

Clouds in The Cloud

Anonymous ACCV 2014 submission

Paper ID 215

Abstract. Light-field imaging can be scaled up to continental size, to map the Earth’s atmosphere in 3D. Multiview spaceborne instruments suffer low spatio-temporal and angular resolution, and are extremely expensive and unscalable. We develop lightfield imaging of the sky, by a scalable network of wide-angle cameras looking upwards, which upload their data to the cloud. This new type of imaging-system poses *new computational vision and photography problems*, some of which generalize prior monocular tasks. These include radiometric self-calibration across a network, overcoming flare by a network, and background estimation. On the other hand, network redundancy offers *solutions* to these problems, which we derive. Based on such solutions, the light-field network enables unprecedented ways to measure nature. We demonstrate this experimentally by 3D recovery of clouds, in high spatio-temporal resolution. It is achieved by space carving of the volumetric distribution of semi-transparent clouds. This sensing approach can complement satellite imagery, be useful to aviation meteorology and study of birds. The radiometric capabilities can make aerosol tomography realizable, and give new, powerful tools to atmospheric scientists.

1 Introduction

Following the introduction of plenoptic imaging to computer vision [1], light-field imaging [2–6] has had a significant impact on sensing. This modality samples the radiance distribution in location and direction. It has so far been used mainly in small-scale setups. However, it can be scaled up to continental size, to map the Earth’s atmosphere in 3D. Sampling the atmospheric radiance spatio-angularly is achieved by a few spaceborne and airborne instruments, including the Multiangle Imaging SpectroRadiometer (MISR) [7, 8], the Airborne Multiangle SpectroPolarimetric Imager (AirMSPI) [9, 10] and POLDER [11–13]. However, these imaging architectures suffer several drawbacks. They have crude resolution spatially (effectively several kilometers per pixel), angularly (≈ 7 angles per view), and temporally (orbit takes several days to return to the same terrestrial spot). Furthermore, spaceborne instruments are extremely expensive and unscalable. Hence, in this work we develop a counter approach: the atmospheric lightfield is captured *from below*, by wide-angle cameras *looking upwards*. Moreover, instead of expensive one-off instruments, our approach is a *scalable sensor network* that captures images simultaneously over a very large area, densely.

Creating and exploiting such a light-field imaging-network poses several requirements: low-cost units, communications, and *tailored computational photography algorithms*. The first two requirements are met thanks to cellular-network infrastructure, cellular-based low-cost cameras and *cloud computing* services. Hence, we can deploy

solar-powered cameras at will, nearly anywhere. By wireless, they upload their sky-images to the “cloud”, from which the light-field data can be analyzed. However, this new type of imaging-system gives rise to new problems and algorithms, part of which we deal with in this paper. In a sense, the network *generalizes some compute vision problems* that had been posed for monocular setups a decade ago. On the other hand, network redundancy offers *solutions* to these problems.

The computational photography problems include radiometric self-calibration across a network of cameras, background estimation, and overcoming saturation and flare by a network. Consequently, the light-field network enables unprecedented 3D imaging of clouds, in high spatio-temporal resolution. This is achieved by *space carving* in a *volume-metric distribution* of semi-transparent clouds. We demonstrate the approach using real field experiments. As a result, we believe this approach can complement multi-angular satellite cloud imagery, perhaps make aerosol tomography [14] realizable, and offer new capabilities to study weather phenomena and birds.

2 Background

2.1 Single-camera radiometric self-calibration

A large network should use low-cost camera units. Such cameras often have spatial and temporal radiometric inconsistencies. For a single camera, consistent readouts can be obtained in the field by self-calibration. The strongest methods rely on redundant images, taken at different settings. In particular, images taken at different exposure times calibrate the radiometric response, up to a global scale factor. Similarly, images taken by slowly panning the camera yield spatial non-uniformity correction (NUC) of offset and gain (e.g., by vignetting [15, 16]). For example, spatial gain is often modeled by a function $M(\mathbf{x})$, where $\mathbf{x} = (x, y)$ is a camera pixel. Thus the image pixel irradiance at time t is

$$\tilde{I}_t(\mathbf{x}) = M(\mathbf{x})I_t(O), \quad (1)$$

where $I_t(O)$ is the pixel irradiance when $M = 1$, for observed object O . Let the camera be panned, so that the object O projects to pixel \mathbf{x}' in time t' . *Correspondence* is established between the two pixels, by image alignment. This yields a redundant measurement $\tilde{I}'_t(\mathbf{x}') = M(\mathbf{x}')I'_t(O)$. Assuming brightness constancy, $I_t(O) = I'_t(O)$, the two measurements yield a linear constraint

$$\log M(\mathbf{x}) - \log M(\mathbf{x}') = \log \tilde{I}_t(\mathbf{x}) - \log \tilde{I}'_t(\mathbf{x}'). \quad (2)$$

Aggregating many such constraints over different pixels and frames recovers $M(\mathbf{x})$, up to a global exponential ambiguity. This recovery makes pixel readout spatially consistent. Similar formulations exist for an additive offset. In Sec. 4.1 we expand this principle to a camera-array.

2.2 Avoiding Blooming, Lens-Flare in a Single Camera

For a wide-angle camera observing the sky, sun-rays are liable to frequently shine directly into the lens. This creates problems. First, the sun projection saturates pixels and

creates blooming that challenges the sensor even in high-dynamic range setups. Second, the sun-rays create an additive, spatially varying lens-flare. Reducing lens flare was suggested by computational photography methods [17–20]: images can be taken by a specialized detector array, for nearby objects. Alternatively, flare features that are sharp can be eliminated by rotating the camera and taking consecutive frames of a static scene. Both approaches complicate the need for simple, low-cost units and operation.

