

Amplitude Modulated Video Camera - Light Separation in Dynamic Scenes

Amir Kolaman, Maxim Lvov, Rami Hagege, and Hugo Guterman
Electrical and Computer Engineering Department Ben-Gurion University of the Negev
POB 653 Beer-Sheva 8410501, Israel

kolaman@post.bgu.ac.il, lvovmax@post.bgu.ac.il, hagege@ee.bgu.ac.il, hugo@ee.bgu.ac.il

Abstract

Controlled light conditions improve considerably the performance of most computer vision algorithms. Dynamic light conditions create varying spatial changes in color and intensity across the scene. These conditions, caused by a moving shadow for example, force developers to create algorithms which are robust to such variations. We suggest a computational camera which produces images that are not influenced by environmental variations in light conditions. The key insight is that many years ago, similar difficulties were already solved in radio communication; As a result each channel is immune to interference from other radio channels. Amplitude Modulated (AM) video camera separates the influence of a modulated light from other unknown light sources in the scene; Causing the AM video camera frame to appear the same - independent of the light conditions in which it was taken. We built a prototype of the AM video camera by using off the shelf hardware and tested it. AM video camera was used to demonstrate color constancy, shadow removal and contrast enhancement in real time. We show theoretically and empirically that: 1. the proposed system can produce images with similar noise levels as a standard camera. 2. The images created by such camera are almost completely immune to temporal, spatial and spectral changes in the background light.

1. Introduction

Many image and video analysis algorithms demonstrate their performance in a dark room with no background lights [1] [2]. Added dynamic background lights, which vary in space, time, and spectra, force these algorithms to compensate for the presence of additional lights. A camera, which creates the same image (Fig.1) anywhere, could solve this problem and promote the common use of many state of the art algorithms in consumer products independent of dynamic light conditions.

An example for such dynamic light condition is a casted shadow upon an object (Fig.2(a)). This generates video

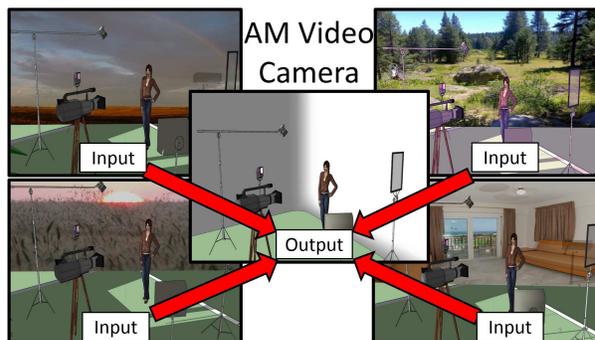


Figure 1. AM video camera enable working in stable illumination conditions independent of the background illumination

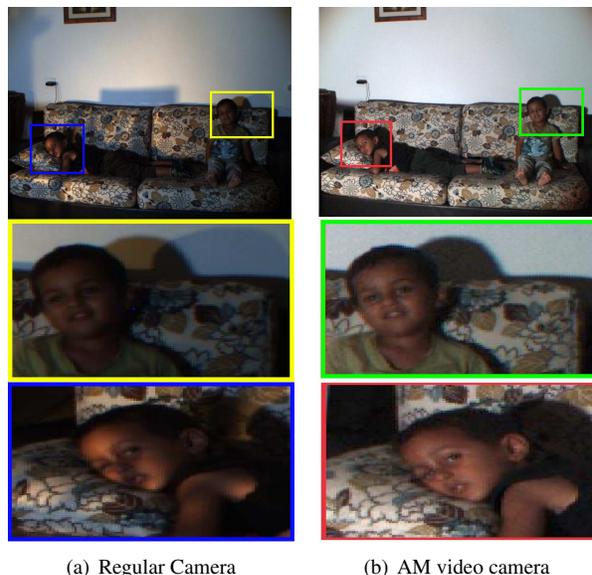


Figure 2. (a) Image with shadows, low contrast, and color cast. (b) Proposed method with no shadows, correct color (white wall), and improved contrast.

frames with changing intensity and color across space and time. Removing these casted shadows is done by separating the effect a singlet light source has on the scene. This light

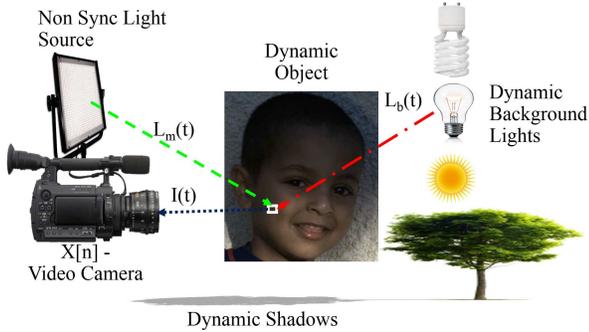


Figure 3. Pixel measurement is a superposition of the incident light of the modulated light source and the background light.

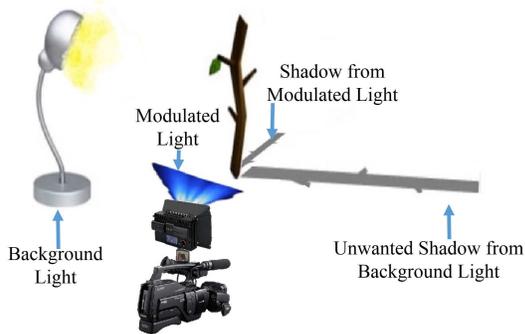


Figure 4. Complex light conditions create multiple shadows.

source separation clears casted shadows, corrects color, and enhances contrast (Fig.2(b)). Thus creating the same image under any light condition is done by light source separation.

