Computational Imaging on the Electric Grid

Mark Sheinin, Yoav Y. Schechner
Viterbi Faculty of Electrical Engineering
Technion - Israel Institute of Technology
markshenin@gmail.com, yoav@ee.technion.ac.il

Kiriakos N. Kutulakos
Dept. of Computer Science
University of Toronto
kyros@cs.toronto.edu

Abstract

Night beats with alternating current (AC) illumination. By passively sensing this beat, we reveal new scene information which includes: the type of bulbs in the scene, the phases of the electric grid up to city scale, and the light transport matrix. This information yields unmixing of reflections and semi-reflections, nocturnal high dynamic range, and scene rendering with bulbs not observed during acquisition. The latter is facilitated by a database of bulb response functions *for a range of sources, which we collected and provide. To do all this, we built a novel coded-exposure high-dynamic-range imaging technique, specifically designed to operate on the grid's AC lighting.*

1. Introduction

For more than a century we have been living in a world that literally pulses with artificial light. Whether outdoors at night or indoors at all hours, most of the light reaching our eyes—and our cameras—originates from artificial sources powered by the electric grid. These light sources change their intensity and spectral power distribution in response to the grid's alternating current (AC) [3, 43] but their flicker is usually too subtle and too fast to notice with the naked eye (100Hz or more) [22]. Artificial lighting produces unnatural-looking colors in photos [13] and temporal aliasing in video [39]. As a result, it is broadly considered undesirable [16, 21, 40, 48].

In this paper we argue that rather than being a mere nuisance, ubiquitous AC-induced lighting variations are a very powerful visual cue—about our indoor and outdoor environments, about the light sources they contain, and the electrical grid itself (Figure 1). To this end, we derive a model of time-varying appearance under AC lighting and describe a novel coded-exposure imaging technique to acquire it.

Our approach yields several never-seen-before capabilities that we demonstrate experimentally with our "ACam" camera prototype: (1) acquiring a scene's transport matrix by passive observation only, (2) computing what a scene would look like if some of its lights were turned off or changed to a different bulb type, (3) recognizing bulb types from their temporal profiles, (4) analyzing city-scale grid phases in the electric grid, and (5) doing all the above under very challenging conditions—nocturnal imaging, an off-the-shelf (30Hz) camera, dimly-lit scenes, uncontrolled environments, distances of meters to kilometers, and operation in two continents using both 110V and 220V AC standards. To enable all this, we compiled a database [35, 36] of temporal lighting response functions (DELIGHT) for a range of bulb types, the first of its kind in computer vision. The only essential constraints in our approach are access to a power outlet and a largely stationary scene.

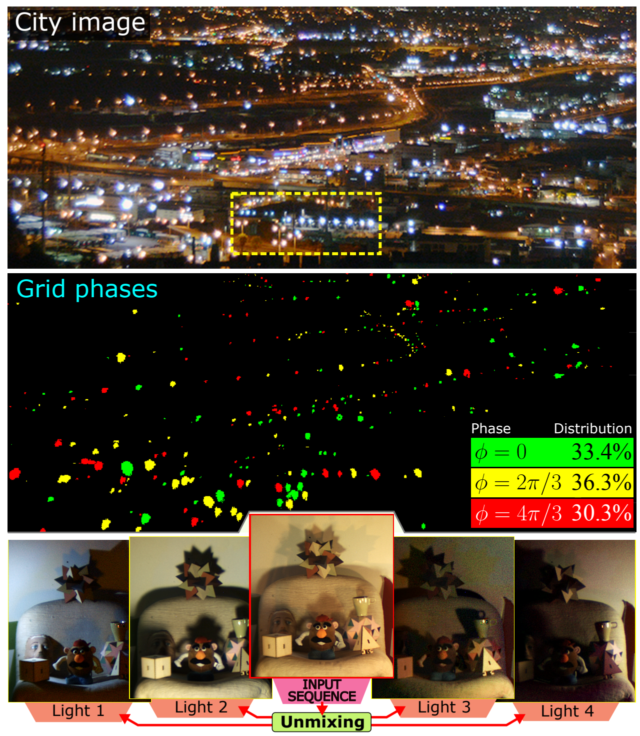

Figure 1. **Top and middle:** City-scale scene. From the AC bulb response function, we can recognize which bulbs are used and the electric-grid phase of each (color coded). This enables statistical analysis of the grid. Inset is magnified in Figure 3. **Bottom:** Unmixing a scene to single-light-source component images.

Our work draws inspiration from the large body of research on actively-controlled light sources. These techniques illuminate a scene with a variety of sources (*e.g.*, projectors [18, 24, 32], lasers [11, 25], computer displays [49], flashes [28] and arrays of point sources [8, 45], *etc.*) in order to impose predictable structure on an otherwise-unstructured visual world. Two lines of research in this area are particularly close to ours. First, methods for computational light transport [7] express the linear relation between controllable light sources and images as a transport matrix that can be acquired [33] or probed [26]. We adopt and extend the transport matrix formulation to AC light sources, and demonstrate scene re-lighting without any access to a programmable source. Second, recent work has treated the imaging process as a radio-like communication channel between active sources and cameras [15, 19]. These techniques transmit periodic signals between lights and cameras at high speed but, like all active methods, they reject ambient AC light rather than use it.

The key observation behind our work is that ambient AC lighting has a great deal of structure already. This is because of two fortunate facts: (1) AC light sources often do not flicker with the same phase even if located in the same space and (2) their temporal intensity profile is different depending on bulb type, make and model. The former comes from a desire to spread evenly the three phases of AC across light sources in order to balance load on the grid and make flicker even less noticeable [43]. The latter comes from differences in power circuitry and in the mechanism of light emission (fluorescence [47], incandescence [4], LED, *etc.*) Thus, the light arriving at a camera pixel is a mixture of differently-shifted and potentially very diverse signals: even among household LED bulbs, we have observed modulations down to 10% of maximum intensity in some products and near-constant intensity in others. The precise mixture of these light signals differs from pixel to pixel in accordance with the scene's light transport properties.

Here we undertake a first systematic study of how to passively record, untangle, and use these signals. In this sense, our work is another example of exploiting visual cues "hidden in plain sight." [41, 46]

The AC cue is very difficult to acquire with high-speed cameras because there is seldom enough light to record useful images at the speed we would need (over 1000 frames per second). On the other hand, long-exposure photography is not an option either because of the cue's transient nature. To overcome these challenges we use a novel coded-exposure imaging [12, 26, 31, 42] technique. Our ACam acquires high-dynamic-range (HDR) images corresponding to fractions of the AC cycle by capturing long exposures while masking and unmasking pixels individually at 2.7kHz, in sync with the AC.

2. Alternating-Current Illumination

2.1. Alternating Current in The Grid

We now describe a model of AC-modulated lighting. Power suppliers strive for a zero-mean sinusoidal AC voltage having a regular peak outlet amplitude[1] $V_{\max}$. There are two exclusive standards, having nominal frequencies 50Hz and 60Hz. The Americas use the former, while Asia and Europe mainly use the latter. Imperfections in electricity generation slightly wiggle the AC frequency randomly. Hence, the AC is *quasi periodic*: for a short time span, the effective frequency is a perturbation of the nominal frequency. The wiggle is practically *spatially invariant* in spatiotemporal scales typical to computer vision: the temporary frequency of the AC is essentially the same in any electrical outlet across the city. The reason is that electricity perturbations propagate at a speed on the order of the speed of light.

