Supplementary information for

Air Quality Monitoring Using Mobile Microscopy and Machine Learning

Authors: Yi-Chen Wu[1,2,3], Ashutosh Shiledar[1], Yi-Cheng Li[1], Jeffrey Wong[4], Steve Feng[1,2,3], Xuan Chen[1], Christine Chen[1], Kevin Jin[1], Saba Janamian[1], Zhe Yang[1], Zachary Scott Ballard[1,2,3], Zoltán Göröcs[1,2,3], Alborz Feizi[1,2,3], and Aydogan Ozcan[1,2,3,5,*]

Affiliations:

[1] Electrical Engineering Department, University of California, Los Angeles, CA, 90095, USA.

[2] Bioengineering Department, University of California, Los Angeles, CA, 90095, USA.

[3] California NanoSystems Institute (CNSI), University of California, Los Angeles, CA, 90095, USA.

[4] Computer Science Department, University of California, Los Angeles, CA, 90095, USA.

[5] David Geffen School of Medicine, University of California, Los Angeles, CA, 90095, USA

[*]Correspondence: Prof. Aydogan Ozcan

E-mail: ozcan@ucla.edu

Address: 420 Westwood Plaza, Engr. IV 68-119, UCLA, Los Angeles, CA 90095, USA

Tel: +1(310)825-0915

Fax: +1(310)206-4685

Supplementary Information

System design overview

The c-Air platform (see Figs. 1(a–b) of main text) is composed of an impaction-based air-sampler (i.e., an impactor) and a lens-free holographic on-chip microscope, as illustrated in Figs. 1(c–d) of main text. The lens-free microscope consists of a partially coherent source with three fiber-coupled light-emitting diodes (LEDs)—at 470 nm, 527 nm, and 624 nm—which are used for illumination. It additionally includes a complementary metal-oxide semiconductor (CMOS) image sensor chip (OV5647, 5 MPix, 1.4 μm pixel size), which is placed ~0.4 mm below the transparent surface of the impactor with no image formation unit between them. As shown in Fig. 1(c) of main text, an embedded vacuum pump drives air through the air-sampler. Aerosols within the airflow are collected on a sticky coverslip of the air-sampler below the impaction nozzle. These collected particles are imaged and quantified by the lens-free mobile microscope.

Processing of the acquired c-Air images is remotely performed. As shown in Fig. 2 of main text, the overall system is composed of three parts: (1) the mobile c-Air device, which takes the air sample and records its image, (2) a remote server, which differentially processes the lens-free images (comparing the images taken before and after the intake of the sample air) to reconstruct phase and amplitude images of the newly captured particles and analyze them using machine learning, and (3) an iOS-based application (app), which controls the device, coordinates the air-sampling process, and displays the server-processed air sample images and corresponding particulate matter (PM) results. The design and functions of this app, as well as the remote server image processing algorithms, are detailed in subsequent sub-sections.

c-Air smartphone app

To control the c-Air device, we developed an iOS-based app using Swift. The app is installed on an iOS device (e.g., iPhone 6s) and is used together with the c-Air device. The app has two basic functions: (1) controlling the device for sampling air; and (2) displaying the results processed by the server. The app automatically establishes a personal hotspot upon launching, to which the device connects through Wi-Fi. Fig. 2(a) of main text presents screen-shots of the c-Air app interface. After the user launches the c-Air app, a welcome screen is presented with the following options: “measure air quality,” “view data,” “device,” and “about.” The “measure air quality” option switches the app to the measurement screen with a logo of the device in the middle. Selecting the logo triggers the air sampling process, as shown in Fig. 2(b) of main text, and the global positioning system (GPS) coordinates of the smartphone are recorded. After the air sampling process is complete, the lens-free holographic images obtained by the sampler are labeled with the GPS coordinates from the smartphone and transferred to a remote server for further processing. The app is designed to pair one smartphone to one c-Air device. To change the device that the app controls, the user can navigate to the screen (iii) and enter the IP address of the new sampler.