Consequently, sky-observing wide-field cameras commonly have a *dynamic sun blocker*. It is essentially an opaque object raised above the camera optics, blocking the Sun from view. There are various configurations, but all of them *move*, as the Sun direction changes during the day and across the year. Motorized solutions [21] that need to work year-around significantly complicate such camera units, making them very expensive. In Sec. 6 we show that a large camera network inherently bypasses the problem, without a need to constantly move a Sun blocker.

2.3 Current 3D Cloud Mapping

Existing sky-imaging systems used for research and operational meteorology¹ are few, relying on high quality cameras and other parts [23–28]. Due to their complexity and costs, they were only used to estimate *cloud-base* over narrow regions right above a narrow-baseline camera pair. Satellite-based estimation of 3D *cloud-tops* has been proposed by MISR [29]. It takes several minutes for MISR to capture multiple viewpoints of a region, during which the clouds generally move. Weather radars sense *raindrops*, which are much larger than cloud-drops and ice crystals. Hence radar does not sense clouds that do not produce rain.

3 Self-Calibration in a Camera Network

Even in high grade equipment, there are generally radiometric inconsistencies between cameras. Hence, the same distant object has slightly different intensity readouts for different cameras. This problem is exacerbated in low-cost cameras, which have somewhat loose quality control. A large light-field network would need to use many such cameras. Estimating and compensating for inter-camera differences is helpful for subsequent computer vision algorithms. It is clear that radiometric consistency would be necessary for radiometric-based tasks, such as tomography [14]. However, it also assists non-radiometric, higher level tasks. For example, consider cloud segmentation. Segmentation (which may rely on image-value thresholds) has to be as consistent as possible across different viewpoints, otherwise contradictions may arise.

Similarly to Sec. 2.1, the solution relies on redundant measurement at *corresponding* points. For correspondence, geometric calibration [30] is a first necessity, which is treated in Sec. 3.1.

¹ There are also ground viewing webcams that happen to see sky parts [22] and weather cameras that are too sparse to be integrated for recovery.

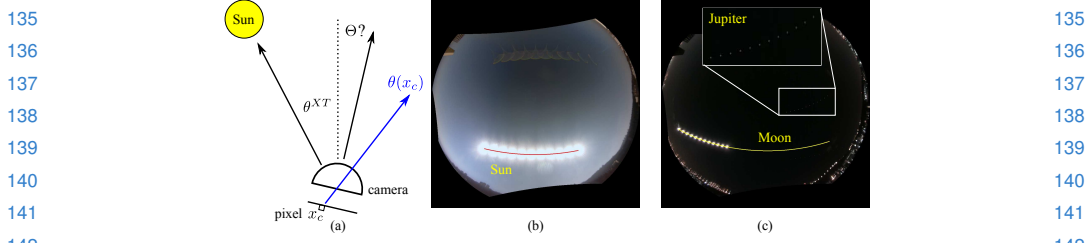


Fig. 1. (a) To estimate the camera yaw-pitch-roll angle vector Θ_c , we rely on image locations of extra-terrestrial objects, whose direction vector $\theta^{XT}(t)$ is known for all time t . (b) Photo-montage of the daylight sky. It shows the Sun at different hours, the expected trajectory based on the estimated Θ_c and lens-flares. (c) Photo-montage of night sky images. It shows the Moon at different times, the expected trajectory based on the estimated Θ_c , and a close-up on the corresponding sampled images of Jupiter.

3.1 Geometric calibration

The cameras are independently deployed in the field. The location vector \mathbf{l}_c of each camera c is known by GPS, and the internal parameters Ψ_c are pre-calibrated in the lab. However, the orientation (yaw, pitch and roll angle vector Θ_c) is loosely set, as the camera is placed in the field. We calibrate the orientation by detecting and tracking extra-terrestrial (XT) objects (Sun and/or Moon) [30, ?], across day or night in the hemispherical field of view. Using astronomical charts per time t and \mathbf{l}_c , an XT object is known to be at angle vector (zenith, azimuth) $\theta^{XT}(t)$ relative to a global coordinate system. Given camera orientation Θ_c , a ray from $\theta^{XT}(t)$ should project to pixel

$$\mathbf{x}_{\text{model}}^{XT}(t) = \mathcal{P}(\theta^{XT}(t)|\Theta_c, \Psi_c), \quad (3)$$

where \mathcal{P} denotes projection (converting ray angles to image pixels). The actual measured pixel $\mathbf{x}_{\text{measured}}^{XT}(t)$ has uncertainty, due to the finite angular size of the XT object and blooming. However, during the course of a day or night, the number of frames N^{frames} is $\mathcal{O}(100)$, leading to a simple optimization formulation:

$$\hat{\Theta}_c = \arg \min_{\Theta_c} \sum_{t=1}^{N^{\text{frames}}} \|\mathcal{P}(\theta^{XT}(t)|\Theta_c, \Psi_c) - \mathbf{x}_{\text{measured}}^{XT}(t)\|^2, \quad (4)$$

The orientation calibration is illustrated in Fig. 1.

Based on $\hat{\Theta}_c$, all captured images $\tilde{I}_{c,t}(\mathbf{x})$ taken by camera c are aligned to the global coordinate system. Any image pixel \mathbf{x} of camera c is backprojected to a ray at direction vector θ in the global system, given by

$$\theta(\mathbf{x}) = \mathcal{P}^{-1}(\mathbf{x}|\hat{\Theta}_c, \Psi_c). \quad (5)$$

The sampled light-field is the radiance measured per location, direction and time, thus given by $\tilde{I}_t[\mathbf{l}_c, \theta(\mathbf{x})]$.