In practice, the temporary frequency of the AC is determined from the time interval Δ between two successive zero crossings (Figure 2[top-left]). Since there are two such crossings per period of the AC, its frequency is given by

$$f \;=\; 1/(2\Delta)\,. \tag{1}$$

The electric grid carries AC in a discrete set $\mathbb{P}$ of *grid phases*, using distinct, exclusive sets of cables. In most scenes there are three such phases spaced $2\pi/3$ apart. Each outlet is connected to one of these grid phases. In our labs, we declared one outlet to be the reference, having phase $\phi = 0$. Hence, $\mathbb{P} = \{0, 2\pi/3, 4\pi/3\}$ (see Figure 1).[2] Now suppose we count time t with a stopwatch, beginning from some negative-to-positive zero crossing of the voltage at the reference outlet (Figure 2[top-left]). The AC voltage is then

$$V(t) \;=\; V_{\max}\sin(2\pi f t - \phi)\,. \tag{2}$$

2.2. From AC Electricity to Light

A bulb β is a system whose input is the voltage $V(t)$ and its output is spectral flux $L_\beta(t,\lambda)$, where λ denotes wavelength. Hypothesize for a moment a bulb which is electrically linear, *i.e.,* the current $J(t)$ satisfies a proportionality $J(t) \propto V(t)$. Then, hypothesize that this bulb is *unmediated*, converting electric power $J(t)V(t) \propto V^2(t)$ to flux directly and instantaneously. Thus, the spectral flux $L_\beta(t,\lambda)$ is equivalent to $V^2(t)$. Consequently, the hypothetical bulb flickers at double the AC frequency and becomes dark whenever $V(t)$ goes to zero. We call this flickering period a *cycle*, whose duration is Δ.

In practice, the transition from electricity to radiance is *mediated* by various mechanisms. Optical mediators include heat, gas discharge and phosphorescence. Non-incandescent bulbs generally have electronic components inside the bulb fixture, to which the lay person is oblivious. These components (diodes, inductors, etc.) mediate

[1] Depending on the country $V_{\max}$ is 170 or 312 zero-to-peak Volts, yielding a root-mean-squared voltage of 120 or 220 Volts, respectively.

[2] Some regions may be linked by a distribution transformer that shifts all phases by a constant ϕ' [9]. If the scene contains two such subregions then $\mathbb{P} = \{0, 2\pi/3, 4\pi/3, \phi', 2\pi/3 + \phi', 4\pi/3 + \phi'\}$. We did not encounter such a scene in our experiments.

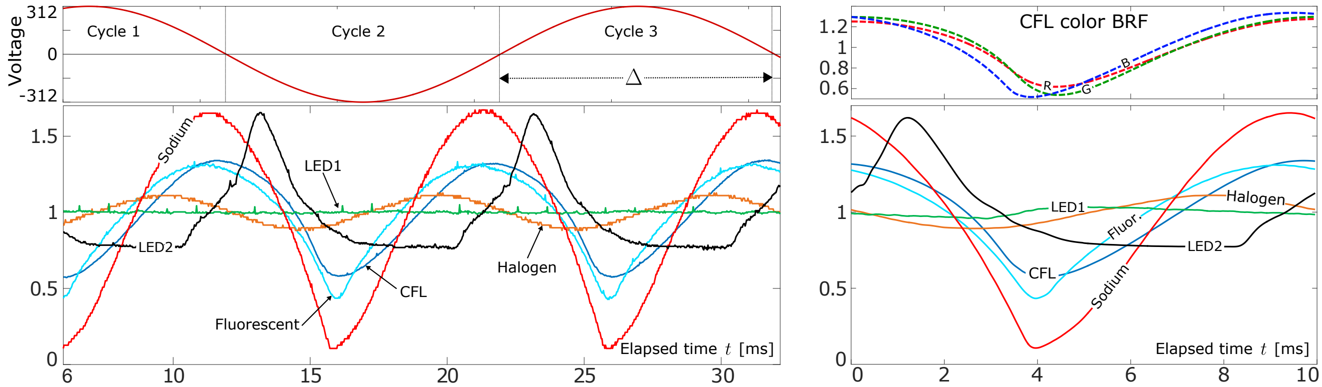

Figure 2. **Top left:** A pure sine wave, fitted to the raw voltage at the reference outlet in Haifa. We count time t starting from a negative-to-positive zero crossing at this outlet. **Bottom left:** Raw signals from a bare photodiode for a sample of the bulbs in DELIGHT, spanning multiple cycles. Each signal was normalized by its temporal average over one flicker cycle. For better visualization, LED2's waveform was attenuated by $1/3$. **Bottom right:** The corresponding monochrome BRFs. These were computed by acquiring signals like those on the bottom-left 200 times, averaging them, and cropping them at $t \in [0, \Delta]$. Here LED1's BRF was amplified $\times 10$ to illustrate that it is not actually constant. **Top right:** The three-band BRF of one of the bulbs, measured by placing color filters in front of the photodiode.

between voltage and spectral flux. Mediators have response times and nonlinearities. Hence the function $L_\beta(t, \lambda)$ is a distortion of $V^2(t)$: there is a delay, and $L_\beta(t, \lambda)$ generally does not go to zero during a cycle.

Denote by $\mathbb{B}$ the finite set of bulbs in use. Consider a bulb $\beta \in \mathbb{B}$, such as a particular fluorescent bulb in a brand fixture, whose time-averaged spectral flux over one cycle is $\overline{L_\beta}(\lambda)$. Relative to this average, at time t the bulb emission fluctuates as:

$$L_\beta(t, \lambda) \;=\; \overline{L_\beta}(\lambda)\; B_\beta(t, \lambda) \;\; . \tag{3}$$

We define the unit-less function $B_\beta(t, \lambda)$ to be the *spectral bulb response function (SBRF)*. This function has a time average of 1 for each wavelength and serves as an intrinsic model of a bulb's temporal behavior.

Acquiring a lamp's SBRF requires specialized equipment like integrating spheres and high-speed spectrometers. As such, measuring the SBRF directly is rather involved. A more practical model of bulb behavior is to consider the time-varying measurements from a camera or photodiode placed nearby (with or without color filters):

$$I_\beta(t, \sigma) \;=\; \overline{I_\beta}(\sigma)\; B^*_\beta(t, \sigma) \;\; . \tag{4}$$

Here $I_\beta(t, \sigma)$ is the intensity measured at a pixel or photodiode at time t and spectral band σ, $\overline{I_\beta}(\sigma)$ is its temporal average and $B^*_\beta(t, \sigma)$ is the unit-less *bulb response function (BRF)*. Unlike the SBRF, the BRF depends on the placement and spectral sensitivity of the device used.[3]

In general, both the SBRF and the BRF may exhibit a slightly different temporal profile across cycles (*e.g.*, due to voltage polarity, warm-up period, ambient temperature, *etc.*) Here we ignore these secondary effects for the sake of simplicity, treating BRFs as essentially invariant to the number of cycles since time zero. Thus, our BRFs are fully specified by their values in very first cycle. In the following we restrict t to lie in the interval $[0, \Delta]$ and treat the BRF as a function that is defined over just that interval.

