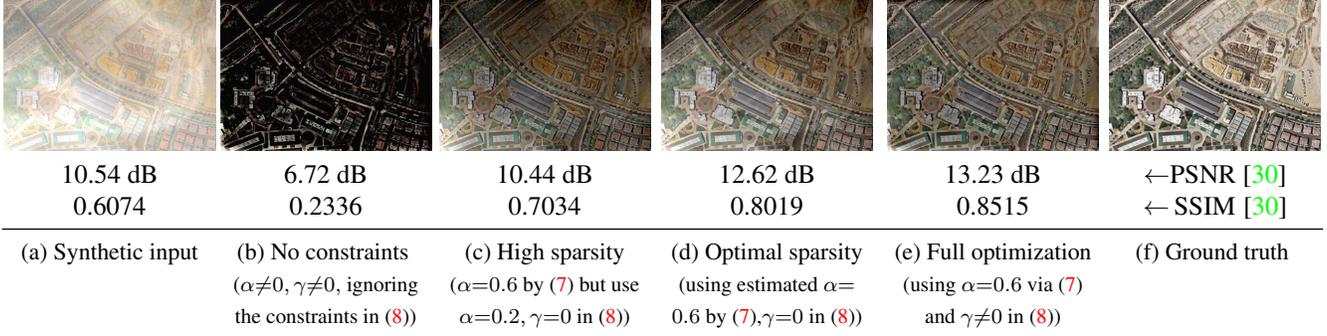


Figure 4: Example showing importance of the different optimization terms in (8) for cloud removal. (b) shows that constraints in (8) are necessary to maintain the sanity of a solution by limiting it within a sensible bound. Highly sparse gradient prior is not always helpful as shown in (c), however by using domain knowledge (sec.2.1, 2.3) we can enforce the optimal sparsity and thereby improving the solution to (d). The prior (5) derived via analysing high-altitude cloud properties helps to boost the performance to (e). PSNR and SSIM are derived after one to one comparing (a) to (e) images with the ground truth in (f).

strained least squares optimization (using [13]) as follows,

$$\hat{\alpha} = \underset{\alpha}{\operatorname{argmin}} \left\| \log(P(x)) - |x|^\alpha - b \right\|^2$$

subject to  $|x| \geq |x_o|$  and  $0 < \alpha < 1$ . (7)

## 2.4. Optimization for Latent Layers

The latent layers can be separated by maximizing the joint probability  $P(B, C)$ , which is equivalent to minimizing  $-\log P(B, C)$ . Similar to [77] assuming the two layers are independent, thus their gradients will also be independent (*i.e.*  $P(B, C) = P(B)P(C)$ ). Background should be non-negative ( $0 \leq (B)_i$ ) and cannot be more than the actual image ( $(B)_i \leq (I)_i$ ). Now by abusing the notations as  $(\partial_j * X)_i = D_i^j X$  and  $\partial_j \in \mathbb{J}_X \equiv j \in \mathbb{J}_X$ ; substituting  $C = I - B$  in (3) and (5); and after integrating the weights ( $\sigma_1, \sigma_2$ ) associated with (2) and (3) together as a single parameter  $\lambda$ , we formulate the energy minimization problem as,

$$\min_B \sum_{i \in \mathbb{I}} \left\{ \sum_{j \in \mathbb{J}_B} |D_i^j B|^\alpha + \sum_{j \in \mathbb{J}_C} \frac{\lambda}{2} \left\| D_i^j B - D_i^j I \right\|^2 \right\} + \frac{\gamma}{2} \|M_B \odot (B - I)\|^2, \quad \text{s.t. } 0 \leq (B)_i \leq (I)_i. \quad (8)$$

In (8),  $\alpha$  is given by (7). As  $\alpha < 1$ , (8) is non-convex, so we use half-quadratic splitting procedure [23, 33, 50] for finding its solution. For casting (8) into half-quadratic splitting framework, we have to make its objective function separable. So introducing auxiliary variables  $y_i^j$  at each pixel, we can split the cost function in (8) as,

$$\min_{B, y^j} \sum_{i \in \mathbb{I}} \sum_{j \in \mathbb{J}_B} \left\{ |y_i^j|^\alpha + \frac{\beta}{2} (D_i^j B - y_i^j)^2 \right\} + \sum_{i \in \mathbb{I}} \sum_{j \in \mathbb{J}_C} \frac{\lambda}{2} \left\| D_i^j B - D_i^j I \right\|^2 + \frac{\gamma}{2} \|M_B \odot (B - I)\|^2, \quad (9)$$

where the positive weight  $\beta$  monotonically increases with each iteration and as  $\beta \rightarrow \infty$ , (9) gets closer to the cost function in (8). Thus the optimization problem (8) is solved iteratively like: for a fixed  $\beta$ , we solve (9) by alternating between updating  $y^j$  and recovering  $B$  while imposing the constraints stated in (8) at the end of every iteration  $t$ . Figure 7 shows the effect of different initial values of  $\beta$ , variation of the costs in (8), (9), how they converge as  $\beta$  increases with  $t$  and how the clouds look like as optimization progresses; whereas Figure 4 analyzes in detail the importance of various terms in the objective function of (8) for cloud removal on a synthetic input.

**Updating- $y$ .** Once  $B$  is recovered at iteration  $t$  then for a fixed  $B_t$ ,  $y^j$  can be updated for each pixel as,

$$y_{t+1}^j = \underset{y^j}{\operatorname{argmin}} \left\{ |y^j|^\alpha + \frac{\beta}{2} (y^j - D_i^j B_t)^2 \right\}. \quad (10)$$

As this is a single variable optimization, it can be quickly solved by using a lookup table (LUT) [33], which maps the values from  $D_i^j B_t$  to  $y^j$ . We fill the LUT by generating  $10^4$  different gradient values between  $-0.5$  to  $0.5$  (this range is determined from Figure 3g), for specified  $\alpha$  and  $\beta$  values. Required value is interpolated if it is missing from the LUT.