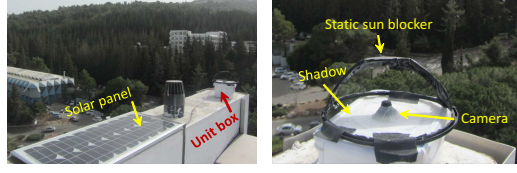


Fig. 2. A unit (network-node) on a rooftop in a hilly area. It is based on a Raspberry-Pi camera and uses solar energy. Its fisheye lens is shadowed here by a mounted static Sun-blocker.

4 Some Details About The Experimental Setup

Before proceeding with mathematical problems and solutions, we now describe a small experimental network, which we built to test the hypotheses of this paper. To show that it can work with very low-cost camera nodes (units), each of the 5 units was built from basic components, and coarse alignment tolerances. Its core is a Raspberry-Pi computer used in DIY electronics, and a 5MP Raspberry-Pi camera. In this camera, the gain, response and white-balance can be fixed, avoiding temporal radiometric variations. We cut-off its OEM lens, and manually mounted a small fisheye lens. Due to this coarse lens-to-chip alignment, *each camera has a different peripheral dead-region*, creating a missing part in the hemispheric view-field and vignetting distinct to each camera. As derived below, a **network as-a-whole can inherently overcome** these issues. Every 30 seconds, synchronously, all units automatically transmit image data using the cellular network to the internet (cloud-service). Each unit is solar powered (Figs. 2,3).

4.1 Radiometric Self-Calibration in a Camera Network

Consider a *fixed* view direction θ observed by several cameras. The values $\{\tilde{I}_t[l_c, \theta]\}_c$ correspond to readouts of parallel rays, back-projected from all cameras in the network. The values in this set generally differ from each other, e.g., $\tilde{I}_t[l_c, \theta] \neq \tilde{I}_t[l_{c'}, \theta]$. There are two causes for this difference:

1. Different camera locations mean different observed objects. Momentarily, camera c may observe a cloud while c' observes a clear sky, or vice versa. Camera c' may momentarily observe a somewhat denser haze volume than c , or vice versa.
 2. Slight inter-camera radiometric inconsistencies, which we need to estimate.
- Cause 1 is usually dominant. We need to overcome it, in order to analyze cause 2.

The assumption we make is of *regional stationarity*. In a region containing the cameras, the chance of a cloud, clear sky, or haziness affecting $\tilde{I}_t[l_c, \theta]$ is *independent* of c . Thus, inter-camera image variations due to atmospheric conditions are *random and unbiased*. Per camera c and view angle θ , bias is due to cause 2. We easily detect and characterize the bias by capturing *statistics over time*.

Consider, for example, Fig. 4. It shows [Top-Left] radiometric inconsistency between cameras a and b . The figure then shows a scatter-plot of $\tilde{I}_t[l_a, \theta]$ vs. $\tilde{I}_t[l_b, \theta]$, $\forall t, \theta$, for the red-channel. From this scatter plot and those of the other color channels,

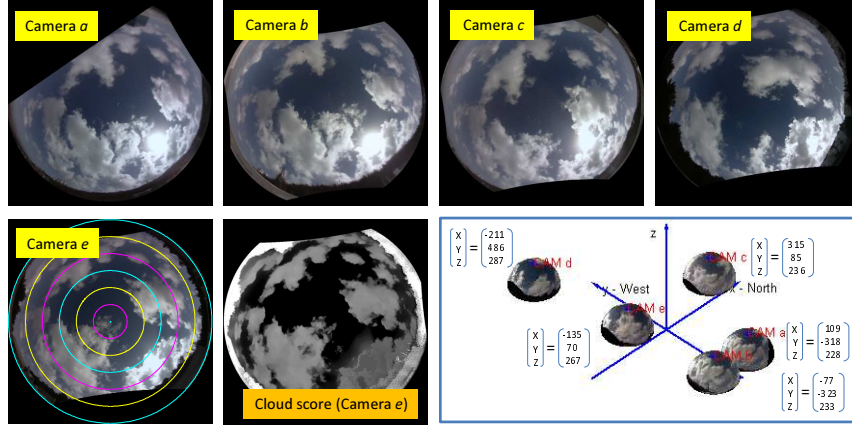


Fig. 3. Images taken simultaneously by a 5-node network. They are geometrically aligned to the zenith and north, and resampled to *polar azimuthal equidistant* projection in this global system. Equidistant zenith angle circles are overlaid on $\tilde{I}_t[l_e, \theta]$ (camera e). Each camera has dead-regions, since the lens-axis had only been roughly centered on the tiny DIY boards. Corresponding to the frame in camera e , a *cloud score* map (Eq. 11) has high values in cloud-pixels, diminishing outside them. [Bottom-right] The 3D setup of the experimental network. It is laterally spread over several hundreds of meters, on rooftops at somewhat different heights.

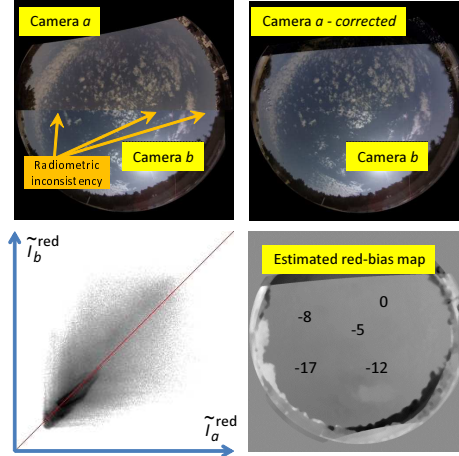


Fig. 4. [Top-Left] Splitting the field of view to upper/lower halves, to pixels corresponding respectively to either $\tilde{I}_t[l_a, \theta]$ or $\tilde{I}_t[l_b, \theta]$. Radiometric inconsistency shows-up as a seam-line across which colors slightly change. [Bottom-Left] Scatter-plot of $\tilde{I}_t[l_a, \theta]$ vs. $\tilde{I}_t[l_b, \theta]$, $\forall t, \theta$, red-channel. [Bottom-right] The estimated offset map $\hat{o}_{b-a}(\theta)$, red channel. It is derived based on a set of images taken during several hours. [Top-right] Splitting the field of view in half, to *corrected* pixels from either $\tilde{I}_t[l_a, \theta]$ or $\tilde{I}_t[l_b, \theta]$: inconsistencies are greatly diminished.