Cameras and photodiodes provide discrete samples of the continuous intensity $I_\beta(t, \sigma)$. Suppose $I_\beta(t, \sigma)$ is resolved into K samples within a cycle. These samples correspond to integrals of $I_\beta(t, \sigma)$ over consecutive time intervals of duration Δ/K. Thus, Eq. (4) becomes

$$\mathbf{i}_\beta(\sigma) \;=\; \overline{I_\beta}(\sigma)\; \mathbf{b}_\beta(\sigma) \tag{5}$$

$$= \underbrace{\left(\sum_\sigma \overline{I_\beta}(\sigma)\right)}_{\text{brightness}} \;\underbrace{\frac{\overline{I_\beta}(\sigma)}{\sum_\sigma \overline{I_\beta}(\sigma)}}_{\text{chromaticity } Q_\beta(\sigma)} \;\underbrace{\mathbf{b}_\beta(\sigma)}_{\text{BRF}} \tag{6}$$

where the K-dimensional row vectors $\mathbf{i}_\beta(\sigma)$ and $\mathbf{b}_\beta(\sigma)$ hold the intensity and BRF samples, respectively.

Figure 2 shows several examples of sampled BRFs. As can be seen, all bulbs flicker at double the AC frequency and are locked to individual cycles.

3. The DELIGHT Database of Bulb Responses

For the tasks in Sections 4 and 5 we created a Database of Electric LIGHTs (DELIGHT). We acquired a variety of bulbs and fixtures. Street lighting is dominated by a few bulb types, mainly high pressure sodium, metal halide, mercury and fluorescent. Each streetlight type is used rather consistently in large areas. Indoor lighting has higher variety, including halogen, fluorescent tubes, different compact fluorescent lamps (CFLs) and simple incandescent. LED lighting has an interesting variety of BRFs, some having very low and some very high BRF amplitudes (Figure 2).

[3] Specifically, the spectral flux and measured intensity are related by the integral $I_\beta(t, \sigma) = G \int L_\beta(t, \lambda) R(\sigma, \lambda) d\lambda$ where $R(\sigma, \lambda)$ is the sensor's spectral sensitivity. The geometric factor G converts emitted flux to pixel/photodiode intensity and depends on their placement, aperture, *etc.*

To keep t common to all BRFs, DELIGHT was acquired by connecting all bulbs and fixtures to a single 50Hz reference outlet in Haifa. The AC voltage $V(t)$ was simultaneously measured at this outlet. We used three sensing schemes: (1) a photodiode with one of three color filters; (2) the same photodiode without any filters; and (3) our ACam prototype described in Section 6, fitted with a color camera. For schemes (1) and (3), we save in DELIGHT the BRF of individual bulbs and their chromaticity. For scheme (2) only a monochrome BRF is saved. In all cases, metadata such as bulb wattage and sensor/filter used are stored as well. See [34, 35, 36] for more information.

4. Recognizing AC Lights and Grid Phase

Let us point a camera at a bulb in the scene. The measured signal $\mathbf{i}(\sigma)$ follows Eq. (6). This signal is normalized by the mean brightness, yielding $\mathbf{i}^{\mathrm{norm}}(\sigma)$. Now, all temporal variations are due to the bulb's BRF, chromaticity and grid phase. We recognize the bulb and its phase using:

$$\{\hat{\beta}, \hat{\phi}\} = \underset{\beta\in\mathbb{B},\phi\in\mathbb{P}}{\arg\min} \sum_{\sigma} \left\| \mathbf{i}^{\mathrm{norm}}(\sigma) - Q_\beta(\sigma)\,\mathrm{shift}\left(\phi, \mathbf{b}_\beta(\sigma)\right) \right\|^2 \tag{7}$$

where $\mathbb{B}$ is the set of bulbs in DELIGHT, $\mathbb{P}$ is the set of possible grid phases, $Q_\beta(\sigma)$ is the chromaticity of bulb β in the database, and $\mathrm{shift}()$ circularly shifts to the right the bulb's sampled BRF by phase ϕ. When using a monochrome camera, there is only one spectral band so $Q_\beta(\sigma) = 1$.

Figures 1 and 3 show results from Haifa Bay, where the ACam was fitted with a monochrome camera. In this metropolitan scale, we recognize the bulb types and their three grid phases. Simple analysis shows that the distribution of grid phases is approximately uniform over the bulbs detected in the field of view.

5. Theory of AC Light Transport

To simplify notation, we drop the spectral band σ wherever we can. A scene contains static objects and is illuminated by S light sources. It is observed by a camera having a linear radiometric response and P pixels. As in Section 2.2, we resolve the time-varying image into K frames.

Now suppose only source s is on, with chromaticity Q_s, BRF $\mathbf{b}_s$ and phase 0. Furthermore, suppose matrix $\mathbf{I}_s$ holds the resulting *single-source image sequence*. Each column of $\mathbf{I}_s$ is a frame and each row is the intensity of one pixel through time. At frame k pixel p's intensity follows Eq. (6):

$$\mathbf{I}_s[p,k] \;=\; \tau_{ps}\, Q_s \mathbf{b}_s[k] \tag{8}$$

where brackets denote individual elements of $\mathbf{I}_s$ and $\mathbf{b}_s$. The factor τ_{ps} expresses *light transport*. This factor specifies the total flux transported from source s to pixel p via all possible paths. This transport encapsulates global factors such as the camera's numerical aperture and spectral

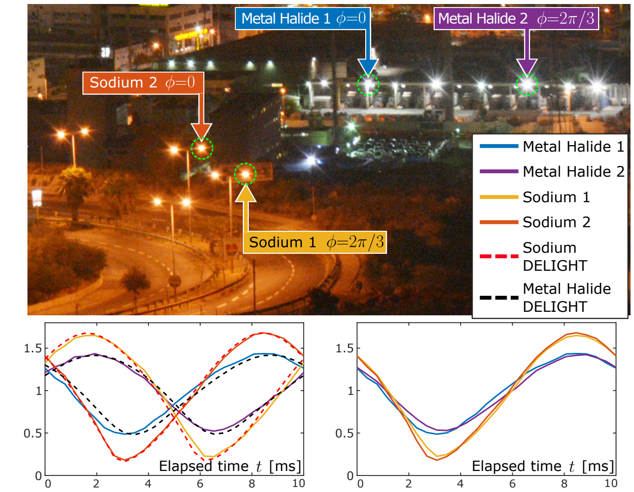

Figure 3. **Top:** Close-up of Haifa bay from Figure 1. Kilometers away, lamps are recognized from DELIGHT and ACam-captured images. In conjunction, the grid phase at each bulb is recognized. **Bottom left:** Raw signals normalized by mean brightness, plotted along with their best-matching BRFs from DELIGHT. Colors correspond to the bulbs indicated above. **Bottom right:** To illustrate how well the measured signals on the bottom left compare to each other, we shifted them by minus their recognized phase, effectively placing them all on grid phase zero. Observe that the signals from bulbs of the same type are indeed very similar.

response; spatial and angular variations in radiance at pixel p by source s; the BRDF at p when illuminated by s; shadows, inter-reflections, *etc.*

Expressing Eq. (8) in matrix form we obtain:

$$\mathbf{I}_s \;=\; \boldsymbol{\tau}_s\, Q_s \mathbf{b}_s\,. \tag{9}$$

Here column vector $\boldsymbol{\tau}_s$ concatenates the transport factors of all pixels for source s. It follows that individual frames of the sequence are just scalings of vector $\boldsymbol{\tau}_s$.