**Recovering- $B$ .** With fixed  $y_t^j$ , (9) becomes quadratic in  $B$ , which can simply be solved by differentiating and setting it to 0. We further apply 2D FFT  $\mathcal{F}$  to quickly recover  $B$  as,

$$Q_t = \lambda \sum_{j \in \mathbb{J}_C} \overline{\mathcal{F}(D^j)} \odot \mathcal{F}(D^j) \odot \mathcal{F}(I) + \beta \sum_{j \in \mathbb{J}_B} \overline{\mathcal{F}(D^j)} \odot \mathcal{F}(y_t^j)$$

$$R = \lambda \sum_{j \in \mathbb{J}_C} \overline{\mathcal{F}(D^j)} \odot \mathcal{F}(D^j) + \beta \sum_{j \in \mathbb{J}_B} \overline{\mathcal{F}(D^j)} \odot \mathcal{F}(D^j),$$

$$B_t = \mathcal{F}^{-1} \left( \frac{Q_t + \gamma \mathcal{F}(M_B^T \odot M_B \odot I)}{R + \gamma \mathcal{F}(M_B^T \odot M_B) + \epsilon} \right), \quad (11)$$

---

**Algorithm 1** Cloud Removal: An Optimization Scheme

---

- 1: **Input:** input image  $I$ ; optimization weights  $\lambda, \gamma$ ; total number of iterations  $T$
  - 2: **Initialize:**  $B_0 \leftarrow I, \beta = \beta_0$
  - 3: **for** iteration  $t$  from 1 to  $T$  **do**
  - 4:   update  $y_t^j$  using (10)
  - 5:   recover  $B_t$  using (11)
  - 6:   update  $B_t : \forall i (B_t)_i \leftarrow \eta + (B_t)_i$  using (12)
  - 7:   close the gap between (8) and (9) via  $\beta \leftarrow 2\beta$
  - 8: **output:** cleared cloud image  $B = B_T$
- 

where the division in (11) is performed element-wise,  $\overline{\mathcal{F}(\cdot)}$  is a complex conjugate of  $\mathcal{F}(\cdot)$  and  $\epsilon (=10^{-10})$  avoids division by 0. Here  $\mathcal{F}(y_t^j)$  need to be computed at each iteration, whereas rest of the terms are precomputed only once.

**Imposing constraints.** The constraints in (8) can be satisfied at each iteration by adding a global normalizing constant  $\eta$  to all elements in  $B_t$  such that  $\eta + (B_t)_i$  falls within  $[0, (I)_i]$ . The suitable  $\eta$  can be obtained by minimizing the following loss function using a gradient descent method,

$$\min_{\eta} \sum_{i \in \mathbb{I}} \mathcal{H}(\eta + (B_t)_i - (I)_i)^2 + \mathcal{H}(-\eta - (B_t)_i)^2, \quad (12)$$

where the Heaviside step (or unit step) function

$$\mathcal{H}(x) = \begin{cases} 1 & \forall x > 0 \\ 0 & \text{otherwise,} \end{cases} \quad (13)$$

penalizes only those terms that violate the constraints given in (8). Thus for each pixel  $i$ ,  $(B_t)_i$  is updated to  $\eta + (B_t)_i$ . Algorithm 1 summarizes the whole iterative optimization process for a high altitude cloud removal.

**Cloud detection.** After executing the Algorithm 1 we can trivially obtain the crude cloudy layer as  $C_T = I - B_T$ . Then we refine it by leveraging the bilateral filtering technique but additionally guiding it via the cloud map  $M_C$  from (4) to obtain the final cloudy layer as,

$$(C)_i = \frac{1}{w} \sum_{k \in \mathbb{N}_i} \mathcal{G}(|\Delta_{i,k} M_C|) \mathcal{G}(|\Delta_{i,k} C_T|) \mathcal{G}(|\Delta_{i,k}|) (C_T)_i, \quad (14)$$

where  $\mathcal{G}(\cdot)$  is a Gaussian kernel;  $\mathbb{N}_i$  is the neighborhood of pixel  $i$  in the image plane;  $\Delta_{i,k} = i - k$ ;  $\Delta_{i,k} X = (X)_i - (X)_k$  and  $w$  is the normalizing term given as,

$$w = \sum_{k \in \mathbb{N}_i} \mathcal{G}(|\Delta_{i,k} M_C|) \mathcal{G}(|\Delta_{i,k} C_T|) \mathcal{G}(|\Delta_{i,k}|). \quad (15)$$

The simultaneous use of spatial blurring factor  $\mathcal{G}(|\Delta_{i,k}|)$ , intensity smoothing weight  $\mathcal{G}(|\Delta_{i,k} C_T|)$  and regularizing term based on cloud properties  $\mathcal{G}(|\Delta_{i,k} M_C|)$ , avoids the smoothing of  $C_T$  across the edges and preserve its fine details to produce the precise cloudy layer  $C$ .

### 3. Experiments

Here, we first describe the experimental setup and implementation details; then verify efficiency of the proposed method using synthetic as well as real world aerial images.