we hypothesized that camera a has a slight offset, particularly in the red channel, relative to camera b . We thus estimated, per color channel, the map of radiometric offset

(across pixels, or ray-directions). A temporal median was used.

$$\hat{o}_{b-a}(\theta) = \text{median}_t \{ \tilde{I}_t[\mathbf{l}_b, \theta] - I_t[\mathbf{l}_a, \theta] \}. \quad (6)$$

The map $\hat{o}_{b-a}(\theta)$ was then spatially smoothed and used to correct $\tilde{I}_t[\mathbf{l}_a, \theta]$. As shown in Fig. 4, the results have much better inter-camera consistency. This is apparent in images taken throughout that day. A similar process was applied to other cameras, but they had negligible radiometric offsets with respect to camera b .

The process was then repeated to detect slight local variations of gain (vignetting). Suppose there is no offset. In analogy to Eq. (6), the gain in b is higher than in a by a factor

$$\hat{M}_{b/a}(\theta) = \text{median}_t \{ \tilde{I}_t[\mathbf{l}_b, \theta] / I_t[\mathbf{l}_a, \theta] \} \quad (7)$$

After radiometric corrections, the light-field samples are denoted tilde $\hat{I}_t[\mathbf{l}_b, \theta]$.

5 Network-assisted Background Estimation

It is often helpful to characterize the background, as motivated by change-detection computer vision algorithms. In monocular settings, such algorithms use temporal filtering to characterize the background: foreground dynamic objects are at *different locations at different times* and are thus pruned. In our case this translates to stating that a cloud in $\tilde{I}_t[\mathbf{l}_c, \theta]$ is often not in $\tilde{I}_{t'}[\mathbf{l}_c, \theta]$, when $t' \neq t$. However, if clouds move slowly, while illumination gradually changes, temporal filtering may be insufficient. This is illustrated in Fig. 5.

A light-field network enhances this principle, with more effective pruning-per-time. After calibration, *simultaneous* images captured by *different camera nodes* are different from each other. Thus, change of viewpoint has an effect similar to change in time: a cloud in $\tilde{I}_t[\mathbf{l}_c, \theta]$ is often not in $\tilde{I}_t[\mathbf{l}_{c'}, \theta]$. This principle relies on the *regional stationarity* mentioned above. Consequently, background sky values are obtained by data filtering over both time and viewpoint.

In broad daylight, clouds are brighter than the blue sky. Hence, an estimator for the sky background is

$$\text{SKY}(\theta) = \arg \min_{t,c} \tilde{I}_t[\mathbf{l}_c, \theta] \quad (8)$$

where $t \in [1 \dots N^{\text{frames}}]$ and $c \in [1 \dots N^{\text{views}}]$. This is illustrated in Fig. 5.

6 Bypassing the Sun through a Camera Network

As Sec. 2.2 explains, in existing full sky-imagers, effects of direct sunlight are mitigated by a small dynamic sun-blocker, which complicates the system and its cost, while having a blind-region. The network offers a different solution, which can be radical, yet simple. On each camera, the sun-blocker is *static*, and has no moving part. The blocker can be large, covering the entire range of directions the Sun may occupy during the year or part of it. In this configuration, each camera unit has a large blind area (See Fig. 6). Nevertheless, the entire *network has no blind spot*, when viewing the atmosphere. This remarkable property is a result of network-redundancy, as we explain.

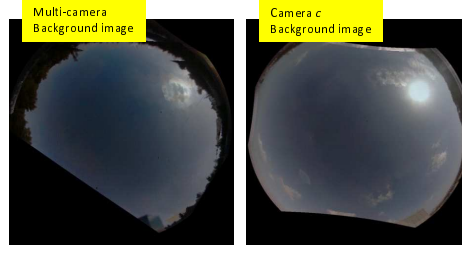


Fig. 5. [Left] Estimation of the sky background, using Eq. (8) based on 5 temporal instances and 5 viewpoints. [Right] Estimation of the sky background, using 5 temporal instances, but just a single viewpoint, resulting in more residual clouds.

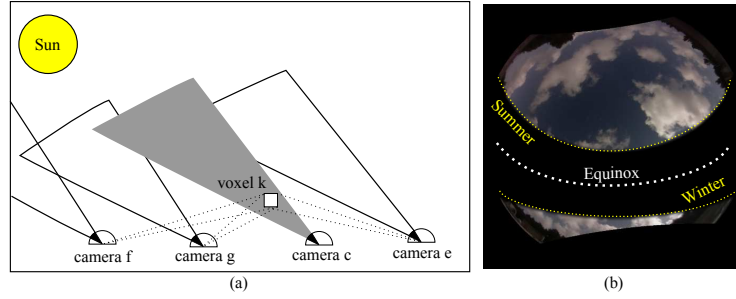


Fig. 6. Camera c has a blind-region, covering Sun directions at l_c . The blind region corresponds to set Γ_c of atmospheric voxels not sensed by camera c . The network as a whole still has coverage of voxels $k \in \Gamma_c$, as they are observed by cameras e, f, g . [Right] Simulation of a whole sky image (polar azimuthal equidistant projection), blocking all solar directions during the year, at a mid-latitude. For tropic cameras, the blind region dissects the field from east to west through the Zenith. In the arctic, the blind region is adjacent to the horizon.