Now, the scene is illuminated by S sources connected to phase zero. The image sequence becomes a superposition of S single-source sequences, one per source s:

$$\mathbf{I} \;=\; \mathbf{I}_1 + \cdots + \mathbf{I}_s + \cdots + \mathbf{I}_S\,. \tag{10}$$

Suppose the chromaticities and BRFs of these sources are $\mathbf{b}_1, \ldots, \mathbf{b}_S$ and $Q_1, \ldots, Q_S$, respectively. Combining Eqs. (9) and (10), factorizing various terms and denoting $^\top$ for transpose we obtain

$$\begin{aligned}
\mathbf{I} &= [\boldsymbol{\tau}_1 \;\cdots\; \boldsymbol{\tau}_S]\,[Q_1\mathbf{b}_1^\top \;\cdots\; Q_S\mathbf{b}_S^\top]^\top && (11)\\
&= \underbrace{[\boldsymbol{\tau}_1 \;\cdots\; \boldsymbol{\tau}_S]}_{\text{transport matrix } \mathbf{T}} \underbrace{\begin{bmatrix} Q_1 & & 0 \\ & \ddots & \\ 0 & & Q_S \end{bmatrix}}_{\text{chromaticity matrix } \mathbf{Q}} \underbrace{\begin{bmatrix} \mathbf{b}_1 \\ \vdots \\ \mathbf{b}_S \end{bmatrix}}_{\text{BRF matrix } \mathbf{B}} && (12)\\
&= \mathbf{TQB}\,. && (13)
\end{aligned}$$

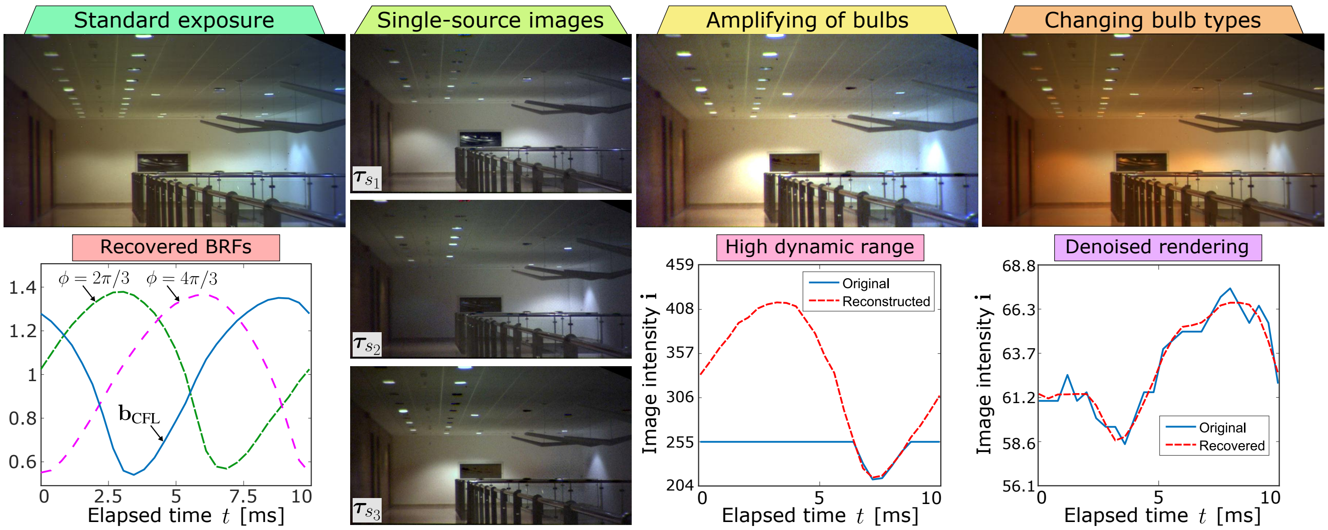

Figure 4. Unmixing results. The second column shows the single-source images reconstructed by unmixing. We simulate bulb amplification/de-amplification by computing $\boldsymbol{\tau}_1 + 3.5\boldsymbol{\tau}_3$, and bulb replacement by replacing $\mathbf{b}_1$ with a Sodium BRF. The left plot shows the monochrome BRFs sampled directly by ACam. On the right we illustrate the successful reconstruction of saturated or noisy pixels.

Matrix $\mathbf{T}$ is the scene's $P \times S$ *transport matrix*. Each column of $\mathbf{T}$ describes the appearance of the scene when a specific source is turned on. This matrix is time-invariant and generally unknown.

Finally, suppose the sources in the scene have phases $\phi_1, \ldots, \phi_S$ instead of being zero. The BRF matrix in Eq. (13) now contains BRFs that have been circularly shifted individually according to their sources' phase:

$$\mathbf{B} \;=\; [\ \mathrm{shift}(\phi_1, \mathbf{b}_1)^\top \ \ \cdots \ \ \mathrm{shift}(\phi_S, \mathbf{b}_S)^\top\ ]^\top \ . \quad (14)$$

5.1. Unmixing: Source Separation

Single-source sequences are linearly mixed in the data $\mathbf{I}$. We seek unmixing, *i.e.*, linear source separation [2]. The key is to estimate the transport matrix $\mathbf{T}$ based on Eq. (13).

Consider any two sources s_1 and s_2 that are connected to the same phase and have the same BRF. According to Eq. (9), the two-source sequence due to these sources is

$$\mathbf{I}_{s_1} + \mathbf{I}_{s_2} \;=\; (\boldsymbol{\tau}_{s_1} Q_{s_1} + \boldsymbol{\tau}_{s_2} Q_{s_2})\,\mathrm{shift}(\phi_{s_1}, \mathbf{b}_{s_1}) \ . \quad (15)$$

Thus the contributions of the two sources add up as if the scene is illuminated by a single source having the same phase and BRF. The contribution of these sources is therefore unseparable. Divide all sources used in the scene into subsets of sources, where each subset has no linear dependency to another. We consider unmixing only across these linearly-independent subsets. For the rest of the paper we refer to these independent subsets as the S "sources.".

Assume we know $\mathbf{QB}$. This is measured in two ways:

(a) Sources are very often in the field of view. Thus their BRFs and chromaticities can be acquired directly by our ACam. This is also possible for pixels dominated by one source (*e.g.*, reflections from nearby surfaces).