Light source separation on video can be performed by one of the two approaches - passive solution or active solution. Passive solutions compensate for the effect of the unwanted background lights by assuming a pre-known behavior of light upon objects in the scene. Spatially varying light conditions, for example, raise the challenge bar on color constancy algorithms [3][4][5][6]. Active solutions use a controlled light source in order to clear the effect of the background lights in the scene. Flash no flash [7][8][9], for example, is a common active solution, which generates results from two captured frames: first with flash on, second with flash off. It assumes a static background [10] and precise synchronization between flash and camera. Video color correction [11] [12] and shadow removal [13] are usually treated as separate problems in the literature. Therefore shadow removal and color correction combined could enhance the performance of many video analysis methods such as object tracking [14] and face detection [15].

1.1. Contributions:

Our active solution proposes the following contributions:

- **Light source separation:** capture a frame that records the contribution of a single light source, and clear

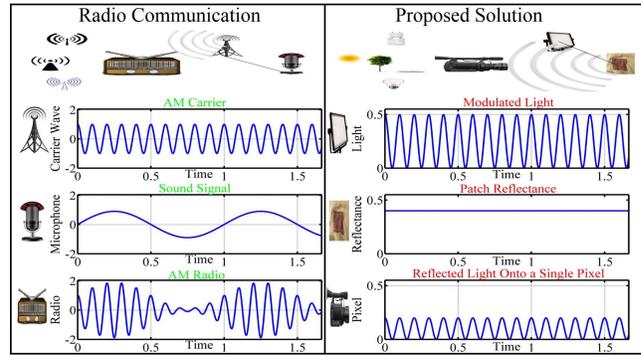


Figure 5. In radio communication (Left Side) the AM Carrier is used to modulate the Sound signal into a Amplitude Modulated (AM) Radio. Proposed method (Right Side) uses a Modulated Light to illuminate a Patch Reflectance in the scene and create a Reflected Light Onto a Single Pixel.

out the contributions of other lights, which change in space, time, or spectra (subsection 2.1).

- **Video shadow removal, contrast enhancement, and color correction:** perform all of the above using a low complexity algorithm in real time video(subsection 2.3).
- **Preserving the appearance of an object independent of the light conditions:** how to maximize reconstruction accuracy (by minimizing reconstruction error) and minimize noise (subsection 2.2).

1.2. Problem formulation

Fig.3 shows a typical video scene with dynamic lights, shadows, and objects. The background light conditions may change unexpectedly in time and space creating non-uniform color and intensity on the moving object. A non-synchronized light is situated on top of the video camera. We would like to know the impact of the non-synchronized light on the scene in order to ease color correction and remove shadows created by the other lights (Fig. 4).

2. Modulated Light Source Separation

Our suggested solution is inspired by Amplitude Modulation (AM) radio (left column of Fig.5), where the voice - modulated by a radio wave - is received by the tuned radio receiver, which filters out all the other frequencies. In photography the scene is modulated by the light source and received by the camera. By tuning the camera to a specific frequency, it should be possible to filter out the background lights from a modulated light oscillating at that frequency (right column of Fig.5). This system has the same limitations as AM radio - where every station must transmit at a separate frequency - the modulated light must oscillate at a unique frequency, different from the background lights.