A static year-round sun blocker on camera c permanently obstructs a set Γ_c of atmospheric voxels. These voxels, however, are generally visible at several other cameras, e.g., those indexed e, f, g in Fig. 6. Consequently, if cameras are spread wide enough, the network as a whole has no blind spot, despite permanent sun-blocking. Voxels that are not obstructed from any camera have a few more viewpoints to constrain their content, than voxels in Γ_c .

What is the limitation of this approach? It would not have worked, had objects of interest reside at extremely high altitude. However, nearly all weather phenomena are under the tropopause, whose altitude is typically 10-20km above sea level. Hence, above the camera network domain, any voxel is observed by an angular range exceeding the solar annual trajectory, if the network is spread over several kilometers. For low level clouds (e.g., cumulus), network span of hundreds of meters suffices.

7 3D Clouds by a Camera Network

Network-based acquisition and the processes described above, yield a calibrated sampled light-field of the sky. Using this light-field, we can estimate the 3D cloud structure

above the network domain, and beyond. This is done by the following steps: (A) Per time t , give a cloud score to each ray $[\mathbf{l}_c, \boldsymbol{\theta}]$, as we explain below. (B) Perform a fuzzy version of space carving [?] (reminiscent of back-projection in tomography).

We first describe (B). The set of all sampled light-field rays is \mathcal{R} , where $|\mathcal{R}| = N^{\text{rays}}$. A ray is indexed by r , and it corresponds to a specific $[\mathbf{l}_c, \boldsymbol{\theta}]$. Voxel k projects to a subset of the rays $\rho_k \subset \mathcal{R}$, that reach ν_k viewpoints. Suppose a ray $r \in \mathcal{R}$ has a cloud-score $s(r) \in [0, 1]$, where $s = 0$ means there is definitely no cloud on the ray, while $s = 1$ means there is confidently a cloud there. Per voxel k , define a back-projected score

$$B_k = \left[\prod_{r \in \rho_k} s(r) \right]^{1/|\rho_k|} \quad \text{if } \nu_k \geq 2. \quad (9)$$

The back-projected score of voxel k is null, if k is not observed by at least two viewpoints. The back-projected score is null also if $s(r) = 0$ for any $r \in \rho_k$. If all $r \in \rho_k$ have same score s , then $B_k = s$. The result of this operation is immediate carving-out of voxles that contradict support for clouds.

Different cloud regions have different gray shades. Ignoring these differences would wrongly allow matching of, say, a darker cloud-bottom to a bright sun-lit side of a cloud. To counter this, we require photometric consistency across viewpoints (the photo-hull concept in space-carving [?]). Clouds are diffuse, cooperating with this requirement. Per voxel k , the set of measured radiances is $\mathcal{I}_t(k) \equiv \{\hat{I}_t(r)\}_{r \in \rho_k}$. Across viewpoints, the measured variance of this set is $\text{VAR}[\mathcal{I}_t(k)]$, per t . Define a photo-consistency score

$$P_k = \exp \left(-\Sigma_{\text{color}} \{ \text{VAR}[\mathcal{I}_t^{\text{color}}(k)] / \sigma^2 \} \right), \quad (10)$$

where σ^2 is a scale parameter, and the index *color* refers to the red, green and blue channels. Overall, the total cloud-score of a voxel is $T_k = B_k P_k$.

The resulting 3D field $\{T_k\}_k$ is a volumetric estimate of cloud occupancy. It is biased to yield clouds larger than they really are: high-altitude voxels occluded by the cloud-base from all viewpoints are interpreted as being cloudy, since for them T_k is high. This is a realization of a basic ambiguity: if a voxel is occluded from all viewpoints, then there is no way of telling if it is cloudy or not, unless auxiliary or prior knowledge is available. Incorporating a visibility prior favors smaller clouds that explain the data. If voxel k is completely occluded by other cloudy voxels, then it is pruned (carved) out. Voxel k can only maintain its T_k if there are at least two camera viewpoints from which k is not occluded by other clouded voxels. This pruning is achieved by sweeping [?] the field $\{T_k\}_k$ iteratively. The pruned 3D cloud occupancy field is denoted $\{\tilde{T}_k\}_k$. We can maintain the non-binary (fuzzy) nature or $\{\tilde{T}_k\}_k$. This way, it possesses the inherent semi-transparency and subtle ambiguity of clouds.

Basic cloud score

For a basic cloud score (step A), there are various functions in the literature. We considered several of them, including using the background derived in Sec. 5. There is certainly room for increased sophistication, including sky modeling and machine learning. However, a ratio of image readout at the red/blue color channels, $\tilde{I}^{\text{red}} / \tilde{I}^{\text{blue}}$, is

widely use in the literature [?,30]. Overall, we found it to be an effective feature in broad daylight: clouds are grey (unit red-blue ratio), and the cloudless sky is significantly biased to blue (below ≈ 0.8 red-blue ratio). Thus, for demonstrations in this paper, the cloud-score we used per ray (pixel) is

$$s(r) = \begin{cases} \frac{6[\tilde{I}^{\text{red}}(r)/\tilde{I}^{\text{blue}}(r)-0.8]}{0.2+\tilde{I}^{\text{red}}(r)/\tilde{I}^{\text{blue}}(r)} & \text{if } \tilde{I}^{\text{red}}(r)/\tilde{I}^{\text{blue}}(r) > 0.8 \\ 0 & \text{otherwise} \end{cases} \quad (11)$$

Here $s \in [0, 1]$, where either bound is achieved at gray clouds or blue sky, respectively. An example of applying this operator on an image is shown in Fig. 3.