(b) Sampling the signal as in (a) and then using DELIGHT and the recognition method of Section 4.

The transport matrix is estimated using

$$\hat{\mathbf{T}} \;=\; \underset{\mathbf{T} \geq 0}{\arg\min} \ \left\| \mathbf{W} \odot (\mathbf{I} - \mathbf{TQB}) \right\|_F^2 \ , \quad (16)$$

where $\odot$ denotes a Hadamard (element-wise) multiplication and $\|\|_F$ is the Frobenius norm. The $P \times K$ weight matrix $\mathbf{W}$ discards saturated data:

$$\mathbf{W}[p,k] = \begin{cases} 0 & \text{if any spectral band is saturated at } \mathbf{I}[p,k] \\ 1 & \text{otherwise .} \end{cases} \quad (17)$$

Eq. (16) is a simple least-squares estimator. Due to noise and minor differences between sources of the same class, the assumption of a known $\mathbf{QB}$ is not precisely met. To counter slight inconsistencies, a refinement allows $\mathbf{B}$ to change a bit. Using $\hat{\mathbf{T}}$ derived in Eq. (16), we compute:

$$\hat{\mathbf{B}} \;=\; \underset{\mathbf{B} \geq 0}{\arg\min} \ \left\| \mathbf{W} \odot (\mathbf{I} - \hat{\mathbf{T}}\mathbf{QB}) \right\|_F^2 \ . \quad (18)$$

After this least-squares estimation of $\hat{\mathbf{B}}$, the estimation in Eq. (16) is applied again using $\hat{\mathbf{B}}$. We have observed in our experiments that, unless this refinement is done, the result may suffer from minor artifacts (see example in [34]).

Each column of $\hat{\mathbf{T}}$ is an unmixed image of the scene. This image is already white balanced because the chromaticities of all sources are factored into $\mathbf{Q}$. Examples are shown in Figures 4 and 5.

5.2. High Dynamic Range, Denoised Rendering

We can now reconstruct the single-source image sequence of a source s using

$$\hat{\mathbf{I}}_s \;=\; \hat{\boldsymbol{\tau}}_s\, Q_s\, \mathrm{shift}(\phi_s, \mathbf{b}_s) \quad (19)$$

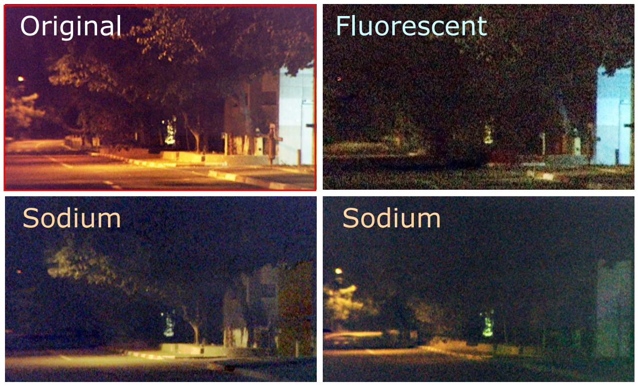

Figure 5. Unmixing results for an outdoor scene.

where $\hat{\tau}_s$ is the corresponding column of $\hat{\mathbf{T}}$. The intensities in this sequence can safely exceed the saturation level of the sensor. This is because Eqs. (16) and (17) bypass saturated data when estimating $\hat{\mathbf{T}}$. We therefore obtain *high dynamic range* results thanks to the AC (Figure 4[middle plot]).

The unmixing process also leads to denoising. Intensities in the captured image sequence suffer from sensor readout noise. Yet, since Eq. (19) forces all pixels to vary in synchrony according to a common BRF, the rendered sequence $\hat{\mathbf{I}}_s$ is less noisy than the input data (Figure 4([right plot])).

Last but not least, light sources can be changed to bulbs that were not seen at all during the acquisition. Changing bulbs means changing their chromaticity and BRF to that of other bulbs (*e.g.*, in DELIGHT or merely hallucinated). Moreover, we can change the grid phase of light sources and can use a diagonal amplification matrix $\mathbf{A}$ to amplify or de-amplify them. This leads to *generalized relighting*:

$$\mathbf{I}^{\text{relight}} = \hat{\mathbf{T}} \, [\mathbf{AQB}]^{\text{relight}} \; . \tag{20}$$

5.3. Semi-Reflection Separation

To separate a semi-reflection from a transmitted scene [1, 20, 37, 38], we show a new principle: passive AC-based unmixing. We realize this principle using either one of the following two mechanisms:

- **AC-illuminated scene:** When all light sources originate from AC-powered bulbs, unmixing is done as described in Section 5.1. See [34] for example results.
- **Natural illumination involved:** Scene illumination contains an outdoor component from daylight.

The indoor environment is illuminated by *two* kinds of sources. First, part of the natural daylight illuminates the indoors through a window. The second light source indoors is connected to the AC grid. In this case $\boldsymbol{\tau}_{\text{out}}$ and $\boldsymbol{\tau}_{\text{ac}}$ correspond to the two sources. Since daylight is approximately time-invariant at timescales of a few thousand cycles, its BRF is a vector of all ones. The bulb's BRF $\mathbf{b}_{\text{ac}}$ is unknown, *i.e.*, we are not relying on any database or known grid phase.

Figure 6. Temporal variations due to AC flicker are attributed to the indoor scene. ICA separates the indoor reflection from the transmitted scene. Both are recovered up to an unknown scale.

As before, our input data is an image sequence $\mathbf{I}$. We ignore chromaticities for brevity. Now consider two frames k_1 and k_2 with $\mathbf{b}_{\text{ac}}[k_1] > \mathbf{b}_{\text{ac}}[k_2]$. The corresponding images, represented by columns of $\mathbf{I}$, are:

$$\mathbf{I}[k_1] = \boldsymbol{\tau}_{\text{out}} + \boldsymbol{\tau}_{\text{ac}} \mathbf{b}_{\text{ac}}[k_1] \tag{21}$$

$$\mathbf{I}[k_2] = \boldsymbol{\tau}_{\text{out}} + \boldsymbol{\tau}_{\text{ac}} \mathbf{b}_{\text{ac}}[k_2] \; . \tag{22}$$

It follows that vectors $\boldsymbol{\tau}_{\text{ac}}$ and $\mathbf{I}[k_1] - \mathbf{I}[k_2]$ are equal up to a scale factor. Along similar lines, it is possible to show that vectors $\boldsymbol{\tau}_{\text{out}}$ and $\mathbf{I}[k_2] - A\mathbf{I}[k_1]$ are also equal up to a scale factor for some unknown scalar A. We estimate this scalar using independent component analysis (ICA) [14]. Specifically, A is optimized to minimize the mutual information of vectors $\mathbf{I}[k_1] - \mathbf{I}[k_2]$ and $\mathbf{I}[k_2] - A\mathbf{I}[k_1]$. This yields the result shown in Figure 6.