**Dataset.** We constructed a dataset with 100 real-world aerial images gathered from authentic satellite image repositories [1, 16, 24, 48, 71]. For incorporating versatility, the dataset includes diverse geological images (*e.g.* flat terrain like ocean, dessert *etc.*, scenes with moderate clutter like agricultural, rocky territory *etc.* and highly cluttered regions like residential, commercial area *etc.*) having variety of clouds (*e.g.* thick translucent, haze like uniformly spread, feathery locally stretched, thin transparent clouds *etc.*) scanned at different time instances (*e.g.* daytime, evening, night). All are color RGB images with the resolution varying from  $600 \times 400$  to  $1600 \times 1200$ . Our Cloudy Aerial Image (CAI) dataset and Matlab implementation will be available for download (<http://pil.snu.ac.kr>).

**Implementation.** All experiments were conducted on a PC equipped with Intel Core i5® 3.4GHz CPU and 8GB RAM. The proposed method (Algorithm 1) was implemented in Matlab without any GPU acceleration. All the cloud removal results (*i.e.* background  $B_T$ ) were obtained via Algorithm 1 without any further contrast enhancement and cloudy layer was detected using (14). The constrained least squares optimization in (7) was solved using inbuilt Matlab function `lsqnonlin` to obtain the parameter  $\alpha$ . First the input image was converted to the YUV color space and only the luminance ( $Y$ ) component is processed via proposed optimization scheme (*i.e.*  $I=Y$ ) while keeping the  $UV$  components intact (*i.e.* to get cloud removed color output:  $Y=B_T$  and  $YUV \rightarrow RGB$ ). The saturation needed in (4) was obtained by RGB to HSV color transformation. The parameters were either empirically fixed or tuned within a range as  $\gamma = 0.1$  to 10,  $\beta_0=30$ ,  $\lambda=4 \times 10^3$ ,  $\sigma_3=10$ ,  $\sigma_4=1$ ,  $T=6$  and  $\mathcal{N}_i$  as 4-pixel neighborhood.

#### 3.1. Quantitative Comparison

We compared our method with the unimodal cloud and haze removal methods [5, 18, 27, 28, 31, 51, 55, 66, 67, 84] by using the implementations provided by the authors. As the ground truth (GT) was necessary for quantitative analysis, we synthesized the input image by mixing a clear background with the Cirrus cloud (Figure 3f) using the image formation model (1). Thus our synthetic inputs (Figure 4 and 5) get rendered with a typical cloudy layer, which is feathery, curvy, globally diffused and locally condensed.

Figure 5 shows the recovered background  $B$  images on top of the discovered cloudy layer  $C$  by [18, 28, 31, 55] and our method. Each image from Figure 5a to 5e was compared with the GT in 5f for evaluating PSNR (peak signal-

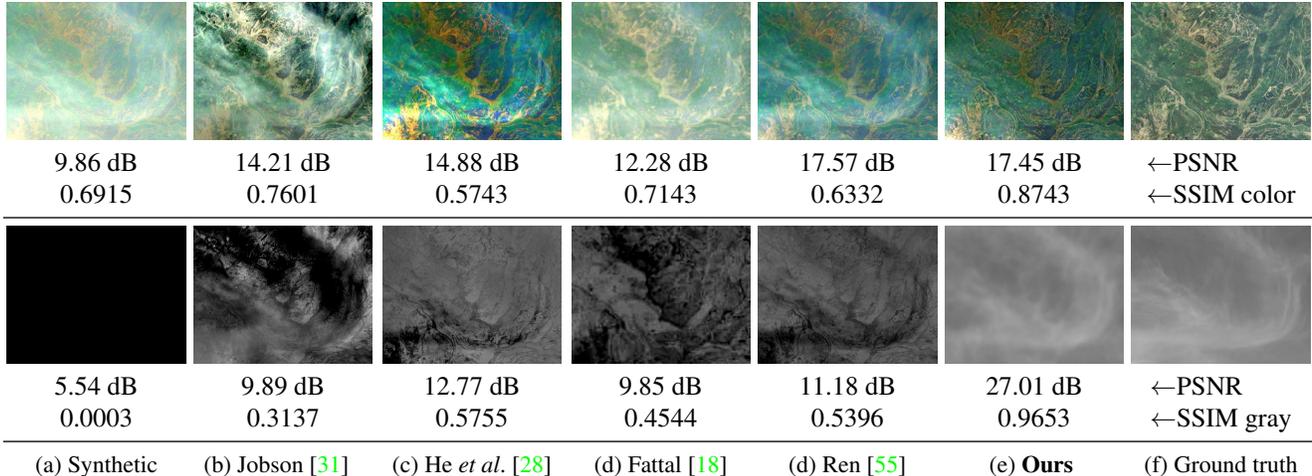


Figure 5: Quantitative comparison using a synthetic input in (a). Images in (b) to (e) show the results by various methods for cloud removal and detection ( $B$  on top of  $C$ ). We show PSNR and SSIM for comparing the color and gray images in (a) to (e) with the ground truth (f). Cloud layer in the synthetic input (a) is completely black image having all pixel values as zero.

to-noise ration) and SSIM (structural similarity index) [30]. Cloudy image in 5a is completely black image, so exhibited extremely low SSIM and 5.54 dB PSNR and served as a lower limit for comparative analysis, thus trivially all the other methods triumphed over baseline for cloud detection; whereas the  $B$  in 5a scored 0.6915 baseline SSIM, which He *et al.* [28] and Ren *et al.* [55] fell short to beat. The haze removal method by He *et al.* [28] nicely cleared the diffused or fuzzy portions of the cloud but failed on removing curvy condensed parts as shown in Figure 5c. Moreover as it enhanced the image contrast with uncleared cloudy area, the image composition greatly diverged from the GT

thereby produced the below baseline 0.5743 SSIM. PSNR is a global measure and does not account for local disparity between images. Peak signal is almost constant in all cases (Figure 5a to 5e) as images always have some saturated pixels (with intensity 255), whereas mean square error compared with GT decreases so all the methods perform better than baseline (Figure 5a) on the PSNR measure and [55] produced the highest PSNR (17.57 dB).