Cloud Reconstruction Results

We applied the approach on various captured scenes. One scene had cumulus clouds, similar to those seen in Fig. 3. Cross-sections of the recovered 3D cloud-occupancy field $\{\tilde{T}_k\}_k$ are shown in Fig. 7. The lateral domain of the clouds is much larger than the camera network. Accounting for the altitude of the cameras above sea-level, the clouds mainly reside between 800m to 1450m above Sea-level. We used two indicators to validate the results. First, there are worldwide balloon-based radiosonde measurements including quantitative vertical humidity profiles. The radiosonde nearest to our network is 88km away, but on a similar coastal strip. It indicated a layer of high humidity that can yield clouds in the range [770, 1881]m above sea-level, consistent with our clouds.

Second, we cross-validated 3D recovery. We used four cameras (indexed a, b, c, d) out of five, for 3D estimation. Camera e was ignored. Then, we projected the estimated 3D cloud distribution into viewpoint e , and compared to the ground truth. The rendered image is as follows. *Ray casting* [?] of field $\{\tilde{T}_k\}_k$ is performed on a ray corresponding to $[\mathbf{l}_e, \boldsymbol{\theta}]$. Ray-casting aggregates $\{\tilde{T}_k\}_k$ on all voxels intersected by the ray. The result is a cloud-score image $w[\mathbf{l}_e, \boldsymbol{\theta}]$. To visualize $w[\mathbf{l}_e, \boldsymbol{\theta}]$, we used it as α -map to the estimated sky-background image (Eq. 8). The alpha-map is

$$\alpha[\mathbf{l}_e, \boldsymbol{\theta}] = \begin{cases} 2w[\mathbf{l}_e, \boldsymbol{\theta}] & \text{if } 2w[\mathbf{l}_e, \boldsymbol{\theta}] < 1 \\ 1 & \text{otherwise} \end{cases} \quad (12)$$

The rendered image is then $J[\mathbf{l}_e, \boldsymbol{\theta}] = \alpha[\mathbf{l}_e, \boldsymbol{\theta}] + (1 - \alpha[\mathbf{l}_e, \boldsymbol{\theta}])\text{SKY}(\boldsymbol{\theta})$. This image does not pretend to properly render clouds in their true shades and effect on the sky. It simply served to visualize the result, as shown in Fig. 7d. It can be compared to the true corresponding image $\hat{I}_t[\mathbf{l}_e, \boldsymbol{\theta}]$, in Fig. 7e.

This rendering relies on the redundancy of the network. Even if a viewpoint is blocked, much of its information can be derived using other viewpoints compounded with 3D recovery. This supports the proposed approach for static Sun-blocking (part of the sky missing from a view), described in Sec. 6.

Another scene had a layer of alto-cumulus clouds. Figure 8 shows sample frames from this scene, and cross-sections of the recovered 3D cloud-occupancy field $\{\tilde{T}_k\}_k$. Accounting for the altitude of the cameras above sea-level, these estimated clouds mainly reside on a horizontal layer 3450m above sea-level, and the layer is about a kilometer thick. This time, the radiosonde readout indicated a layer of high humidity that can yield clouds in height range [3072, 4180]m above sea-level. This is in strong agreement with our results.

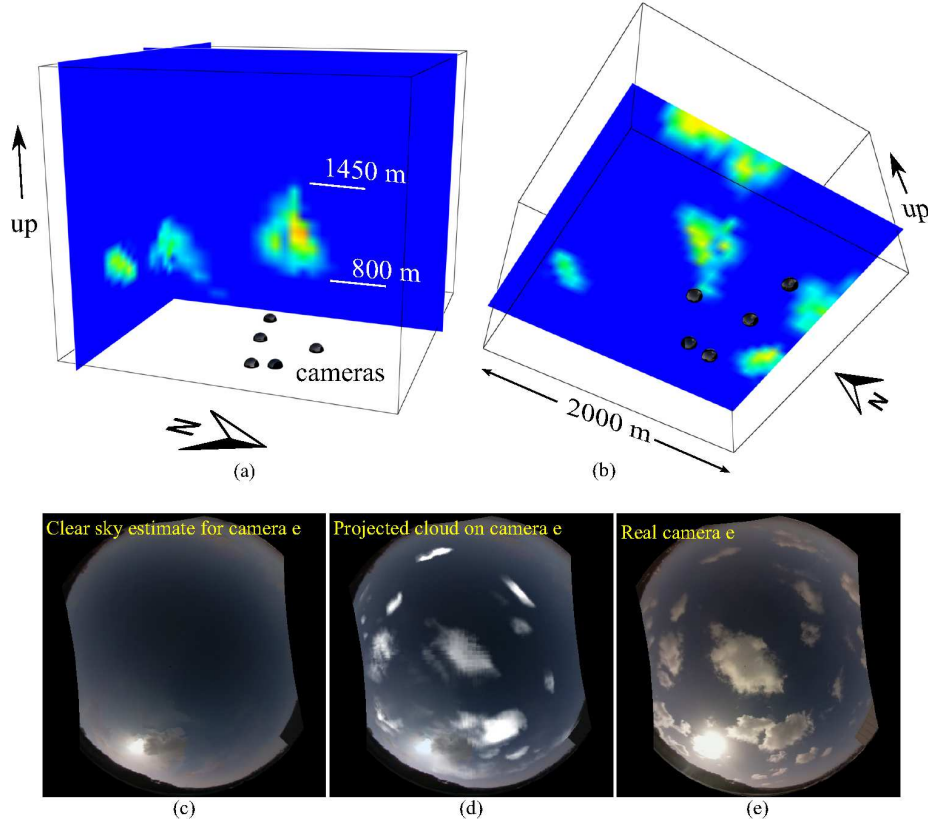


Fig. 7. 3D cumulus cloud recovery results. (a,b) Cross-sections of the recovered cloud-occupancy field $\{\tilde{T}_k\}_k$. Note how the domain of the clouds is much larger than the camera network. Cloud altitude is above sea-level. (c) Estimated sky-background image. Based on four viewpoints (indexed a, b, c, d), the 3D volumetric cloud-occupancy field $\{\tilde{T}_k\}_k$ was derived. The field $\{\tilde{T}_k\}_k$ was projected to viewpoint e , and overlaid on the estimated sky-background image. The resulting synthetic cloud-score image $J[\mathbf{l}_e, \boldsymbol{\theta}]$ is shown in (d). This can be compared to the real captured image $\hat{I}_t[\mathbf{l}_e, \boldsymbol{\theta}]$, shown in (e).