6. The Alternating-Current Camera (ACam)

The previous section relies on a key image acquisition task: capturing a sequence of K frames that spans one cycle. Very little light, however, enters the camera at the timescale of $1/K$-th the AC cycle.[4] This is especially problematic at night and indoors where light levels are usually low and sensor readout noise overwhelms the signal. Moreover, frame acquisition must support HDR imaging. This is because the field of view may include both bright light sources and poorly-lit surfaces (*e.g.*, from shadows, squared-distance light fall-off, AC flicker, *etc.*) These issues make capturing K-frame sequences impractical with a high-speed camera.

To overcome them, our ACam keeps its electronic shutter open for hundreds of cycles while optically blocking its sensor at all times *except* during the same brief interval in each cycle. This is illustrated in Figure 7[top]. Since the

[4] For instance, acquiring $K = 20$ images per cycle in North America, where light flickers at 120Hz, requires a frame exposure time of 416μsec.

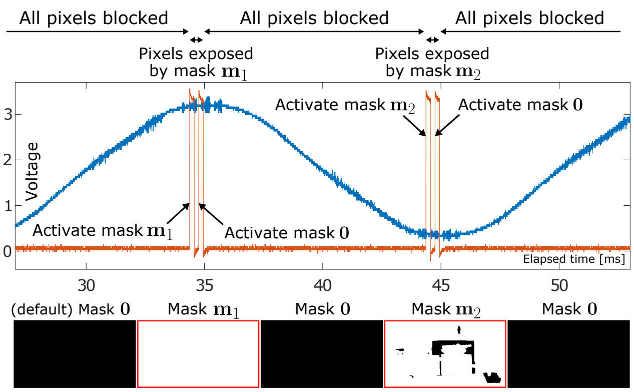

Figure 7. ACam operation. **Top:** The camera's pixels are repeatedly blocked and unblocked over C cycles so that they can integrate light only during *the same* brief interval in each cycle. Because each cycle's duration varies slightly, the timing of these events is controlled precisely with an Arduino that tracks AC zero-crossings in real time. Here we show the Arduino's input voltage (blue) and the mask-switching signal it generates (red), measured simultaneously with a high-speed oscilloscope. Masks are switched at the signal's rising edge and must persist for at least Δ/K microseconds. The ACam supports $K \leq 26$ for 50Hz grids and $K \leq 22$ for 60Hz grids. **Bottom:** The corresponding DMD masks. Mask $\mathbf{0}$ is active most of the time and acts like a global shutter. Mask $\mathbf{m}_1$ briefly exposes all pixels to light. Mask $\mathbf{m}_2$, on the other hand, blocks light from some of the pixels in the next cycle to prevent their saturation.

light collected by the sensor is proportional to the number of cycles the electronic shutter is open, the ACam trades off acquisition speed for enhanced signal-to-noise ratio. Moreover, it can handle large variations in light level across the field of view by allowing some sensor pixels to integrate light for fewer cycles than others (Figure 7[bottom]).

Just like other coded-exposure techniques [12, 31, 42], we implement high-speed pixel masking with a digital micromirror device (DMD) that is optically coupled to an off-the-shelf camera. We adopt the overall design proposed in [26], modifying it for the purpose of passive AC-modulated imaging. Figure 8 shows our ACam and highlights its main differences from the system in [26]. It operates correctly on 60Hz/120V and 50Hz/220V grids.

Each ACam image yields exactly one frame of the K-frame sequence, indexed by $k \in [1 \dots K]$. The procedure is applied K times to acquire all frames —and is potentially applied more times if HDR frames are needed.

Acquiring frame k without HDR To capture a frame we (1) define a sequence of M binary DMD masks, (2) open the electronic shutter for C cycles while the DMD is locked to the AC, and (3) close the shutter and read out the image. In practice, C ranges from 100 to 1500 cycles depending on light levels. During this period the DMD repeatedly goes through its M masks. ACam imaging is therefore controlled

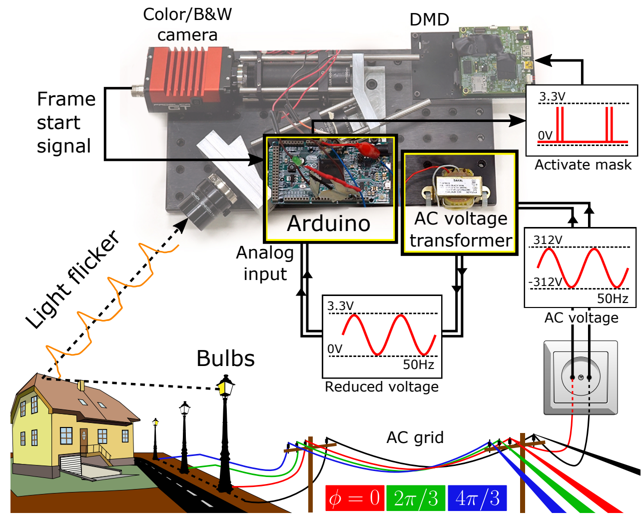

Figure 8. Our ACam combines an Arduino and voltage transformer with the DMD-based programmable mask in [26]. The Arduino senses the AC grid in real-time, switching between DMD masks over hundreds of cycles. The masks are loaded to the DMD by a PC (not shown). In each AC cycle, these masks expose individual pixels for a small fraction of the cycle duration.

by three quantities: the number of cycles C, the matrix $\mathbf{M}$ holding the mask sequence, and the timing signal that forces the DMD to switch from one mask to the next.

Our mask matrix has the following general form:

$$\mathbf{M} = [\,\mathbf{m}_1 \;\; \mathbf{0} \;\; \mathbf{m}_2 \;\; \mathbf{0} \;\; \dots \;\; \mathbf{m}_{M/2} \;\; \mathbf{0}\,] \tag{23}$$

where $\mathbf{m}_m$ is a column vector representing a binary pixel mask and $\mathbf{0}$ is a mask of all zeros. The zero mask blocks the sensor completely and is active at all times except during the interval corresponding to frame k. The non-zero mask, on the other hand, determines which pixels are actually exposed to light during that interval. To acquire a non-HDR image we set $\mathbf{m}_m = \mathbf{1}$ for all m. This forces the DMD to act like a "flutter-shutter" [29] synchronized with the AC. To acquire an HDR image we modify $\mathbf{M}$ adaptively over repeated long-exposure acquisitions (see below).

AC-locked mask switching We generate the mask-switching signal with an Arduino plugged into the reference outlet (Figure 7[top]). We found it very important to generate this signal in a closed loop, locked to the last-detected zero-crossing. Given that the duration of each cycle varies slightly, switching masks without accounting for this variation causes their position within a cycle to drift over time and leads to poor results (Figure 9). In contrast, locking the signal onto the zero-crossings gives temporal-blur-free images even after thousands of cycles.