The same app is additionally used to view the server-processed results of the air samples captured from different locations. The full history of the samples obtained by this device can be accessed in (iv) “map view” or (v) “list view.” Selecting a sample in “list view” or a pinpoint in “map view” creates a summary of the server-processed results, as shown in (vi). The results can be viewed in two aspects using

the app: a reconstructed microscopic image of the captured aerosols on the substrate, with an option to zoom into individual particles, and a histogram of the particle size and density distribution.

Remote processing of c-Air images

Processing of the captured holographic c-Air images is performed on a Linux-based server (Intel Xeon E5-1630 3.70-GHz quad-core central processing unit, CPU) running Matlab. As illustrated in Fig. 2(c) of main text, the server processes the captured lens-free holographic images in the following steps. 1) It receives two sets of holograms (background and sample). 2) Pre-processing of these images is performed, and a differential hologram is obtained. 3) After holographic reconstruction of the resulting differential hologram, particle detection is conducted using particle peeling and machine learning algorithms, which are detailed below. 4) The detected particles are sized, and the particle size statistics are generated. 5) The reconstructed images and the processed PM statistics are transferred to the smartphone app for visualization.

In Step 1, the raw holograms, each approximately 5 MB in size, are transferred from the device to the server at 1 s or less per image using Wi-Fi. In Step 5, the processed information is packaged as a JPEG file of the reconstructed image plus a vector containing the particle density of each size range, which is later rendered and displayed as a histogram on the smartphone app.

Pre-processing and differential hologram formation

The server receives the two sets of holograms (background and sample) in three colors: red (R), green (G), and blue (B). For pre-processing before the hologram reconstruction and particle detection steps, we first extract the raw format images, which are then de-Bayered, i.e., only the information of the corresponding color channel is maintained. Next, to isolate the current aerosol sample collected during the latest sampling period, three differential holograms in R, G, and B are generated by digitally subtracting the corresponding background image from the sample image and normalizing it to a range of zero to two with a background mean centered at one.

Holographic reconstruction

A 2D distribution of captured particles, $O(x,y)$, can be reconstructed through digital propagation of its measured hologram, $A(x,y)$, to the image plane using the angular spectrum method:[1]

$$O(x,y) = F^{-1}\{F\{A(x,y)\} \cdot H(f_x, f_y)\} \quad \text{(S1)}$$

where F and F^{-1} define the spatial Fourier transform and its inverse, respectively, and $H(f_x, f_y)$ is the propagation kernel (i.e., the angular spectrum) in the Fourier domain, which is defined as:

$$H(f_x, f_y; \lambda, n, z) = \begin{cases} \exp\left[-j2\pi \frac{nz}{\lambda} \cdot \sqrt{1 - \left(\frac{\lambda}{n} \cdot f_x\right)^2 - \left(\frac{\lambda}{n} \cdot f_y\right)^2}\right], & if\ {f_x}^2 + {f_y}^2 \leq \left(\frac{n}{\lambda}\right)^2 \\ 0, & if\ {f_x}^2 + {f_y}^2 > \left(\frac{n}{\lambda}\right)^2 \end{cases} \quad \text{(S2)}$$

where f_x and f_y represent the spatial frequencies of the image in the Fourier domain. The propagation kernel $H(f_x, f_y)$ is uniquely defined, given the illumination wavelength λ, refractive index of the medium, n, and the propagation distance z. Without further clarification, all the propagation-related terms herein and in the main text refer to this angular-spectrum-based digital propagation of a complex object.

Digital auto-focusing on aerosols

In our lens-free on-chip imaging geometry, the specific distance from the sample to the sensor plane is usually unknown and must be digitally estimated for accurate reconstruction and particle analysis. Here, we use a digital measure, termed the Tamura coefficient,[2] for autofocusing and estimation of the vertical distance, z, from the particle to the sensor plane. It is defined as the square root of the standard deviation of an image over its mean:

$$T(I_z) = \sqrt{\frac{\sigma(I_z)}{\langle I_z \rangle}} \quad \text{(S3)}$$

where I_z is the intensity image of a hologram after propagation by distance z.