Intensity in the cloudy layers indirectly indicates at which image parts the method has devoted more efforts for the recovery of the latent background. So  $C$  indicates that [31] worked on condensed clouds but intermittently, whereas contrastingly [28] and [55] operated on diffused clouds. Considering SSIM, though [18] and [31] enhanced the  $B$ , they did not match their  $C$  with the GT. As our priors ((2), (3), (5) and (14)) finely align with properties of the high-altitude cloud (sec. 2.2) and the background (sec. 2.1) in synthetic image, the proposed method shows (Figure 5e) significantly higher SSIM (0.8743) and PSNR (17.45 dB) for cloud removal.

Extensive comparison with the related methods is shown in Table 1. Contrast Limited Adaptive Histogram Equalization (CLAHE [84]) outperformed all the methods on cloud removal, however disastrously failed on the cloud detection; plausibly because it removed the diffuse clouds by marginally enhancing the input to  $B$ , so the image color configuration around remaining condensed clouds did not deviate much from the GT (unlike [28] in Figure 5c). This contrasting results indicate the importance of simultaneous evaluation on cloud detection and removal. Our method quantitatively outperformed other methods on simultaneous cloud detection and removal from an image.

Method	Cloud Detection		Cloud Removal	
	PSNR	SSIM	PSNR	SSIM
Homomorphic [51]	9.49 dB	0.3944	12.93 dB	0.5903
CLAHE [84]	6.54 dB	0.1099	18.44 dB	0.8915
He <i>et al.</i> [27]	13.01 dB	0.6071	15.22 dB	0.5577
Jobson <i>et al.</i> [31]	9.89 dB	0.3137	14.21 dB	0.7601
Tarel <i>et al.</i> [66]	8.34 dB	0.2168	11.55 dB	0.6865
He <i>et al.</i> [28]	12.77 dB	0.5755	14.88 dB	0.5743
Fattal [18]	9.85 dB	0.4544	12.28 dB	0.7143
Ren <i>et al.</i> [55]	11.18 dB	0.5396	17.57 dB	0.6332
Berman <i>et al.</i> [5]	11.74 dB	0.4017	14.73 dB	0.6336
Tarel <i>et al.</i> [67]	17.22 dB	0.5629	17.51 dB	0.7243
<b>Ours</b>	<b>27.01 dB</b>	<b>0.9653</b>	17.45 dB	<b>0.8743</b>

Table 1: Quantitative comparison (using input Figure 5a) with the various cloud and haze removal methods (top 3 results in each column are color coded as RGB respectively).