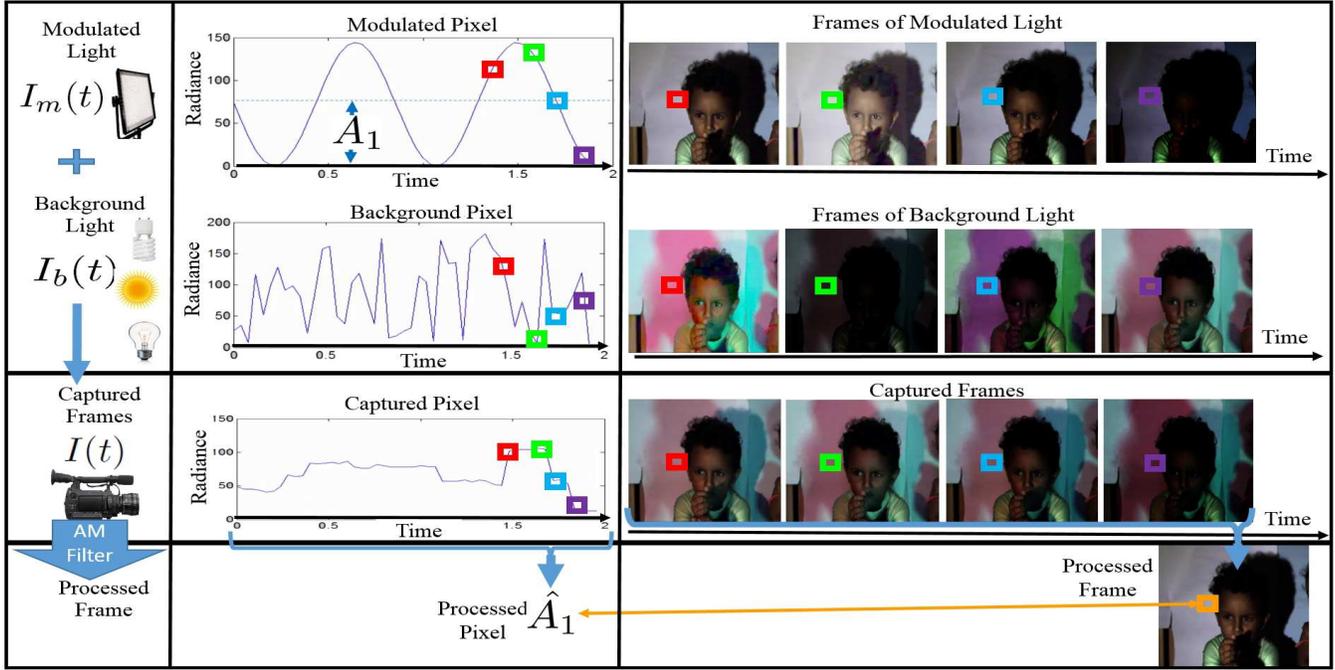


Figure 6. Data of a patch/pixel across time. A patch affected by the modulated light $I_m(t)$ (first line) having an amplitude A_1 , is also affected by background lights $I_b(t)$ (second line) and captured by a video camera $I(t)$ (third line) in time. The captured values in time are processed into a single reconstructed amplitude \hat{A}_1 (last line).

2.1. Mathematical Formulation

Modulated light source converts the objects in the scene into modulated signals in time. Consider a light path (figure 3) that begins at the light sources, reflects from a patch, and is measured by a camera pixel. Light sources divide into two groups: 1. Background sources $I_b(t)$ with an unknown behavior in time and space 2. Ideal modulated light (non-synchronized) source $I_m(f_1, t)$, which is modeled by the following:

$$I_m(f_1, t) = a_0 + a_1 \cos(2\pi f_1 t) \quad (1)$$

where t represents time, a_0 is the constant intensity over time, a_1 is the amplitude of the main harmonic oscillating at $f_1 = \frac{1}{T_1}$.

Total light in the scene is reflected by the object patch generates a radiance $I(t)$ equal to:

$$I(t) = C + A_1 \cdot \cos(2\pi f_1 t) + I_b(t) \quad (2)$$

where C depends on the patch reflectance and constant part of all the lights (modulated and background), radiance coefficient A_1 depends on the patch reflectance and intensity amplitude a_1 from Eq. (1), and $I_b(t)$ are the dynamic background lights.

Modulated radiance $A_1 \cdot \cos(2\pi f_1 t)$ has two important properties: 1. The frequency f_1 of $\cos(2\pi f_1 t)$ is the same as the frequency of the modulated light source. 2. A_1 is

constant in time. These properties help to separate the influence of the modulated light (A_1) from the influence of the dynamic background lights ($I_b(t)$) and constant part (C).

Radiance $I(t)$ is sampled by a camera pixel at discrete times $n \in \{0, 1, \dots, N-1\}$:

$$X[n] = C + I_b[nT_s] + A_1 \cdot \cos[2\pi f_1 nT_s + \varphi_1] \quad (3)$$

where C is the measured radiance¹ of constant part (modulated and background), $I_b[nT_s]$ is the intensity of the dynamic background radiance, $T_s = \frac{1}{f_s}$ is the sample time of the camera (the sample frequency f_s is also referred as Frames Per Second (FPS)), A_1 is the amplitude of the modulated radiance, $\cos[2\pi f_1 nT_s + \varphi_1]$ is a discrete sample of $\cos(2\pi f_1 t)$, and φ_1 is the unknown phase difference between modulated light and camera. Note that this sample model is ideal without noise artifacts, which will be discussed in subsection 2.2

The aim of the AM video camera system is to reconstruct A_1 using pre-known information on the frequency of the modulated light f_1 . This can be done by various methods, one of which is the inner product using a Finite Impulse Response (FIR) filter:

$$\hat{A}_1 = \left| \frac{2}{N} \sum_{n=0}^{N-1} X[n] e^{-i2\pi f_1 T_s n} \right| \quad (4)$$

¹Camera can measure radiance by normalizing its measurements with Exposure Value (EV)

This filter will attenuate all the terms in Eq. (3) except for the amplitude of the oscillating part at the target frequency f_1 i.e. A_1 . The purpose of the absolute value is get rid of the phase term $e^{i\varphi_1}$ which is unknown.

In summary AM video camera creates a single processed frame (last row of Fig.6) using Eq. (4) from N captured frames (third row of Fig. 6), which are illuminated by both modulated light (first row of Fig. 6) and a random background light (second row of Fig.6). The processed frame \hat{A}_1 (right side of the last row of Fig. (6)) is a reconstruction of the amplitude of the AM light source A_1 up to a scale (second image from the left of the first row of Fig. (6)).

2.2. Performance Analysis

Performance of this reconstruction system is measured by two factors: 1.Reconstruction Error 2.Noise levels. The real modulated light source has unwanted harmonics - which change Eq. (1) into

$$L_m(f_1, t) = a_0 + a_1 \cos(2\pi f_1 t) + \sum_{k=2}^{\infty} a_k \cos(2\pi f_k t) \quad (5)$$

where $\{a_k\}_{k=2}^M$ are amplitudes of the parasitic harmonics, and $\{f_k\}_{k=2}^M$ are their frequencies.