For speeding up this auto-focusing process, we employ a fast searching algorithm for this Tamura coefficient, which is illustrated in Supplementary Fig. S7. This searching method consists of two steps: finding a concave region around the Tamura coefficient peak, and performing a refined search in the concave region. In the first step, for an initial height interval of (l_0, u_0), we propagate the hologram to $4N + 1$ equally spaced vertical distances between l_0 and u_0. We then measure the corresponding Tamura coefficients at these heights. At this point, we check if the Tamura coefficient curve formed by these $4N + 1$ points is concave. If it is not, we determine the peak location and retain the $2N$ points around the peak. We then uniformly add $2N$ new points in the middle of the intervals defined by these $2N + 1$ points. The overall height interval that is searched is decreased by one-half in each step. This process is repeated until a concave height interval, (l, u), is identified.

In the second step, after a concave height interval (l, u) is identified, we run a golden-ratio search[3] to find the correct depth focus for each particle. To this end, we define $\alpha = (\sqrt{5} - 1)/2$ as the golden ratio[4] and let $p = u - \alpha(u - l)$ and $q = l + \alpha(u - l)$ be the golden-ratio division point on each side of a given height interval. After propagating the hologram to these new heights (p and q), we compare the Tamura coefficient, T_p and T_q, at these two heights, p and q, respectively. If $T_p < T_q$, we move the lower bound, $l = p$, and then let $p = q, T_p = T_q, q = l + \alpha(u - l)$. Otherwise, we move the upper bound, $u = q$, and then let $q = p, T_q = T_p$, and $p = u - \alpha(u - l)$. This process is repeated until the length of the height interval is smaller than a predefined threshold, $u - l < \delta$, e.g., $\delta = 0.1$ μm.

Detection and exclusion of moved particles

We conducted multiple differential imaging experiments on the same substrate. It was observed that a small number of aerosols (~3%) moved or changed their positions on the substrate in the later runs, as shown in Supplementary Fig. S8. In the reconstruction of the differential hologram, the moved particle appears as a pair of white–black points, where the "white" particle appears because it was present in the previous image but is absent in the current one. To avoid over-counting the aerosols on account of these moved particles, a threshold-based algorithm was used as part of the four peeling cycles to detect these "white" particles. Then, the nearest black particles that were similar in size and intensity to the detected white particles were marked to define a "moved particle". The moved particle was then removed from the total particle density distribution, thereby avoiding double counting of the same particle.

Converting particle count to particle density

For each sample (and the corresponding c-Air differential hologram), the particle detection and sizing algorithm, as described in the main text, provides the particle count in the number of particles for different sizes (diameters). To facilitate a more universal unit, we convert the sampling particle count, N_i, to particle density n_i (in $\text{counts} \cdot \text{L}^{-1}$) using the following equation:

$$n_i = \frac{N_i}{Q \cdot t} \cdot \frac{L_{sensor}}{L_{total}} \quad \text{(S4)}$$

where Q = 13 L·min^{-1} is the flow rate of air, and $t = 0.5$ min is the typical sampling duration. In addition, $L_{total} = 15.5$ mm is the total length of the impactor nozzle slit, and $L_{sensor} = 3.67$ mm is the part of the slit being imaged, which equals the longer edge of the CMOS sensor active area. The conversion equation here assumes that the particle distribution is uniform along the sampler nozzle length, which is a valid assumption because the nozzle tapering is in its orthogonal direction, while the structure of the sampler in this direction is spatially invariant.

References for supplementary information

1. Goodman, J. W. *Introduction to Fourier optics, 3rd Edition*. (Roberts and Company Publishers, 2005).
2. Memmolo, P. *et al.* Automatic focusing in digital holography and its application to stretched holograms. *Opt. Lett.* **36,** 1945–1947 (2011).
3. Golden section search. *Wikipedia* (2016). Available at: https://en.wikipedia.org/w/index.php?title=Golden_section_search&oldid=745658475. (Accessed: 14th November 2016)
4. Golden ratio. *Wikipedia* (2016). Available at: https://en.wikipedia.org/w/index.php?title=Golden_ratio&oldid=746669942. (Accessed: 14th November 2016)

Supplementary Tables:

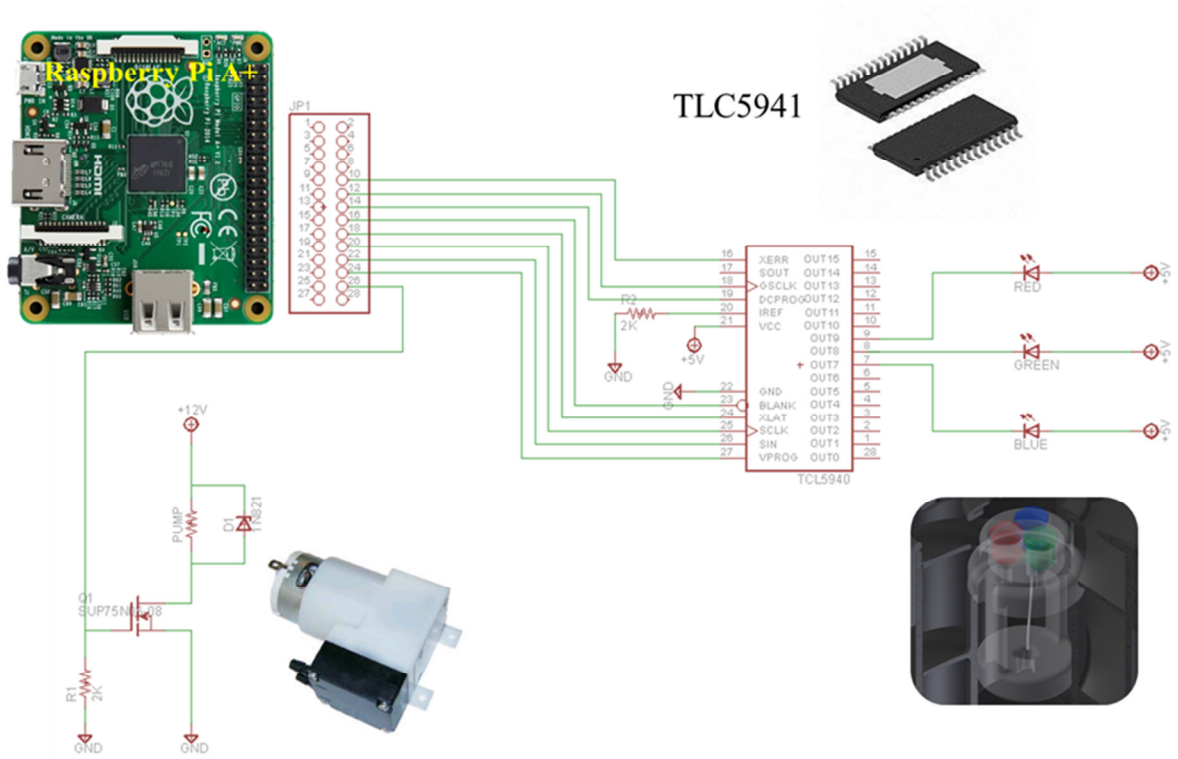
Device A		*Class-100 Clean Room*	*Class-1000 Clean Room*	*Indoor*	*Outdoor*
Total (count/L)	mean μ	10.76	27.32	114.14	196.52
	std σ	3.34	27.94	17.09	55.25
PM10 (count/L)	mean μ	10.76	26.92	111.59	195.84
	std σ	3.34	27.41	17.22	55.43
PM2.5 (count/L)	mean μ	7.95	14.88	67.99	113.36
	std σ	2.87	14.33	12.03	43.45
Device B		***Class-100 Clean Room***	***Class-1000 Clean Room***	***Indoor***	***Outdoor***
Total (count/L)	mean μ	6.14	23.57	151.68	190.89
	std σ	6.17	22.02	67.8	18
PM10 (count/L)	mean μ	6.14	23.57	147.76	190.79
	std σ	6.17	22.02	67.01	18.06
PM2.5 (count/L)	mean μ	4.4	16.19	89.93	116.59
	std σ	3.24	14.67	39.31	13.53

Supplementary Table S1. Intra-device repeatability.