Acquiring frame k with HDR We first acquire the frame without HDR, using a long enough exposure time to achieve good signal-to-noise ratio at dimly-lit surfaces. If this frame

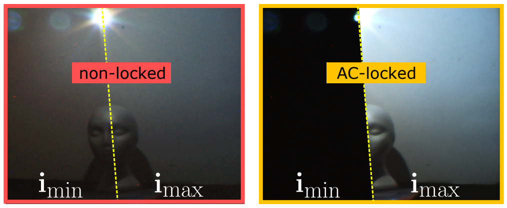

Figure 9. Non-locked versus AC-locked imaging. Using a 1500-cycle integration to acquire frames corresponding to the maximum and minimum intensity of a bulb (LED2). **Left:** Without AC locking the integration time window drifts, causing temporally-blurred results. **Right:** When the ACam is continuously synchronized to the AC zero-crossings, temporal blur is minimal.

has saturated pixels, we repeat the acquisition with a modified mask matrix that exposes saturated pixels to light for a lot less. Specifically, let p be a saturated pixel and let $\mathbf{M}[p,:]$ be the row corresponding to p. We modify $\mathbf{M}[p,:]$ by zeroing out half its non-zero elements. This cuts in half the time that pixel p will be exposed to light. In contrast, the rows of $\mathbf{M}$ associated with unsaturated pixels are left as-is. The process of modifying $\mathbf{M}$ and re-acquiring the frame is repeated until either the number of saturated pixels falls below a threshold or $\mathbf{M}$ has rows with only one non-zero element.[5] In this way, the brightest points in a scene can be exposed up to $M/2$ times less than the darkest ones.

7. Discussion

We believe we have only scratched the surface of imaging on the electric grid. Our unmixing of scene appearance to components associated with distinct bulb sets opens the door to further photometric processing. In particular, *photometric stereo* [45] can possibly be obtained in the wild using as few as $S = 4$ sources (bulbs in three AC phases and daylight). Because objects can be large relative to their distance to light sources, near-lighting effects will need to be accounted for [17, 27]. Unmixing can also be followed by intrinsic image recovery [44], shape from shadows [10], surface texture and BRDF chracterization [6]. Moreover, flicker using different bulbs and AC phases can be intentionally used for controlled illumination of objects. This way, multiplexed illumination [5, 30] is easily implemented.

We call for more sophisticated algorithms that are robust to deviations from assumptions. Deviations include situations where some scene bulb types or the number of sources S are unknown, as in Figure 10. Robustness is required for operation in the presence of non-AC temporal distractions: moving cars, blinking-advertisement neon lights and building-sized dynamic screens. Such non-stationary distractions are often unavoidable, because low light conditions demand acquisition times of seconds to minutes.

[5] Our ACam's DMD can handle up to $M = 96$ masks so the maximum number of iterations is $\lfloor \log_2(M/2) \rfloor = 6$.

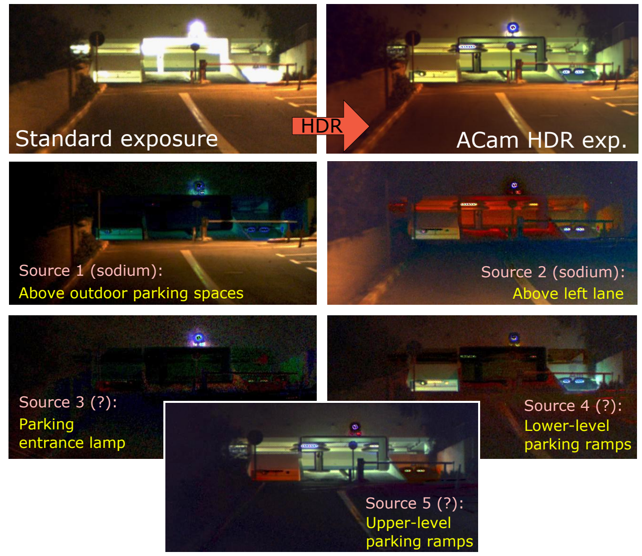

Figure 10. An unmixing experiment for a scene that deviates from our assumptions. Here, some scene bulbs are not in DELIGHT and are not observed directly due to their location deep inside the building. Lacking knowledge of the number of independent sources, BRFs and chromaticities, we set $S = 5$ for unmixing but in reality S is likely higher. The results suffer from residual crosstalk and color distortion, *e.g.*, notice that some signal from sources 4 and 5 falsely appears in parts of sources 1 and 2.

A system that yields denser image sequences, enlarging K, is desirable. This goal may be achieved by more elaborate coded imaging and processing, *e.g.*, allowing different pixels in each 5×5 neighborhood to sample a different interval of the AC cycle. Alternatively, compressed-sensing codes [23] can be used. Enhanced temporal resolution and signal-to-noise ratio expand the applications as well. These can involve distinguishing bulbs of the same type *and* AC phase, followed by their unmixing, as well as finer characterization of the electric grid.

Acknowledgements: The authors thank M. O'Toole, Y. Levron, R. Swanson, P. Lehn, Z. Tate, A. Levin and G. Tennenholtz for useful discussions, V. Holodovsky and R. Zamir for help with experiments, K. Lavia and I. Talmon for technical support. Y. Y. Schechner is a Landau Fellow—supported by the Taub Foundation. This research was supported by the Israel Science Foundation (Grant 542/16) and conducted in the Ollendorff Minerva Center. Minerva is funded through the BMBF. M. Sheinin was partly supported by the Mitacs Canada-Israel Globalink Innovation Initiative. K. N. Kutulakos gratefully acknowledges the support of the Natural Sciences and Engineering Research Council of Canada under the RGPIN and SGP programs, as well as the support of DARPA under the REVEAL program.

References

[1] A. Agrawal, R. Raskar, S. K. Nayar, and Y. Li. Removing photography artifacts using gradient projection and flash-exposure sampling. In *ACM SIGGRAPH*, 24(3):828–835, 2005.

[2] M. Alterman, Y. Y. Schechner and A. Weiss. Multiplexed fluorescence unmixing. In *Proc. IEEE ICCP*, 2010.

[3] N. Andersson, M. Sandström, A. Berglund, and K. Hansson. Amplitude modulation of light from various sources. *Lighting Research and Technology*, 26(3):157–160, Sept. 1994.

[4] B. G. Batchelor. Illumination sources. In *Machine Vision Handbook*, 283–317. Springer London, 2012.

[5] O. G. Cula, K. J. Dana, D. K. Pai and D. Wang. Polarization multiplexing for bidirectional imaging. In *Proc. IEEE CVPR*, 2:1116–1123, 2005.

[6] K. J. Dana. BRDF/BTF measurement device. In *Proc. IEEE ICCV*, 2:460–466, 2001.

[7] P. Debevec, T. Hawkins, C. Tchou, H.-P. Duiker, W. Sarokin, and M. Sagar. Acquiring the reflectance field of a human face. In *ACM SIGGRAPH*, 145–156, 2000.

[8] A. Ghosh, G. Fyffe, J. Busch, B. Tunwattanapong, X. Yu, and P. Debevec. Multiview face capture using polarized spherical gradient illumination. In *ACM SIGGRAPH ASIA*, 2011.

[9] J. J. Grainger, and W. D. Stevenson. Power system analysis. *McGraw-Hill*, 1994.

[10] M. Hatzitheodorou. Shape from shadows: a Hilbert space setting. In Journal of Complexity, 14.1:63–84, 1998.