This changes light radiance Eq. (2) into:

$$I(t) = C + \sum_{k=1}^M A_k \cos(2\pi f_k t + \varphi_k), \quad (6)$$

where the term C is a constant term of the illumination, $\{A_k\}_{k=1}^M$ are the amplitudes of different harmonics (including modulated light source and background illuminations).

A camera captures N frames of the scene at a frame rate f_s . Denote by $\{t_n\}_{n=0}^{N-1}$ the acquisition time of frame number $n \in \{0, 1, \dots, N-1\}$. The time between consequent frame acquisitions is not constant due to noise, and can be described by:

$$t_n = t_{n-1} + (1/f_s)(1 + q_n), \quad (7)$$

where $\{q_n\}_{n=0}^{N-1}$ is a zero mean white Gaussian noise with variance σ_q^2 .

Non ideal radiance Eq. (6) change the ideal sampled signal, presented in Eq. (3), into

$$X[n] = C + \sum_{k=1}^M A_k \cos(\omega_k T_s(n + r_n) + \varphi_k) + Z_n, \quad (8)$$

where A_1 is the radiance amplitude of the modulated light from Eq. (3), $\{A_k\}_{k=2}^M$ are the radiance amplitudes of background lights and parasitic harmonics of the modulated light, with their frequencies $f_k = \frac{\omega_k}{2\pi}$, $\{r_n\}_{n=0}^{N-1}$ is a

Gaussian random walk process defined by $r_n = \sum_{m=0}^n q_m$, $\{\varphi_k\}_{k=2}^M$ are the random phases of the additional harmonics distributed uniformly on the interval $[0, 2\pi]$ and independent of $\{r_n\}_{n=0}^{N-1}$ ², and Z_n is a zero mean additive noise.

Reconstruction error is important in many applications such as spectral measurements and radiance evaluations, which gather information from a digital camera or a photometric sensor. This creates the need to evaluate the reconstruction error of our method in order to justify its use in precise measurement tools. The reconstruction error can be measured by:

$$MSE = \mathbb{E} \left[\left| \hat{A}_1 - A_1 \right|^2 \right] \quad (9)$$

where MSE is the Mean Square Error between the reconstructed signal \hat{A}_1 , and the amplitude intensity A_1 .

Precise derivation of the MSE is difficult due to the non linearity of the reconstruction formula (caused by the absolute value operation). A simple bound, however, on the MSE can be derived:

$$\mathbb{E} \left[\left| \hat{A}_1 - A_1 \right|^2 \right] \leq MSE_C + MSE_{A_1} + MSE_{A_k} + \frac{2}{N} \mathbb{E} [Z_n^2] \quad (10)$$

where MSE_C is due to the constant term C , MSE_{A_1} is due to the modulated light harmonic $\{A_1, f_1\}$, MSE_{A_k} due to all the other harmonics $\{A_k, f_k\}_{k=2}^M$, and $\mathbb{E}[Z_n^2]$ is the variance of the additive noise in Eq. (8).

- $MSE_C = C^2 \left| \frac{2}{N} \cdot \frac{\sin(\pi N f_1 / f_s)}{\sin(\pi f_1 / f_s)} \right|^2$.
- $MSE_{A_k} = \sum_{k=2}^M A_k^2 (I_k^+ + I_k^-)$ where $I_k^\pm = \frac{1}{N^2} \sum_{n,m=0}^{N-1} e^{i2\pi(n-m)(f_k \pm f_1)T_s - 2|n-m|(\sigma_q \pi f_k T_s)}$. It can be shown that I_k^\pm decays as $O(1/N)$ if $\sigma_q > 0$. If $\sigma_q = 0$ then I_k^\pm simplifies to: $I_k^\pm = \left| \frac{1}{N} \cdot \frac{\sin(\pi N(f_1 \pm f_k)/f_s)}{\sin(\pi(f_1 \pm f_k)/f_s)} \right|^2$.
- $MSE_{A_1} = A_1^2 (I_1 + I_1^+)$ where $I_1 = I_1^- + 1 - \frac{2}{N} \sum_{n=0}^{N-1} e^{-2n(\pi \sigma_q f_1 T_s)^2}$ is due to the unwanted phase noise. $I_1 \rightarrow 0$ as $(\sigma_q f_1 T_s)^2 N \rightarrow 0$ and $I_1 \rightarrow 1$ as $(\sigma_q f_1 T_s)^2 N \rightarrow \infty$.

Reconstruction error Eq. (10) explains several visible phenomena of its graph (Fig. 10): 1. Local minimums are generated by the variance of $MSE_C \rightarrow 0$ when $N \cdot f_1 / f_s \in \mathbb{N}$ and f_1 / f_s is far from an integer. 2. MSE diminishes as N gets larger up to a limit.

²For each pair of phases φ_j, φ_k the expected value $\mathbb{E} \left[e^{i(\varphi_j - \varphi_k)} \right]$ is zero.

The model can be generalized to contain additional random processes (such as white/colorful noise), not necessary periodic. If, for instance, a wide-sense-stationary noise $\{b[n]\}$ with a power spectral density $S_b(\theta)$ is added to $X[n]$ then its contribution to the MSE is

$$MSE_b = \frac{1}{2\pi} \int_{[-\pi, \pi]} S_b(\theta) K_N(2\pi f_1 T_s - \theta) d\theta, \quad (11)$$

where $K_N(\theta) = \left| \frac{1}{N} \cdot \frac{\sin(\theta N/2)}{\sin(\theta/2)} \right|^2$. If the noise $\{b[n]\}$ does not contain high spectral power near the frequency $\theta = 2\pi f_1 T_s$ then its contribution to the MSE will be small.