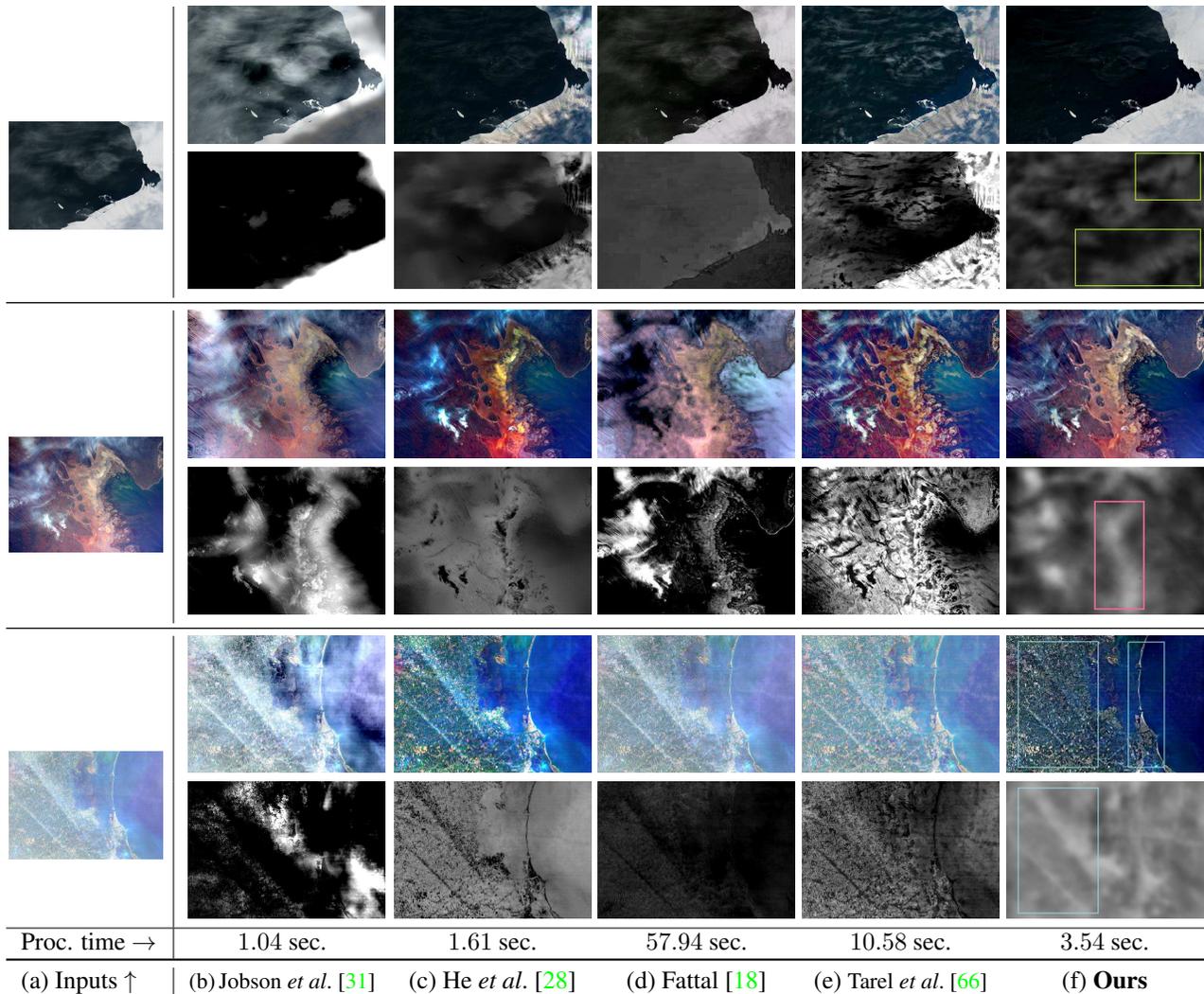


Figure 6: Qualitative comparison with competitive methods. (a) shows the real-world satellite input images covered with various clouds (from 1<sup>st</sup> to 3<sup>rd</sup> row as thin locally condensed, thick local and thin diffused with thick striped) and illuminated differently (from 1<sup>st</sup> to 3<sup>rd</sup> row as dark, bright and faint). For these inputs, the (b) to (e) show the cloud removal and detection ( $B$  on top of  $C'$ ) results. Average computation time for processing all the inputs from (a) is compared in the bottom row.

### 3.2. Results on Real-World Inputs

Figure 6 shows the comparison on real-worlds images from [24, 48, 71] having variety of clouds and illuminations. Top row shows that the bright Antarctic iceberg [48] camouflages the spongy clouds whereas the dark calm sea unmasks them. Thus it serves as an excellent input for assessing the cloud detection ability. As in Figure 6c, [28] correctly detected and removed clouds over the sea but confused ice parts for the white clouds. Instead of detecting the clouds, [18] and [31] mistakenly detected complementary regions sea and iceberg respectively. Along with evident clouds over the sea, our method also correctly found the obscure cloudy patches over the icebergs (green boxes in 6f).

In Figure 6a, middle row has relatively thicker clouds where the transmission is negligible ( $t_i \approx 0$ ). It is an example of ‘low’-altitude clouds, which can only be removed via inpainting based methods [26, 42, 80]. However it presents a good test case for analyzing how the image enhancement based cloud removal methods perform in the presence of saturated image regions (*i.e.* thick clouds) within the colorful background. The contrast of the input was diminished by [18] whereas over enhanced by [28] and appropriately amplified by [66] to produce visually pleasing  $B$ . [18] detected more precise cloudy regions than the rest of methods. Our method preserved the contrast but as it is tailored towards high-altitude clouds, it made a mistake similar to [31] in detecting low-altitude thick clouds (red box in Figure 6f).

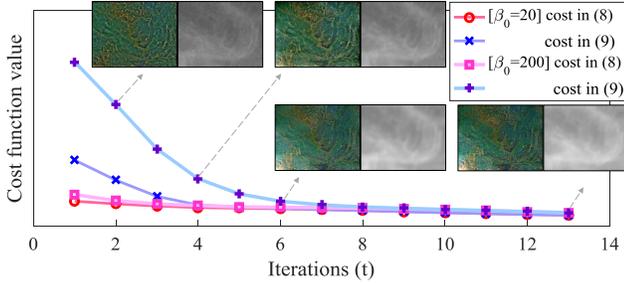


Figure 7: Effect of initializer  $\beta_0$  and variation of the costs in (8), (9) with the iterations given in Algorithm 1. Intermittent results ( $B$  on top of  $C$ ) are shown in the inset.

The last row input in Figure 6a has both globally diffused and locally striped high-altitude clouds. As in Figure 6c, [28] cleared the diffused clouds but failed to remove condensed clouds, which can also be seen as the void strips in  $C$ . Previously [18] did the best job on detecting low-altitude clouds but here on high-altitude clouds it broke down ( $C$  in Figure 6d). Our method not only cleared the diffused clouds but also removed to much extent the bright striped clouds (left-side blue box in  $B$  of Figure 6f) along with preserving the original image structure of bright striped coastal line (right-side blue box in  $B$  of Figure 6f).

Our method arguably shows the best results as given in Figure 6f, considering detection of the thin camouflaged clouds, detecting the feathery, striped structures of the clouds and preservation of the accurate color configuration while removing the diffused as well as locally condensed high-altitude clouds.

**Computation-time.** Average computational time on the images of about HD resolution (CAI dataset) is reported in the Figure 6. The proposed method is not the fastest but consumes decent time about 3.5 seconds. The quicker methods Jobson *et al.* [31] and He *et al.* [28] enhance the fuzzy cloud images within a second but suffer on structured cloudy images. On the other hand Fattal [18] and Tarel *et al.* [66] takes about 58 and 10 seconds respectively. So the faster convergence (sec.3.3) and considerably lower processing time of the proposed method can be credited to the LUT and FFT based iterative optimization scheme.