8 Discussion

Scaling light-field imaging hugely to sense the sky, would use a large network of camera nodes, each having a hemispherical field of view, deployed over a wide area. Such a network can reach anywhere wireless communication exists, and in principle can be a continent-wide light-field system. This sensing approach offers significant advantages over existing technologies (experimental and operational) of atmospheric sensing, particularly 3D imaging, and doing so in high spatio-temporal resolution. This sensing approach poses new questions for computer vision and computational photography. These include network-based extensions to monocular tasks including network-based radio-

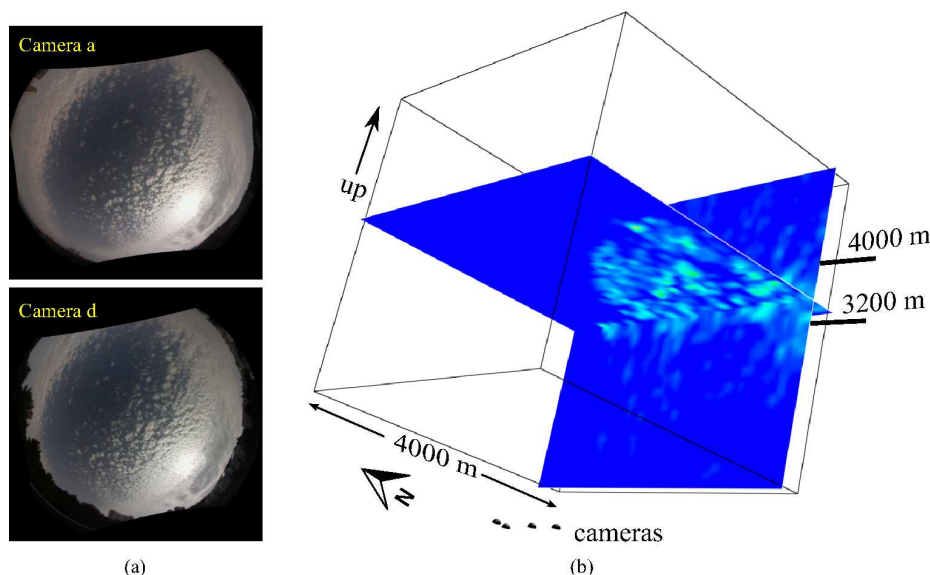


Fig. 8. 3D altocumulus cloud recovery results. (a,b) Sample frames. (c,d) Cross-sections of the recovered cloud-occupancy field $\{\tilde{T}_k\}_k$. Cloud altitude is above sea-level.

metric calibration and background estimation. Network redundancy offers the ability of by-passing saturation or blind spots, as those created by the sun, without moving parts.

To enable a massive network, each node should have very low-cost. To demonstrate this can work, we built the units in our experimental small-network using very basic components and coarse alignment tolerances. In a long term, mass producing consistent units would ensure proper alignment performance (non dead-regions) and operational robustness over long periods. Still, this concept opens the door to more interesting research. In the direction of the sun blocker, a diffraction grating can disperse direct sunlight (currently blocked) or skylight, to measure their spectra. This principle is already used by Aeronet [?] to characterize aerosols. Night-time operation is another interesting challenge: if a known planet and star appears missing on ray r , this indicated a cloud on r . Furthermore, such a light-field system can be used for studying lightnings and bird life, in 3D time-lapses.

References

1. Adelson, E., Wang, J.: Single lens stereo with a plenoptic camera. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **14** (1992) 99–106
2. Basha, T., Avidan, S., Hornung, a., Matusik, W.: Structure and motion from scene registration. *2012 IEEE Conference on Computer Vision and Pattern Recognition* (2012) 1426–1433
3. Bishop, T.E., Zanetti, S., Favaro, P.: Light field superresolution. *IEEE ICCP* **129** (2009) 1–9
4. Horstmeyer, R., Euliss, G., Athale, R., Levoy, M.: Flexible Multimodal Camera Using a Light Field Architecture. In: *ICCP*. (2009) 1–8