Noise Level is one of the important factors for measuring the quality of an output color image [16].

Amplitude of the modulated intensity A_1 should be bigger than the camera noise levels in order to have no apparent noise artifacts [17]; meaning the radiance of the AM light should be in the same order of magnitude as the background light. In addition the noise level is inverse-proportional to the number of captured frames N . Apparent noise levels depend on intensity relation between AM light and background lights.

Z_n is a zero mean additive noise. Its standard deviation (STD) depends on many factors, such as the temperature, exposure time and the average light intensity during the frame acquisition. Since only the light intensity changes from one frame to another, we can define a function $g(\mu)$ to be the conditional STD of Z_n given that $X[n] - Z[n]$ is equal to μ . Examples of the functions $g(\mu)$ for red, green and blue pixels are shown in Fig. 11. The conditional mean of Z_n given that $(X[n] - Z[n])$ is equal to μ , is zero. The noise terms $\{Z_n\}_{n=0}^{N-1}$ are statistically independent of each other.

2.3. Example Applications

Background lights generate shadows, which are captured by the camera. A light source situated on-top of the camera creates shadows that do not appear in the frame because this light path is almost aligned with the optical axis of the camera (Fig.4). Thus AM video camera removes shadows by generating an image influenced only by the modulated light (middle images of Fig. 8(b) and 8(d)).

Background light removal causes the output frames of AM video camera to have a single light source. This single light source helps color constancy algorithms to perform better since most of them assume a dominant light in the scene [6]. Am video frame fixate the light source in the scene - reducing the need for sophisticated color constancy/white balance algorithms - and may be replaced by a fixed color correction matrix. Color correction results are better in AM video frames compared to the standard camera (top two images in Fig. 8(d) and bottom image at Fig. 8(d)).

Uneven light across the scene may generate low contrast at some parts of the image even though the entire dynamic

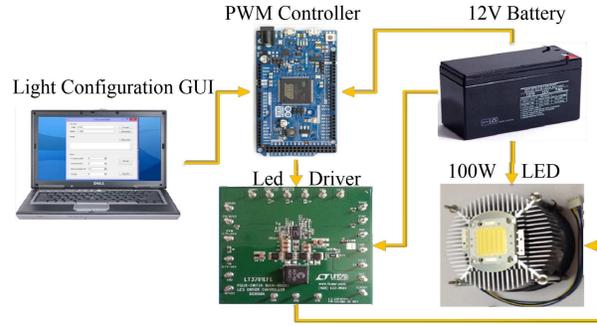


Figure 7. The prototype - System diagram of the modulated light source.

range of the camera is used (two button images of Fig. 8(c)). Uniform light conditions, generated by the removed background lights and shadows, improve local image contrast (two button images of Fig. 8(d)).

3. Live Modulated Light Source Separation

3.1. System Overview of the Prototype

An online video system was built using off the shelf products, and includes three parts - usb3 camera, laptop, and a modulated light source. The laptop controls the AM light, captures the frame from the camera, and performs the post-processing. The laptop sets the required frequency and amplitude of the AM light (Fig.7) by configuring a PWM sine generator, which is an input to a driver board of the 100W LED light. The system is capable of generating sine waves from 1 Hz to 600 Hz with up to 256 points per cycle and varying amplitude from 10% to 100% of the 100W LED light source.

3.2. Constraints

AM light frequency should be unique enough to make sure that the reconstructed frame contains only the AM light source and no background lights. The required modulated light f_1 is set according to the available frequencies in the captured scene. The system finds the required frequency f_1 by capturing a set of N frames - prior to turning on the AM light - and finding the minimal power on the $FFT\{X[n]\}$.

Camera exposure was set to 1 ms in order to be able to capture up to 1000 FPS (our camera could effectively reach 700 FPS). This fast exposure time forces the lens aperture to open at its maximum value - in order to get enough light into the camera. Frames used in all of the experiments were captured by a camera with 400 frames per second and a resolution of 640x480.

4. Experimental Results

This section presents experimental results showing applications and analysis of AM video camera output.



(a) Original image

(b) Proposed method

(c) Original image

(d) Proposed method

Figure 8. (a) and (c) Original scene (b) and (d) AM video camera frames performing real time video contrast enhancement, color correction, and shadow removal

4.1. Applications

The application part shows color correction, shadow removal, and contrast enhancement under different types of background illumination conditions and objects. Dynamic background light varied between natural (sunlight) and artificial (tungsten), while scene type varied between static and dynamic - shadows and objects. The analysis part assessed noise and reconstruction error on the output.