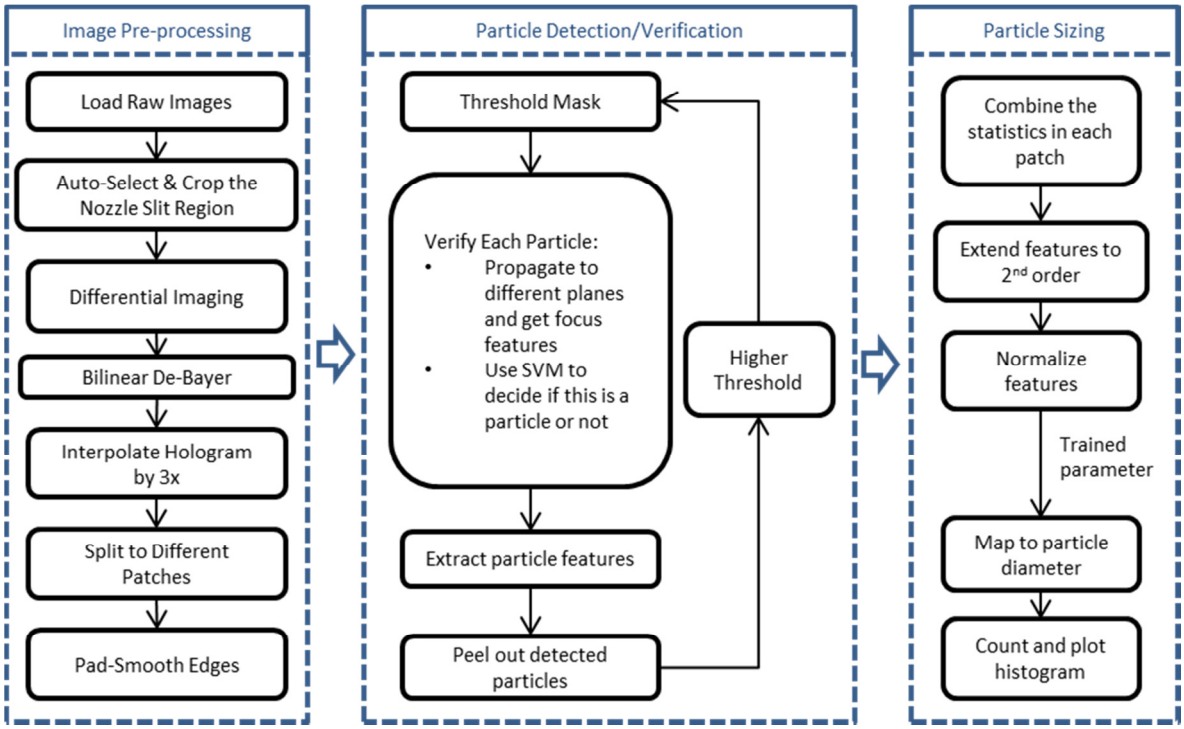
		Class-100 Clean Room	*Class-1000 Clean Room*	*Indoor*	*Outdoor*
Total (count/L)	**H**	0	0	0	0
	P-value	0.07	0.46	0.13	0.53
PM10 (count/L)	**H**	0	0	0	0
	P-value	0.07	0.46	0.21	0.53
PM2.5 (count/L)	**H**	0	0	0	0
	P-value	0.09	0.43	0.16	0.25

Supplementary Table S2. Inter-device repeatability.