### 3.3. Optimization Convergence Analysis

Figure 7 shows the values of the cost function given in (8) and (9) at each iteration  $t$  of the Algorithm 1. Starting from  $\beta_0$ , the  $\beta$  doubles at every iteration. As  $\beta$  increases, the discrepancy between  $D_i^j B$  and  $y_i^j$  decreases in (9). This can be seen in Figure 7 *i.e.* the gap between (8) and (9) declines with the progress of iteration. The higher initializer value ( $\beta_0=200$ ) makes (9) to converge quicker but to a slightly higher value than the lower initializer ( $\beta_0=20$ ) *i.e.*

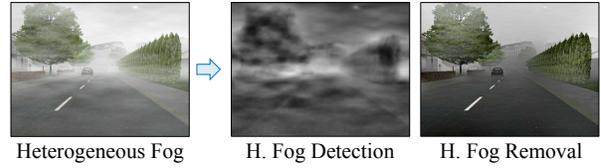


Figure 8: Testing our method on heterogeneous fog.

$$\text{convergence} \Rightarrow \begin{cases} \text{gap-rate: } \mathcal{R} & \mathcal{R}_{20}^{\beta_0} < \mathcal{R}_{200}^{\beta_0} \\ \text{limit-value: } \mathcal{L} & \mathcal{L}_{20}^{\beta_0} < \mathcal{L}_{200}^{\beta_0}, \end{cases} \quad (16)$$

where the rate of gap reduction is the change in cost difference per unit iteration change *i.e.*  $\mathcal{R} = \frac{\Delta|[(8)-(9)]}{\Delta t}$  and limit  $\mathcal{L}$  is the value to which the objective function converges. For this cost splitting based minimization problem, lower  $\mathcal{L}$  and higher  $\mathcal{R}$  is preferred, so with some tradeoff we have used  $\beta_0=30$ . Four pairs of in-between results for cloud removal and detection ( $B, C$ ) are also shown in Figure 7 after 2, 4, 6 and 13 iterations. We can see that the results at  $t=6$  are qualitatively similar to those at  $t=13$ . So by trading between the accuracy and the speed we set  $T=6$  in our implementation. Thus empirically the Algorithm 1 rapidly converges to a good solution.

## 4. Discussion

We have proposed a novel algorithm for simultaneous detection and removal of the high-altitude clouds from a single image. We have comprehensively analyzed the properties of geological images and high-altitude clouds to derive the appropriate image priors. Our method exploits the statistical gradient disparity of cloud and non-cloud regions, non-uniformity, low-saturation of the clouds and gradient sparsity of the aerial images, by incorporating all the priors in a unified optimization framework; whose FFT and LUT based iterative solution quickly removes the clouds within a few seconds on real-world satellite images.

Visually heterogeneous fog [67] looks like clouds descended to ground. However intrinsically it differs from high-altitude clouds on many aspects such as: ground imaging process does not follow the model in Figure 2; some portion of the fog is diffused throughout the scene like haze; background remains at variable distance from the camera; and gradient statistics is not globally consistent as image contains some flat sky regions. Figure 8 shows that though our method can reduce the non-diffused portions of the fog, the heterogeneous fog removal needs a separate treatment.

## Acknowledgement

This work was supported by the ICT R&D program of MSIP/IITP (2017-0-00306, Development of Multimodal Sensor-based Intelligent Systems for Outdoor Surveillance Robots) and Brain Korea 21 Plus (BK21 plus) Project.