[11] F. Heide, M. B. Hullin, J. Gregson, and W. Heidrich. Low-budget transient imaging using photonic mixer devices. In *ACM SIGGRAPH*, 2013.

[12] Y. Hitomi, J. Gu, M. Gupta, T. Mitsunaga, and S. K. Nayar. Video from a single coded exposure photograph using a learned over-complete dictionary. In *Proc. IEEE ICCV*, 287–294, 2011.

[13] E. Hsu, T. Mertens, S. Paris, S. Avidan, and F. Durand. Light mixture estimation for spatially varying white balance. In *ACM SIGGRAPH*, 2008.

[14] A. Hyvärinen, and E. Oja. Independent component analysis: algorithms and applications. In *Neural Networks*,13(4):411-430, 2008.

[15] K. Jo, M. Gupta, and S. K. Nayar. DisCo: Display-camera communication using rolling shutter sensors. *ACM TOG*, 35(5), article 150, 2016.

[16] H. R. V. Joze and M. S. Drew. Exemplar-based color constancy and multiple illumination. *IEEE TPAMI*, 36(5):860–873, 2014.

[17] B. Kim and P. Burger. Depth and shape from shading using the photometric stereo method. *CVGIP: Image understanding*, 54(3):416427, 1991.

[18] A. Kolaman, R. Hagege, and H. Guterman. Light source separation from image sequences of oscillating lights. In *Proc. IEEE Elect. & Electronics Eng. in Israel*, 1–5, 2014.

[19] A. Kolaman, M. Lvov, R. Hagege, and H. Guterman. Amplitude modulated video camera-light separation in dynamic scenes. In *Proc. IEEE CVPR*, 3698–3706, 2016.

[20] A. Levin and Y. Weiss. User assisted separation of reflections from a single image using a sparsity prior. In *IEEE TPAMI*, 29(9):1647-1654, 2011.

[21] B. Li, W. Xiong, W. Hu, and B. Funt. Evaluating Combinational Illumination Estimation Methods on Real-World Images. *IEEE TIP*, 23(3):1194–1209, Mar. 2014.

[22] M. G. Masi, L. Peretto, R. Tinarelli, and L. Rovati. Modeling of the physiological behavior of human vision system under flicker condition. In *IEEE Int. Conf. Harmonics & Quality of Power*, 2008.

[23] M. Mordechay and Y. Y. Schechner. Matrix Optimization for poisson compressed sensing. In *Proc. IEEE GlobalSIP*, 684–688, 2014.

[24] D. Moreno, K. Son, and G. Taubin. Embedded phase shifting: Robust phase shifting with embedded signals. In *Proc. IEEE CVPR*, 2301–2309, 2015.

[25] M. O'Toole, S. Achar, S. G. Narasimhan, and K. N. Kutulakos. Homogeneous codes for energy-efficient illumination and imaging. *ACM SIGGRAPH*, 2015.

[26] M. O'Toole, J. Mather, and K. N. Kutulakos. 3D shape and indirect appearance by structured light transport. *IEEE TPAMI*, 38(7):1298–1312, 2016.

[27] T. Papadhimitri and P. Favaro. Uncalibrated near-light photometric stereo. In *Proc. BMVA*, 2014.

[28] G. Petschnigg, R. Szeliski, M. Agrawala, M. F. Cohen, H. Hoppe, and K. Toyama. Digital photography with flash and no-flash image pairs. In *ACM SIGGRAPH*, 2004.

[29] R. Raskar, A. Agrawal, J. Tumblin. Coded exposure photography: motion deblurring using a fluttered shutter. In *ACM SIGGRAPH*, 2006.

[30] N. Ratner, Y. Y. Schechner, and F. Goldberg. Optimal multiplexed sensing: bounds, conditions and a graph theory link. In *Optics Express*, 15(25):17072–17092, 2007.

[31] D. Reddy, A. Veeraraghavan, and R. Chellappa. P2C2: Programmable pixel compressive camera for high speed imaging. In *Proc. IEEE CVPR*, 329–336, 2011.

[32] J. Salvi, J. Pages, and J. Batlle. Pattern codification strategies in structured light systems. *Pattern Recogn.*, 37(4):827–849, 2004.

[33] P. Sen, B. Chen, G. Garg, S. Marschner, M. Horowitz, M. Levoy, and H. P. A. Lensch. Dual photography. In *ACM SIGGRAPH*, 745–755, 2005.

[34] M. Sheinin, Y. Y. Schechner, and K. Kutulakos. Computational imaging on the electric grid: supplementary material. In *Proc. IEEE CVPR*, 2017.

[35] *Computational Imaging on the Electric Grid: Webpage*. [Online] (2017). Available at: http://webee.technion.ac.il/˜yoav/research/ACam.html

[36] *Computational Imaging on the Electric Grid: Project webpage*. [Online] (2017). Available at: http://www.dgp.toronto.edu/ACam

[37] S. Shwartz, Y. Y. Schechner, and M. Zibulevsky. Efficient separation of convolutive image mixtures. In *Proc. Int. Conf. on ICA and BSS*, 246–253, 2006.

[38] S. N. Sinha, J. Kopf, M. Goesele, D. Scharstein, and R. Szeliski. Image-based rendering for scenes with reflections. In *ACM TOG*, 31(4):1–10, 2006.

[39] H. Su, A. Hajj-Ahmad, R. Garg, and M. Wu. Exploiting rolling shutter for ENF signal extraction from video. In *Proc. IEEE ICIP*, 5367–5371, 2014.

[40] T. Tajbakhsh and R.-R. Grigat. Illumination flicker correction and frequency classification methods. In *Proc. SPIE Electronic Imaging*, 650210, 2007.

[41] A. Torralba and W. T. Freeman. Accidental pinhole and pinspeck cameras. *IJCV*, 110(2):92–112, 2014.

[42] A. Veeraraghavan, D. Reddy, and R. Raskar. Coded strobing photography: compressive sensing of high speed periodic videos. *IEEE TPAMI*, 33(4):671–686, 2011.

[43] M. Vollmer and K.-P. Moellmann. Flickering lamps. *European Journal of Physics*, 36(3), 2015.

[44] Y. Weiss. Deriving intrinsic images from image sequences. In *Proc. IEEE ICCV*, 2:68–75, 2001.

[45] R. J. Woodham. Photometric method for determining surface orientation from multiple images. *Opt. Eng.*, 19(1), 1980.

[46] H.-Y. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand, and W. T. Freeman. Eulerian video magnification for revealing subtle changes in the world. *ACM SIGGRAPH*, 31(4), 2012.

[47] W. Xiaoming, Z. Ke, Z. Ying, and S. Wenxiang. Study on fluorescent lamp illumination and flicker. In *Proc. IEEE Int. Conf. on Power Electronics and Drive Systems*, 1529–1532, 2003.

[48] Y. Yoo, J. Im, and J. Paik. Flicker removal for CMOS wide dynamic range imaging based on alternating current component analysis. *IEEE Trans. on Consumer Electronics*, 60(3):294–301, 2014.

[49] D. Zongker, D. Werner, B. Curless, and D. Salesin. Environment matting and compositing. In *ACM SIGGRAPH*, 1999.