AM video camera output shows several example applications in Fig.8, where captured frames are presented in Fig. 8(a) and Fig. 8(c) and processed frames are presented in Fig. 8(b) and Fig. 8(d). Contrast enhancement occurs when the input image has high intensity differences between the modulated light and the background light as in the second row of Fig. 8(a) and fourth and third rows of Fig. 8(c). Color correction occurs when the background light and the modulated light have large color differences, as in the last row of Fig. 8(a) and first and second row of Fig. 8(c). Shadow caused by the background lights are removed in all of the resulting AM frames.³

³Real time videos available in the paper site

4.2. Performance Measurements

This part compares noise and reconstruction error of the AM video camera, presented in 2.2, with the standard camera. A static scene was captured at a frame rate of $f_s = 704Hz$ with different constant light illuminations and modulated light with a frequency of $f_1 = 105Hz$. Fig.9 presents the values of a single pixel as a function of time, and demonstrates how its values are only due to the modulated illumination. The blue curve shows the pixel value captured by the camera when no background lights are present. The green and the red curves show the processed pixel's values \hat{A}_1 . $N = 20$ frames were used to calculate the green curve \hat{A}_1 , and $N = 67$ frames was used to calculate the red curve. For these values of N the value of Nf_1/f_s is close to an integer and the value of MSE_C in Eq. (10) is very small. Note $N = 67$ and $N = 20$ were chosen according to the local minima of the relative root mean squared error (RRMSE) defined by $\frac{\sqrt{E[\hat{A}_1 - A_1]^2}}{A_1}$ (Fig. 10).

Performance of the reconstruction system was evaluated by the following features:

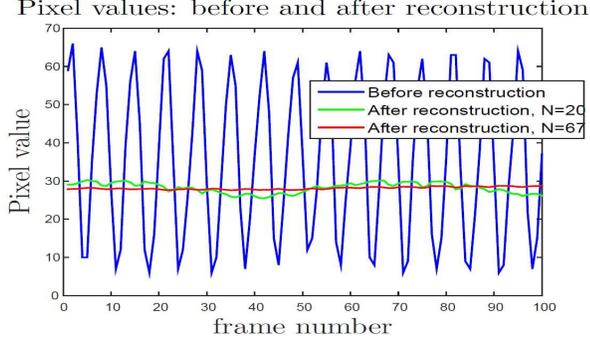


Figure 9. Pixel values for different times. In blue: the captured by the camera pixel values, without background light. In green and red: the processed pixel values \hat{A}_1 for $N = 20$ and $N = 67$, respectively.

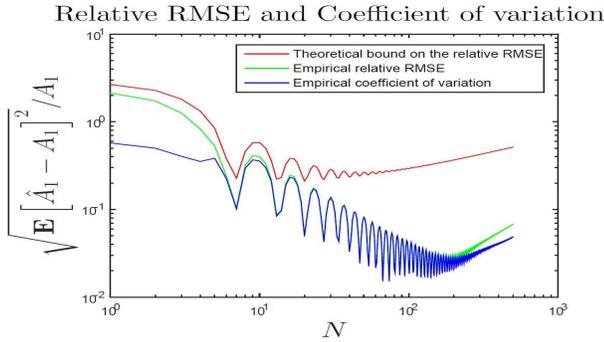


Figure 10. The green curve: the empirical RRMSE of \hat{A}_1 . The red curve: a theoretical upper bound for the RRMSE of \hat{A}_1 . The blue curve: the empirical coefficient of variation of \hat{A}_1 .

- Standard deviation of a reconstructed pixel \hat{A}_1 when the scene is static.
- Coefficient of variation defined by $\frac{STD(\hat{A}_1)}{\mu_{\hat{A}_1}}$, where $\mu_{\hat{A}_1}$ is the expected value of \hat{A}_1 .
- RRMSE

Fig. 10 shows the RRMSE as a function of N (number of frames for calculation of \hat{A}_1). The green curve represents the empirical RRMSE, The red curve represents a theoretical upper bound - based on Eq. (10) - and the blue curve represents the empirical coefficient of variation. To calculate the theoretical upper bound for the RRMSE the following parameters were estimated from the blue curve in Fig. 9: $\frac{C}{A_1} = 1.23$, $\frac{A_2}{A_1} = 0.13$, $\frac{A_3}{A_2} = 0.03$, where A_2, A_3 are the amplitudes of additional harmonics of our modulated illumination. The frequencies of these harmonics are $f_k = k \cdot f_1$, for $k = 2, 3$. The value of $(\sigma_q f_1 T_s)^2$ is taken to be $1.6 \cdot 10^{-3}$. The function $g(\mu)$ used to estimate the last term of Eq. (10) is the $g(\mu)$ for blue pixels shown in Fig. 11

Since σ_q is not zero, \hat{A}_1 tends to zero and the RRMSE tends to 1 as N tends to infinity. The reason is because

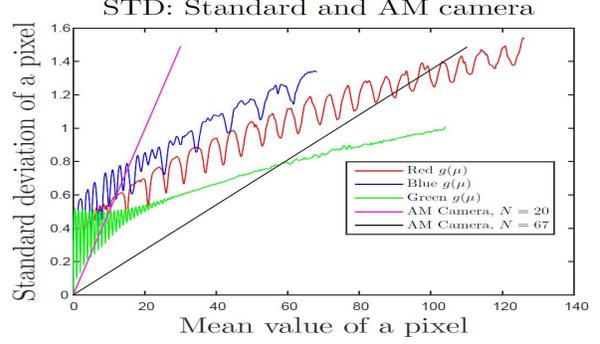


Figure 11. In red/green/blue: the STD of a red/green/blue pixel in a standard camera as a function of its mean value. In black and magenta: the STD of both red, green and blue pixels in AM video camera (the STD of \hat{A}_1) as a function of its mean value, when no background illumination is present.

in the inner product $\sum_{n=0}^{N-1} X[n]e^{-i2\pi f_1 T_s n}$ the component of $X[n]$ that should be proportional to $e^{i2\pi f_1 T_s n}$ contains a phase noise. That noise makes this component to be in the same phase as $e^{i2\pi f_1 T_s n}$ for some times and in the opposite phase for other times. The sum of these terms would cancel, leading the whole sum to grow slower than N . The multiplication of the whole sum by $2/N$ makes it tend to zero as N tends to infinity. If, on the other hand, $\sigma_q = 0$ then \hat{A}_1 would tend to A_1 and the RRMSE would tend to zero, since the upper bound in Eq. (10) would tend to zero.