## References

- [1] Landsat missions. <https://landsat.usgs.gov/>.
- [2] McGraw-hill yearbook of science and technology, 2005.
- [3] L. Battan. Weather. *Foundations of Earth Science Series, Prentice Hall*, page 74, 1974.
- [4] S. Becker, J. Bobin, and E. J. Candes. Nesta: A fast and accurate first-order method for sparse recovery. In *technical report*. California Institute of Technology, 2009.
- [5] D. Berman, T. Treibitz, and S. Avidan. Non-local image dehazing. In *IEEE CVPR*, 2016.
- [6] E. Berry and G. G. Mace. Cirrus cloud properties and the large-scale meteorological environment: Relationships derived from a-train and ncepncar reanalysis data. *American Meteorological Society*, 2013.
- [7] E. C and J. Romberg. L1-magic: Recovery of sparse signals via convex programming, 2005.
- [8] G. Camps-Valls, T. B. Marsheva, and D. Zhou. Semi-supervised graph-based hyperspectral image classification. *IEEE Tran on Geosci. and Remote Sensing*, 45, 2007.
- [9] T. Chan, S. Esedoglu, F. Park, and A. Yip. Recent developments in total variation image restoration. In *Math Models of Computer Vision*. Springer Verlag, 2005.
- [10] F. Chen, D. Yan, and Z. Zhao. Haze detection and removal in remote sensing images based on undecimated wavelet transform. *Geomatics and info sci*, 2007.
- [11] J. Chen, X. Zhu, J. E. Vogelmann, F. Gao, and S. Jin. A simple and effective method for filling gaps in landsat etm+ slc-off images. *Remote Sensing of Environment*, 2011.
- [12] J. Cihlar and J. Howarth. Detection and removal of cloud contamination from avhrr images. *IEEE Trn. GRS*, 1994.
- [13] T. Coleman and Y. Li. An interior, trust region approach for nonlinear minimization subject to bounds. *SIAM Journal on Optimization*, 1996.
- [14] P. L. Combettes and J. C. Pesquet. Image restoration subject to a total variation constraint. *IEEE Tran on IP*, 13, 2004.
- [15] J. A. Day. The book of clouds, 2005.
- [16] D.G. (DigitalGlobe) high-resolution earth imagery. <https://www.digitalglobe.com/>.
- [17] D. R. Dowling and L. F. Radke. A summary of the physical properties of cirrus cloud. *Journal of Applied Meteorology*, 29, 1990.
- [18] R. Fattal. Single image dehazing. In *ACM SIGGRAPH*, 2008.
- [19] M. Ferencat and N. Boujemaa. Interactive remote-sensing image retrieval using active relevance feedback. *IEEE Transactions on Geoscience and Remote Sensing*, 45, 2007.
- [20] R. A. Frey, S. A. Ackerman, Y. Liu, K. I. Strabala, H. Zhang, J. R. Key, and X. Wang. Cloud detection with modis part I: Improvements in the modis cloud mask for collection 5. *Journal of Atmospheric and Oceanic Technology*, 2008.
- [21] S. Gabarda and G. Cristobal. Cloud covering denoising through image fusion. *Image and Vision Comp*, 2007.
- [22] S. Gallegos, R. Holyer, S. Peckinpaugh, and C. Cheng. Apparatus and method for cloud masking, US patent 5612901, 1997.
- [23] D. Geman and C. Yang. Nonlinear image recovery with half-quadratic regularization. *IEEE Tran IP*, 1995.
- [24] G.L.C.F. Global land cover facility. <http://landcover.org/>.
- [25] N. Greeshma, M. Baburaj, and S. George. Reconstruction of cloud-contaminated satellite remote sensing images using kernel pca-based image modelling. *Arab Journal of Geoscience*, 2016.
- [26] K. He and J. Sun. Statistics of patch offsets for image completion. *ECCV*, 2012.
- [27] K. He, J. Sun, and X. Tang. Single image haze removal using dark channel prior. In *IEEE CVPR*, 2009.
- [28] K. He, J. Sun, and X. Tang. Single image haze removal using dark channel prior. *IEEE Trans on PAMI*, 2011.
- [29] E. Helmer and B. Ruefenacht. Cloud-free satellite image mosaics with regression trees and histogram matching. *Photogramm. Eng. Remote Sensing*, 71, 2005.
- [30] A. Hore and D. Ziou. Image quality metrics: Psnr vs. ssim. In *ICPR*, 2010.
- [31] D. J. Jobson, Z. Rahman, and G. A. Woodell. A multiscale retinex for bridging the gap between color images and the human observation of scenes. *IEEE TIP*, 1997.
- [32] J. Ju and D. P. Roy. The availability of cloud-free landsat etm+ data over the conterminous united states and globally. *Remote Sensing of Environment*, 112, 2008.
- [33] D. Krishnan and R. Fergus. Fast image deconvolution using hyper-laplacian priors. In *NIPS*, 2009.
- [34] A. Levin and Y. Weiss. User assisted separation of reflections from a single image using a sparsity prior. *IEEE Tran on Pattern Analysis and Machine Intelligence*, 29, 2007.
- [35] M. Li, S. C. Liew, and L. K. Kwok. Generating cloud free and cloud-shadow free mosaic for spot panchromatic images. In *IEEE International Geoscience and Remote Sensing Symposium*, volume 4, 2002.
- [36] Q. Li, W. Lu, J. Yang, and J. Z. Wang. Thin cloud detection of all-sky images using markov random fields. *IEEE Geoscience And Remote Sensing Letters*, 2012.
- [37] S. Liang, H. Fang, and M. Chen. Atmospheric correction of landsat ETM+ land surface imagery part-I: methods. *IEEE Geoscience and Remote Sensing*, 39, 2001.
- [38] C. H. Lin, P. H. Tsai, K. H. Lai, and J. Y. Chen. Cloud removal from multitemporal satellite images using information cloning. *IEEE Tran on GRS*, 51, 2013.
- [39] K.-N. Liou. Influence of cirrus clouds on weather and climate processes: A global perspective. *Monthly Weather Review*, 1986.
- [40] J. Liu, X. Wang, M. Chen, S. Liu, X. Zhou, Z. Shao, and P. Liu. Thin cloud removal from single satellite images. *OPTICS EXPRESS*, 22, 2014.
- [41] J. Long, Z. Shi, W. Tang, and C. Zhang. Single remote sensing image dehazing. *IEEE GRSL*, 11, 2014.
- [42] L. Lorenzi, F. Melgani, and G. Mercier. Inpainting strategies for reconstruction of missing data in vhr images. *IEEE Geoscience and Remote Sensing Letters*, 8, 2011.
- [43] D. M. Ludlum. National audubon society field guide to north american weather. *Random House, New York*, 2001.