- 540 5. Levoy, M., Ng, R., Adams, A., Footer, M., Horowitz, M.: Light field microscopy. *ACM* 540
541 *Transactions on Graphics (TOG)* **25** (2006) 924—934 541
- 542 6. Kim, J., Lanman, D., Mukaigawa, Y., Raskar, R.: Descattering transmission via angular 542
543 filtering. In: *Proc. ECCV' 10*, Berlin, Heidelberg, Springer-Verlag (2010) 86–99 543
- 544 7. Diner, D.J., Martonchik, J.V.: Atmospheric transmittance from spacecraft using multiple 544
545 view angle imagery. *Appl. Opt.* **24** (1985) 3503–3511 545
- 546 8. Diner, D.J., Beckert, J.C., Reilly, T.H., Bruegge, C.J., Conel, J.E., Kahn, R.A., Martonchik, 546
547 J.V., Ackerman, T.P., Davies, R., Gerstl, S.A.: Multi-angle Imaging SpectroRadiometer 547
548 (MISR) instrument description and experiment overview. *IEEE Transactions on Geoscience* 548
549 *and Remote Sensing* **36** (1998) 1072–1087 549
- 550 9. Diner, D.J., Davis, A., Hancock, B., Gutt, G., Chipman, R.A., Cairns, B.: Dual-photoelastic- 550
551 modulator-based polarimetric imaging concept for aerosol remote sensing. *Appl. Opt.* **46** 551
552 (2007) 8428–8445 552
- 553 10. Diner, D.J., Davis, A., Hancock, B., Geier, S., Rheingans, B., Jovanovic, V., Bull, M., Rider, 553
554 D.M., Chipman, R.A., Mahler, A.B., McClain, S.C.: First results from a dual photoelastic- 554
555 modulator-based polarimetric camera. *Applied Optics* **49** (2010) 2929 555
- 556 11. Baxter, B., Hooper, B.A., Williams, J.Z., Dugan, J.P.: Polarimetric remote sensing of ocean 556
557 waves. In: *OCEANS 2009, MTS/IEEE Biloxi - Marine Technology for Our Future: Global* 557
558 *and Local Challenges.* (2009) 1–5 558
- 559 12. Brdon, E.M., Bréon, F.M.: An analytical model for the cloud-free atmosphere/ocean system 559
560 reflectance. *Remote Sensing of Environment* **43** (1993) 179–192 560
- 561 13. Mol, B.V., Ruddick, K., van Mol, B., K.Ruddick: The Compact High Resolution Imaging 561
562 Spectrometer (CHRIS): the future of hyperspectral satellite sensors . Imagery of Oostende 562
563 coastal and inland waters. In: *Proceedings of the Airborne Imaging Spectroscopy workshop,* 563
564 Brugge. Number October (2004) 564
- 565 14. Aides, A., Schechner, Y.Y., Holodovsky, V., Garay, M.J., Davis, A.B.: Multi sky-view 3D 565
566 aerosol distribution recovery. *Opt. Express* **21** (2013) 25820–25833 566
- 567 15. Litvinov, A., Schechner, Y.Y.: Addressing radiometric nonidealities: a unified framework. 567
568 In: *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society* 568
569 *Conference on. Volume 2.* (2005) 52–59 vol. 2 569
- 570 16. Kang, S., Weiss, R.: Can we calibrate a camera using an image of a flat, textureless Lamber- 570
571 tian surface? *Computer VisionECCV 2000* (2000) 640–653 571
- 572 17. Koreban, F., Schechner, Y.Y.: Geometry by deflaring. 2009 IEEE International Conference 572
573 on Computational Photography (ICCP) (2009) 1–8 573
- 574 18. Raskar, R., Agrawal, A., Wilson, C.a., Veeraraghavan, A.: Glare aware photography. *ACM* 574
575 *Transactions on Graphics* **27** (2008) 1 575
- 576 19. Talvala, E., Adams, A.: Veiling glare in high dynamic range imaging. *ACM Transactions on* 576
577 *Graphics (TOG)* **26** (2007) 1–10 577
- 578 20. Rouf, M., Mantiuk, R., Heidrich, W., Trentacoste, M., Lau, C.: Glare encoding of high 578
579 dynamic range images. *CVPR 2011* (2011) 289–296 579
- 580 21. Pust, N.J., Shaw, J.a.: Digital all-sky polarization imaging of partly cloudy skies. *Applied* 580
581 *optics* **47** (2008) H190–8 581
- 582 22. Bradley, E.S., Toomey, M.P., Still, C.J., Roberts, D.A.: Multi-Scale Sensor Fusion With an 582
583 Online Application: Integrating GOES, MODIS, and Webcam Imagery for Environmental 583
584 Monitoring. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote* 584
585 *Sensing* **3** (2010) 497–506 585
23. Allmen, M.C., Kegelmeyer Jr., P.: The computation of cloud base height from paired whole-
sky imaging cameras. *Machine Vision and Applications* **9** (1997) 160–165
24. Baumgarten, G., Fiedler, J., Fricke, K.H., Gerding, M., Hervig, M., Hoffmann, P., Müller, N.,
Pautet, P.D., Rapp, M., Robert, C., Rusch, D., von Savigny, C., Singer, W.: The noctilucent

585	cloud (NLC) display during the ECOMA/MASS sounding rocket flights on 3 August 2007:	585
586	morphology on global to local scales. <i>Annales Geophysicae</i> 27 (2009) 953–965	586
587	25. Cazorla, a., Olmo, F.J., Aladosarboledas, L., Alados-Arboledas, L.: Using a Sky Imager for	587
588	aerosol characterization. <i>Atmospheric Environment</i> 42 (2008) 2739–2745	588
589	26. Long, C.N., Sabburg, J.M., Calbo, J., Pages, D., Calbó, J., Pagès, D.: Retrieving cloud	589
590	characteristics from ground-based daytime color all-sky images. <i>Journal of Atmospheric</i>	590
591	and Oceanic Technology 23 (2006) 633–652	591
592	27. Seiz, G., Shields, J., Feister, U., Baltsavias, E.P., Gruen, A.: Cloud mapping with ground-	592
593	-based photogrammetric cameras. <i>Int. J. Remote Sens.</i> 28 (2007) 2001–2032	593
594	28. Kassianov, E., Long, C., Christy, J.: Cloud-base-height estimation from paired ground-based	594
595	hemispherical observations. <i>Journal of Applied Meteorology</i> 44 (2005) 1221–1233	595
596	29. Seiz, G., Davies, R.: Reconstruction of cloud geometry from multi-view satellite images.	596
597	<i>Remote Sensing of Environment</i> 100 (2006) 143–149	597
598	30. Seiz, G., Baltsavias, E., Gruen, A.A.: Cloud mapping from the ground: Use of photogram-	598
599	-metric methods. <i>Photogrammetric engineering and remote sensing</i> 68 (2002) 941–951	599
600		600
601		601
602		602
603		603
604		604
605		605
606		606
607		607
608		608
609		609
610		610
611		611
612		612
613		613
614		614
615		615
616		616
617		617
618		618
619		619
620		620
621		621
622		622
623		623
624		624
625		625
626		626
627		627
628		628
629		629