Fig. 11 shows the STD of the pixels as a function of their mean value for standard and AM video camera. The red, green and blue curves represent the function $g(\mu)$ (for red, green and blue pixels), which is the STD of a pixel in a standard camera given that its mean value is μ . The STD of \hat{A}_1 are shown in the magenta curve (for $N = 20$) and the black curve (for $N = 67$) for the scenario when there is no background illumination. These graphs are the same for the red, green and blue colors, and grow linearly as a function of μ . This can be explained by Eq. (10), since all terms in that bound, except for the last term that is much smaller than the others, are proportional to A_1^2 .

5. Discussion

Our AM video camera system demonstrates a low complexity and effective system performing shadow removal, color correction, and contrast enhancement on real time video frames (Fig. 8). Performance analysis demonstrated how precision of the AM video camera grows as number of sampled frames N is higher up to a limit - which was proven analytically and experimentally (Fig. 10). AM video camera is unaffected by the intensity of the background assuming the modulated amplitude A_1 is bigger than the camera noise values. In practice, AM light intensities surpassing 20 percent of background light level gave appreciable results. For cases of very high intensities of background light our

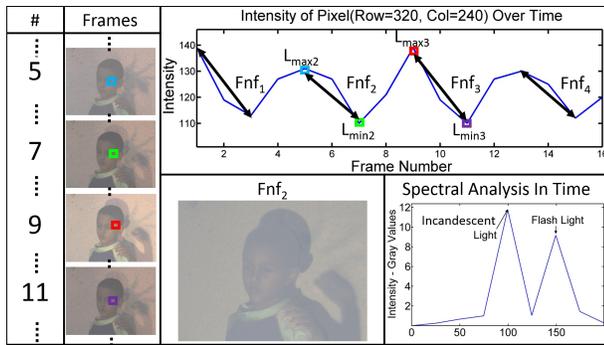


Figure 12. A sampled pixel in time showing flash points L_{max} and no-flash points L_{min} creating an image with low contrast F_{nf} (middle bottom part) using Flash No-Flash method. A spectral analysis of the sampled pixel is shown at bottom right.

Performance Comparison Under Changing Background Light

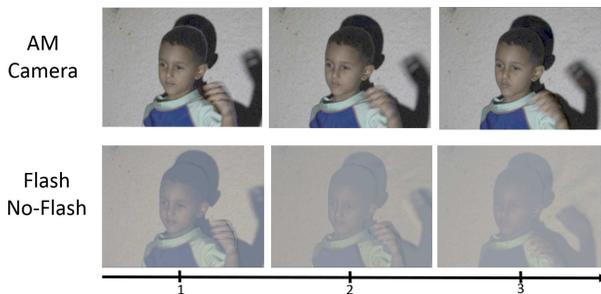


Figure 13. Inconsistency of Flash No-Flash when the background light changes compared with our method

system could work in sunlight using method such as [18], with some minor adjustments to work with a rolling shutter camera. Reconstruction analysis proved that there is an analytical and experimental upper bound to the MSE (Fig. 10). Noise levels of AM video camera resemble the noise level of the standard camera (Fig. 11) and get closer to its performance as N grows.

This paper differs from Schechner et. al [19] by assuming a dynamic background light. Our work is closely related to [20], but in our case non-synchronized light was used without any spatial patterns. It also improves [21] by capturing dynamic video scenes outside the lab, using less frames to generate a single reconstruction, and no need for synchronization between the light source and the camera.

Flash no flash assume constant illumination between two subsequent frames. Therefore when such techniques apply to indoor video with changing background lights, such as incandescent/florecent light, they may produce a video sequence with

1. Low contrast output (Fig. 12) - due to negative change in the background lights between two subsequent frames.
2. Inconsistent flickering video (Fig. 13) - due to incon-

sistent changes in the background lights between two subsequent frames.

The AM camera technique actively eliminate the influence of illumination changes. Therefore the video produced by the AM camera is much more consistent than the one produced by flash no flash techniques.

6. Conclusions and Future work

The highlights of the AM video camera are:

- We used a principle from the AM radio field and applied it to computational photography. AM demodulation filtered out all of the background lights and reconstructed the scene illuminated only by the AM light source.
- Light source separation can be used as an application for shadow removal, color correction, and contrast enhancement. The shadows in the reconstructed image are minimized by placing the modulated light source near the camera. Color correction is easier to perform because the reconstructed AM frame contains a single light source. Contrast is enhanced due to uniform light conditions in the AM frame.
- A highly parallelizable algorithm, which has good potential to work on currently available smart phones, was presented. The algorithm works separately on each pixel and its results are non-dependent on neighboring pixels.
- The presented method requires no synchronization between light and video camera, thus needing less hardware, and is simpler to implement.
- Prototype was built and tested extensively under different light conditions showing real time video color correction, shadow removal, and contrast enhancement.