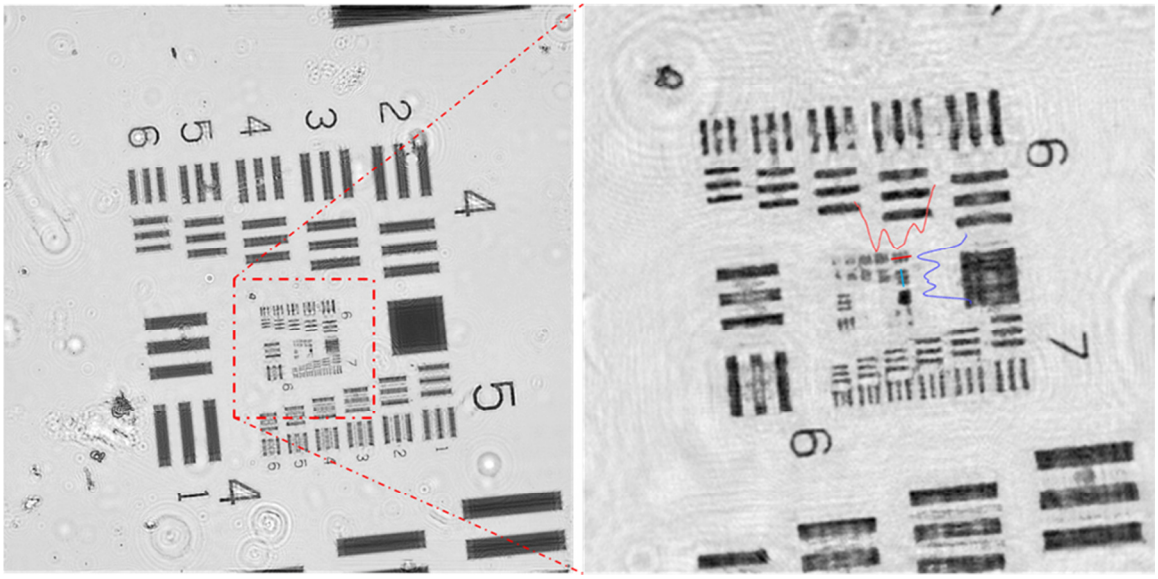
Supplementary Figures



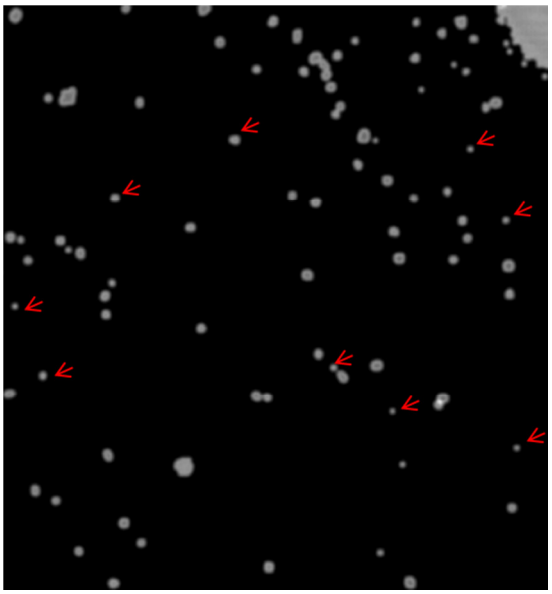
Supplementary Fig. S1. Controlling circuit of the pump and fiber coupled LEDs, based on Raspberry Pi A+ as the core micro-computer.



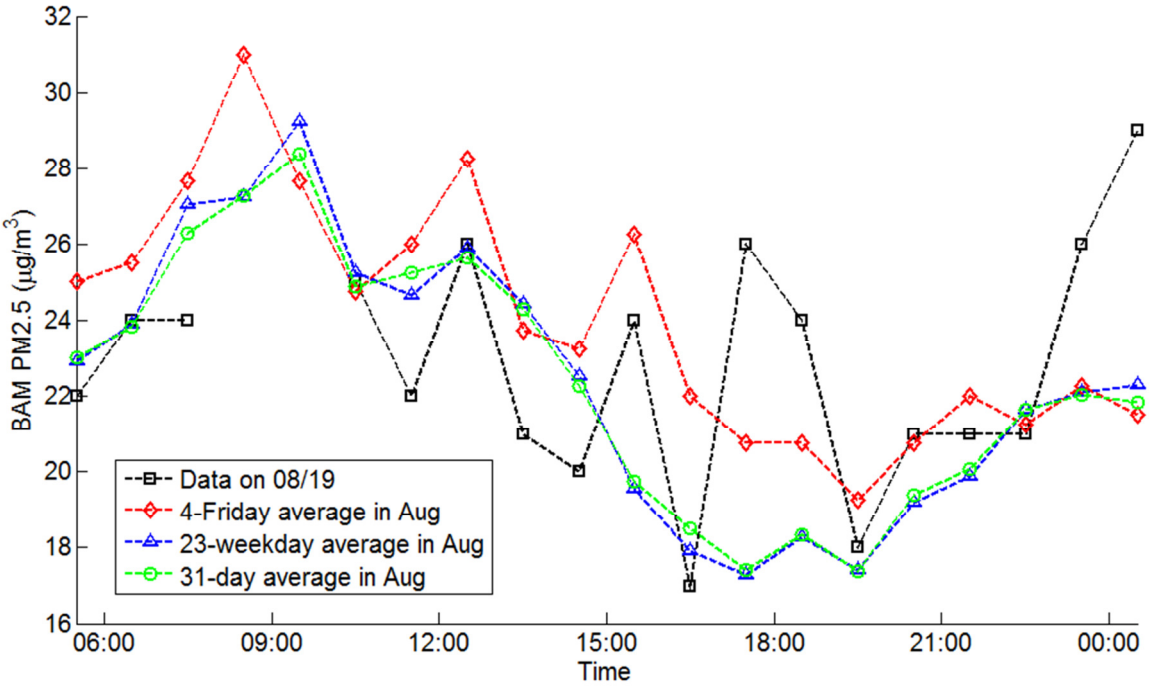
Supplementary Fig. S2. Particle detection and counting algorithm at the server side.