- [44] A. Maalouf, P. Carre, B. Augereau, and C. Fernandez-Maloigne. A bandelet-based inpainting technique for clouds removal from remotely sensed images. *IEEE Transactions on Geoscience and Remote Sensing*, 47, 2009.
- [45] S. L. H. Mascle and C. Andre. Use of markov random fields for automatic cloud/shadow detection on high resolution optical images. *ISPRS J of PRS*, 2009.
- [46] F. Melgani. Contextual reconstruction of cloud-contaminated multitemporal multispectral images. *IEEE Trans on Geoscience and Remote Sensing*, 44, 2006.
- [47] O. R. Mitchell, E. J. Delp, and P. L. Chen. Filtering to remove cloud cover in satellite imagery. *IEEE T. GE*, 1977.
- [48] N.A.S.A. National aeronautics and space administration. <https://www.nasa.gov/>.
- [49] S. L. M. Neto, A. V. Wangenheim, E. B. Pereira, and E. Comunello. The use of euclidean geometric distance on rgb color space for the classification of sky and cloud patterns. *J. Atmos. Ocean. Technol.*, 2010.
- [50] M. Nikolova and M. K. Ng. Analysis of half-quadratic minimization methods for signal and image recovery. *SIAM Journal on Scientific Computing*, 27, 2005.
- [51] U. Nnolim and P. Lee. Homomorphic filtering of colour images using a spatial filter kernel in the hsi colour space. In *IEEE Instru and Measure. Tech Conf*, 2008.
- [52] I. Pitas and A. N. Venetsanopoulos. *Homomorphic Filters*. 1990.
- [53] L. Poggio, A. Gimona, and I. Brown. Spatio-temporal modis evi gap filling under cloud cover: An example in scotland. *ISPRS J of Photogrammetry and Remote Sensing*, 2012.
- [54] P. Rakwatin, W. Takeuchi, and Y. Yasuoka. Restoration of aqua modis band 6 using histogram matching and local least squares fitting. *IEEE Tran GRS*, 47, 2009.
- [55] W. Ren, S. Liu, H. Zhang, J. Pan, X. Cao, and M.-H. Yang. Single image dehazing via multi-scale convolutional neural networks. In *ECCV*, 2016.
- [56] H. Richard and H. R. Keith. *Boater's bowditch: The small craft american practical navigator*. MRM Press, 2000.
- [57] J. Roerink, M. Menenti, and W. Verhoef. Reconstructing cloudfree ndvi composites using fourier analysis of time series. *Int. J. Remote Sens.*, 21, 2000.
- [58] R. E. Rossi, J. L. Dungan, and L. R. Beck. Kriging in the shadows: Geostatistical interpolation for remote sensing. *Remote Sensing of Environment*, 1994.
- [59] A. B. Salberg. Land cover classification of cloud-contaminated multitemporal high-resolution images. *IEEE Tran on Geoscience and Remote Sensing*, 49, 2011.
- [60] R. Saunders and K. Kriebel. An improved method for detecting clear sky and cloudy radiances from avhrr data. *International Journal of Remote Sensing*, 1988.
- [61] G. Shaw and H. Burke. Spectral imaging for remote sensing. *Lincoln Lab Journal*, 14, 2003.
- [62] H. Shen, H. Li, Y. Qian, L. Zhang, and Q. Yuan. An effective thin cloud removal procedure for visible remote sensing images. *ISPRS J of PRS*, 2014.
- [63] 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. *J. Atmos. Ocean. Technol.*, 2006.
- [64] H. Sundqvist, E. Berge, and J. E. Kristjansson. Condensation and cloud parameterization studies with a mesoscale numerical weather prediction model. *Weather review*, 1989.
- [65] M. F. Tappen, W. T. Freeman, and E. H. Adelson. Recovering intrinsic images from a single image. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27, 2005.
- [66] J.-P. Tarel and N. Hautiere. Fast visibility restoration from a single color or gray level image. In *ICCV*, 2009.
- [67] J.-P. Tarel, N. Hautiere, L. Caraffa, A. Cord, H. Halmaoui, and D. Gruyer. Vision enhancement in homogeneous and heterogeneous fog. *IEEE ITS Magazine*, 2012.
- [68] G. Tselioudis and C. Jakob. Evaluation of midlatitude cloud properties in a weather and a climate model: Dependence on dynamic regime and spatial resolution. *Journal of Geophysical research, Atmospheres*, 107, 2002.
- [69] D. Tseng and C. Chien. A cloud removal approach for aerial image visualization. *Int J of ICIC*, 2013.
- [70] D.-C. Tseng, H.-T. Tseng, and C.-L. Chien. Automatic cloud removal from multi-temporal spot images. *AMC*, 2008.
- [71] U.S.G.S. US geological survey long term archive. <https://earthexplorer.usgs.gov/>.
- [72] van der Meer. Remote-sensing image analysis and geostatistics. *International Journal of Remote Sensing*, 33, 2012.
- [73] M. Wan and X. Li. Removing thin cloud on single remote sensing image based on SVVF. In *ICOACS*, 2016.
- [74] B. Wang, A. Ono, K. Muramatsu, and N. Fuhiwara. Automated detection and removal of clouds and their shadows from landsat tm images. *IEICE Trans. Inf. Syst.*, 1999.
- [75] Z. Wang, J. Jin, J. Liang, K. Yan, and Q. Peng. A new cloud removal algorithm for multi-spectral images. *MIPPR SAR and multispectral image processing*, 2005.
- [76] Z. Wei, S. Jiali, W. Meng, and H. Dapeng. A thin cloud removal method from remote sensing image for water body identification. *China Geogra. Sci.*, 23, 2013.
- [77] Y. Weiss. Deriving intrinsic images from image sequences. In *IEEE ICCV*, 2001.
- [78] Y. Xiong, H. Yan, and C. Yu. Algorithm of removing thin cloud-fog cover from single remote sensing image. *Journal of Information and Computational Science*, 2014.
- [79] C. Zeng, H. Shen, and L. Zhang. Recovering missing pixels for landsat etm + sle-off imagery using multi-temporal regression analysis and a regularization method. *Remote Sensing of Environment*, 2013.
- [80] J. Zhang, D. Zhao, and W. Gao. Group-based sparse representation for image restoration. *IEEE TIP*, 2014.
- [81] Q. Zhang and C. Xiao. Cloud detection of rgb color aerial photographs by progressive refinement scheme. *IEEE Transactions on Geoscience and Remote Sensing*, 52, 2014.
- [82] Y. Zhang, B. Guindona, and J. Cihlar. An image transform to characterize and compensate for spatial variations in thin cloud contamination of landsat images. *Remote Sensing of Environment*, 2002.
- [83] X. Zhu, F. Gao, D. Liu, and J. Chen. A modified neighborhood similar pixel interpolator approach for removing thick clouds in landsat images. *IEEE GRSL*, 9, 2012.
- [84] K. Zuiderveld. *Graphics gems iv. chapter Contrast Limited Adaptive Histogram Equalization*. Academic Press Professional, Inc., 1994.