The suggested technique can be implemented directly on available hardware using software alone. Future work will include methods for integration of the processing stage into the sensor, improved control of the AM light source, and mitigating the edge artifacts⁴.

7. Acknowledgment

The authors would like to thank undergraduate students Boris Bessarabov and Eitam Man for their help in building the modulated light system used in this paper.

⁴Real time videos available in the paper site

References

- [1] M. O. et. al. Homogeneous codes for energy-efficient illumination and imaging. **1**
- [2] C. Tsotsios, M. E. Angelopoulou, T.-K. Kim, and A. J. Davison, “Backscatter compensated photometric stereo with 3 sources,” in *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*. IEEE, 2014, pp. 2259–2266. **1**
- [3] H. R. V. Joze and M. S. Drew, “Exemplar-based color constancy and multiple illumination,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 36, no. 5, pp. 860–873, 2014. **2**
- [4] E. Hsu, T. Mertens, S. Paris, S. Avidan, and F. Durand, “Light mixture estimation for spatially varying white balance,” in *ACM Transactions on Graphics (TOG)*, vol. 27, no. 3. ACM, 2008, p. 70. **2**
- [5] B. Li, W. Xiong, W. Hu, and B. Funt, “Evaluating combinational illumination estimation methods on real-world images,” *Image Processing, IEEE Transactions on*, vol. 23, no. 3, pp. 1194–1209, 2014. **2**
- [6] A. Gijsenij, T. Gevers, and J. Van De Weijer, “Computational color constancy: Survey and experiments,” *Image Processing, IEEE Transactions on*, vol. 20, no. 9, pp. 2475–2489, 2011. **2, 5**
- [7] S. He and R. W. Lau, “Saliency detection with flash and no-flash image pairs,” in *Computer Vision—ECCV 2014*. Springer, 2014, pp. 110–124. **2**
- [8] O. Suominen and A. Gotchev, “Preserving natural scene lighting by strobe-lit video,” in *IS&T/SPIE Electronic Imaging*. International Society for Optics and Photonics, 2015, pp. 939 919–939 919. **2**
- [9] C. Wu, R. Samadani, and P. Gunawardane, “Same frame rate ir to enhance visible video conference lighting,” in *Image Processing (ICIP), 2011 18th IEEE International Conference on*. IEEE, 2011, pp. 1521–1524. **2**
- [10] J. Sun, J. Sun, S. B. Kang, Z.-B. Xu, X. Tang, and H.-Y. Shum, “Flash cut: Foreground extraction with flash and no-flash image pairs,” in *Computer Vision and Pattern Recognition, 2007. CVPR’07. IEEE Conference on*. IEEE, 2007, pp. 1–8. **2**
- [11] V. Prinet, D. Lischinski, and M. Werman, “Illuminant chromaticity from image sequences,” in *Computer Vision (ICCV), 2013 IEEE International Conference on*. IEEE, 2013, pp. 3320–3327. **2**
- [12] A. Chakrabarti, K. Hirakawa, and T. Zickler, “Color constancy with spatio-spectral statistics,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 34, no. 8, pp. 1509–1519, 2012. **2**
- [13] S. H. Khan, M. Bennamoun, F. Sohel, and R. Togneri, “Automatic feature learning for robust shadow detection,” in *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*. IEEE, 2014, pp. 1939–1946. **2**
- [14] A. Prati, I. Mikic, M. M. Trivedi, and R. Cucchiara, “Detecting moving shadows: algorithms and evaluation,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 25, no. 7, pp. 918–923, 2003. **2**
- [15] S. Hauberg, A. Feragen, and M. J. Black, “Grassmann averages for scalable robust pca,” in *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*. IEEE, 2014, pp. 3810–3817. **2**
- [16] A. Kolaman and O. Yadid-Pecht, “Quaternion structural similarity: a new quality index for color images,” *Image Processing, IEEE Transactions on*, vol. 21, no. 4, pp. 1526–1536, 2012. **5**
- [17] S. W. Hasinoff, F. Durand, and W. T. Freeman, “Noise-optimal capture for high dynamic range photography,” in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*. IEEE, 2010, pp. 553–560. **5**
- [18] M. O’Toole, S. Achar, S. G. Narasimhan, and K. N. Kutulakos, “Homogeneous codes for energy-efficient illumination and imaging,” *ACM Transactions on Graphics (TOG)*, vol. 34, no. 4, p. 35, 2015. **8**
- [19] Y. Y. Schechner, S. K. Nayar, and P. N. Belhumeur, “A theory of multiplexed illumination,” in *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*. IEEE, 2003, pp. 808–815. **8**
- [20] J. Gu, T. Kobayashi, M. Gupta, and S. K. Nayar, “Multiplexed illumination for scene recovery in the presence of global illumination,” in *Computer Vision (ICCV), 2011 IEEE International Conference on*. IEEE, 2011, pp. 691–698. **8**
- [21] A. Kolaman, R. Hagege, and H. Guterman, “Light source separation from image sequences of oscillating lights,” in *Electrical & Electronics Engineers in Israel (IEEEI), 2014 IEEE 28th Convention of*. IEEE, 2014, pp. 1–5. **8**