Supplementary Fig. S3. c-Air spatial resolution test. USAF-1951 resolution test target was used to quantify the spatial resolution of the c-Air platform. The reconstructed image of this test target is shown here. The finest resolved feature is group 8, element 2, which has a line-width of 1.74 µm.

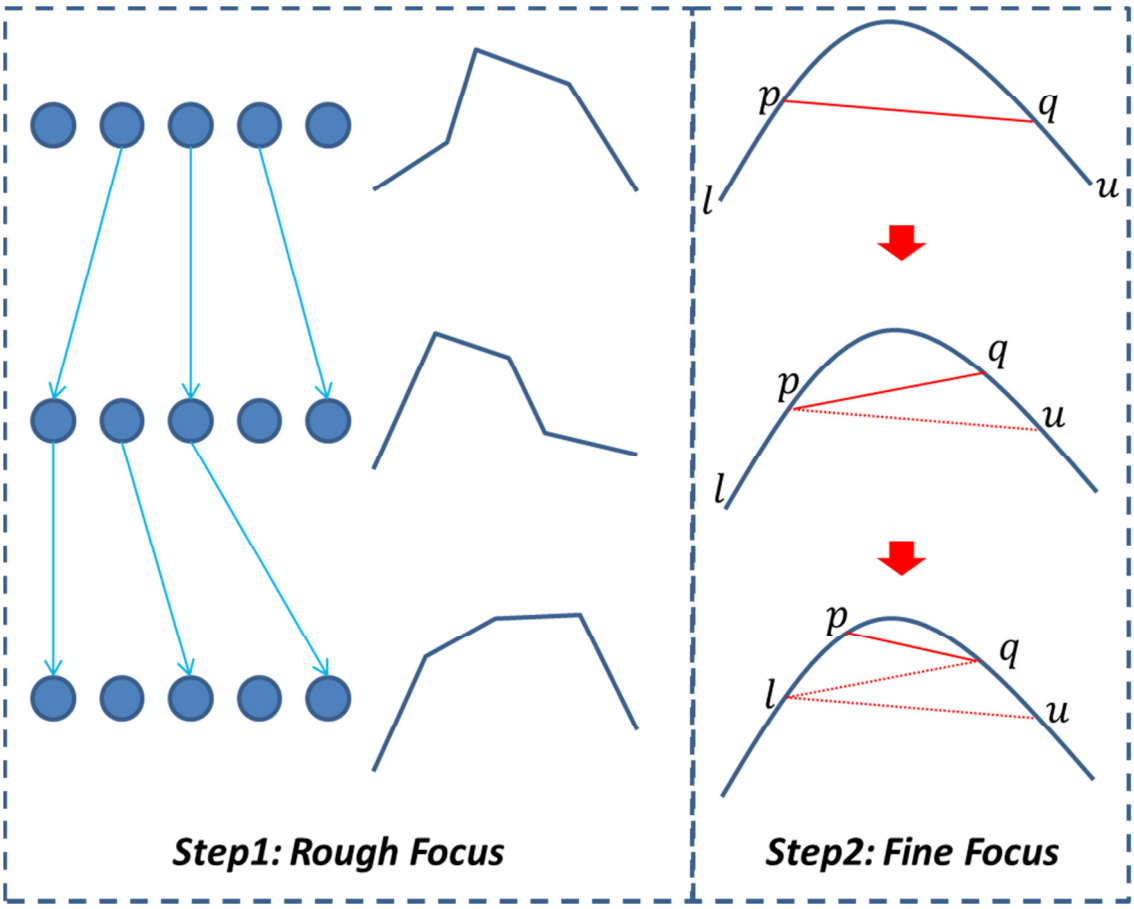


Supplementary Fig. S4. Detection of 1 µm particles using c-Air. Particles detected using c-Air during a digital peeling cycle are shown – refer to the Methods section of the main text for further details. The locations of 1 µm particles, which were independently verified with a 100x objective-lens (NA=0.95), are marked by red arrows. Also refer to Fig. 3 of main text for c-Air particle sizing accuracy.

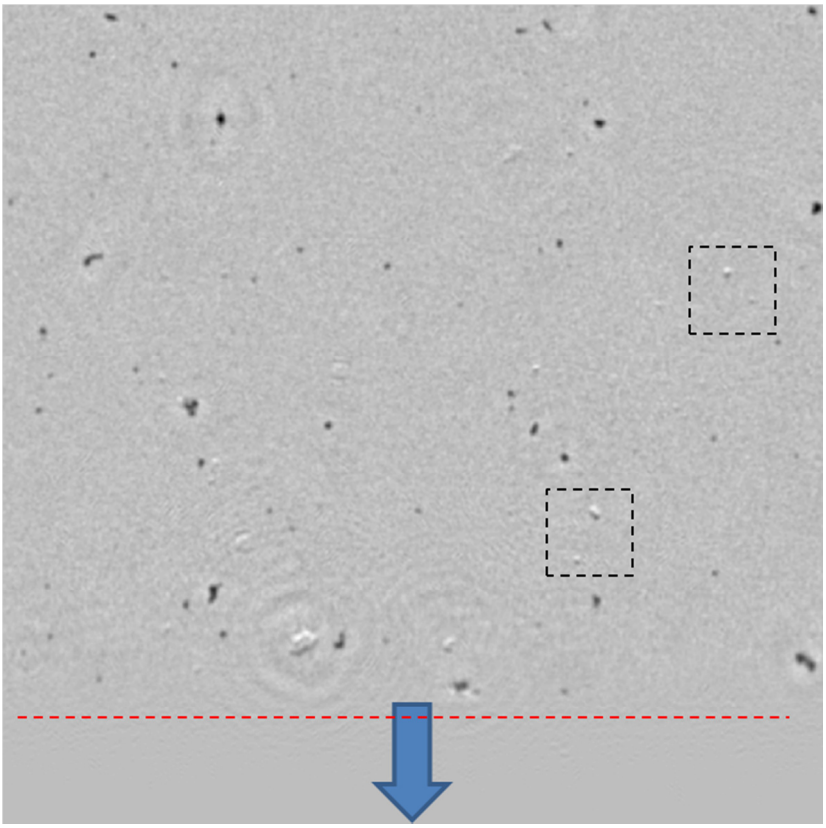


Supplementary Fig. S5. Averaged BAM PM2.5 daily variation measured at Reseda station in August, 2016. BAM PM2.5 measurements on 08/19 (black curve) between 7 am and 9 am are missing because of the maintenance of the device.

Supplementary Fig. S6. (a) Google Maps image (accessed on 10/06/2016) of the locations near our measurement Route I. (b) Zoomed-in region of the Google Maps image near our third measurement location (marked by a yellow star) shows a large parking lot with ~3,400 car spots. (c) Google earth map (accessed on 12/19/2016) of the green dashed region in (a) showing the construction zone. (d)-(f) Photos of the locations marked in (c) showing the construction at this location. This parking lot and the related construction nearby might have caused an additional source of air pollution as also measured by c-Air, see Figs. 7(b,c) of the main text.



Supplementary Fig. S7. Auto-focusing fast search algorithm.



Supplementary Fig. S8. Detection of moved particles in c-Air data. A cropped region of the reconstructed c-Air image is shown. The boundary of the imaging cell is marked by a dashed red line while the direction of the air flow is denoted by a blue arrow. The regions marked by dashed squares show some examples of moved particles that are automatically detected to avoid over-counting of particles. Refer to the Methods section of the main text for further details.