Femto-photography: capturing and visualizing the propagation of light

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.

<table>
<thead>
<tr>
<th>Citation</th>
<th>Andreas Velten, Di Wu, Adrian Jarabo, Belen Masia, Christopher Barsi, Chinmaya Joshi, Everett Lawson, Mouni Bawendi, Diego Gutierrez, and Ramesh Raskar. 2013. Femto-photography: capturing and visualizing the propagation of light. ACM Trans. Graph. 32, 4, Article 44 (July 2013), 8 pages.</th>
</tr>
</thead>
<tbody>
<tr>
<td>As Published</td>
<td><a href="http://dx.doi.org/10.1145/2461912.2461928">http://dx.doi.org/10.1145/2461912.2461928</a></td>
</tr>
<tr>
<td>Publisher</td>
<td>Association for Computing Machinery (ACM)</td>
</tr>
<tr>
<td>Version</td>
<td>Author's final manuscript</td>
</tr>
<tr>
<td>Accessed</td>
<td>Thu Feb 07 00:30:18 EST 2019</td>
</tr>
<tr>
<td>Citable Link</td>
<td><a href="http://hdl.handle.net/1721.1/82039">http://hdl.handle.net/1721.1/82039</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>Creative Commons Attribution-Noncommercial-Share Alike 3.0</td>
</tr>
<tr>
<td>Detailed Terms</td>
<td><a href="http://creativecommons.org/licenses/by-nc-sa/3.0/">http://creativecommons.org/licenses/by-nc-sa/3.0/</a></td>
</tr>
</tbody>
</table>
Femto-Photography: Capturing and Visualizing the Propagation of Light

Andreas Velten\textsuperscript{1∗} \hspace{1cm} Di Wu\textsuperscript{1†} \hspace{1cm} Adrian Jarabo\textsuperscript{2} \hspace{1cm} Belen Masia\textsuperscript{1,2} \hspace{1cm} Christopher Barsi\textsuperscript{1}

Chinmaya Joshi\textsuperscript{1†} \hspace{1cm} Everett Lawson\textsuperscript{1} \hspace{1cm} Mounig Bawendi\textsuperscript{3} \hspace{1cm} Diego Gutierrez\textsuperscript{2} \hspace{1cm} Ramesh Raskar\textsuperscript{1}

1 MIT Media Lab 2 Universidad de Zaragoza 3 MIT Department of Chemistry

Abstract

We present \textit{femto-photography}, a novel imaging technique to capture and visualize the propagation of light. With an effective exposure time of 1.85 picoseconds (ps) per frame, we reconstruct movies of ultrafast events at an equivalent resolution of about one half trillion frames per second. Because cameras with this shutter speed do not exist, we re-purpose modern imaging hardware to record an ensemble average of repeatable events that are synchronized to a streak sensor, in which the time of arrival of light from the scene is coded in one of the sensor’s spatial dimensions. We introduce reconstruction methods that allow us to visualize the propagation of femtosecond light pulses through macroscopic scenes; at such fast resolution, we must consider the notion of \textit{time-unwarping} between the camera’s and the world’s space-time coordinate systems to take into account effects associated with the finite speed of light. We apply our femto-photography technique to visualizations of very different scenes, which allow us to observe the rich dynamics of time-resolved light transport effects, including scattering, specular reflections, diffuse interreflections, diffraction, caustics, and subsurface scattering. Our work has potential applications in artistic, educational, and scientific visualizations; industrial imaging to analyze material properties; and medical imaging to reconstruct subsurface elements. In addition, our time-resolved technique may motivate new forms of computational photography.


Keywords: ultrafast imaging, computational photography

1 Introduction

Forward and inverse analysis of light transport plays an important role in diverse fields, such as computer graphics, computer vision, and scientific imaging. Because conventional imaging hardware is slow compared to the speed of light, traditional computer graphics and computer vision algorithms typically analyze transport using low time-resolution photos. Consequently, any information that is encoded in the time delays of light propagation is lost. Whereas the joint design of novel optical hardware and smart computation, i.e., computational photography, has expanded the way we capture, ana-

\textsuperscript{∗}Currently at Morgridge Institute for Research, University of Wisconsin at Madison.

\textsuperscript{†}Currently at Tsinghua University.

\textsuperscript{‡}Currently at College of Engineering, Pune.
Figure 2: Our setup for capturing a single 1D space-time photo. (a) A laser beam strikes a diffuser, which converts the beam into a spherical energy front that illuminates the scene; a beam splitter and a synchronization detector enable synchronization between the laser and the streak sensor. (b) After interacting with the scene, photons enter a horizontal slit in the camera and strike a photocathode, which generates electrons. These are deflected at different angles as they pass through a microchannel plate, by means of rapidly changing the voltage between the electrodes. (c) The CCD records the horizontal position of each pulse and maps its arrival time to the vertical axis, depending on how much the electrons have been deflected. (d) We focus the streak sensor on a single narrow scan line of the scene. (e) Sample image taken by the streak sensor. The horizontal axis (672 pixels) records the photons’ spatial locations in the acquired scanline, while the vertical axis (1 nanosecond window in our implementation) codes their arrival time. Rotating the adjustable mirrors shown in (a) allows for scanning of the scene in the y-axis and generation of ultrafast 2D movies such as the one visualized in Figure 1. (Figures (a)-(d), credit: [Gbur 2012])

Challenges Developing such time-resolved system is a challenging problem for several reasons that are under-appreciated in conventional methods: (a) brute-force time exposures under 2 ps yield an impractical signal-to-noise (SNR) ratio; (b) suitable cameras to record 2D image sequences at this time resolution do not exist due to sensor bandwidth limitations; (c) comprehensible visualization of the captured time-resolved data is non-trivial; and (d) direct measurements of events appear warped in space-time, because the finite speed of light implies that the recorded light propagation delay depends on camera position relative to the scene.

Contributions Our main contribution is in addressing these challenges and creating a first prototype as follows:

• We exploit the statistical similarity of periodic light transport events to record multiple, ultrashort exposure times of one-dimensional views (Section 3).

• We introduce a novel hardware implementation to sweep the exposures across a vertical field of view, to build 3D space-time data volumes (Section 4).

• We create techniques for comprehensible visualization, including movies showing the dynamics of real-world light transport phenomena (including reflections, scattering, diffuse inter-reflections, or beam diffraction) and the notion of peak-time, which partially overcomes the low-frequency appearance of integrated global light transport (Section 5).

• We introduce a time-unwarping technique to correct the distortions in captured time-resolved information due to the finite speed of light (Section 6).

Limitations Although not conceptual, our setup has several practical limitations, primarily due to the limited SNR of scattered light.

Since the hardware elements in our system were originally designed for different purposes, it is not optimized for efficiency and suffers from low optical throughput (e.g., the detector is optimized for 500 nm visible light, while the infrared laser wavelength we use is 795 nm), and from dynamic range limitations. This lengthens the total recording time to approximately one hour. Furthermore, the scanning mirror, rotating continuously, introduces some blurring in the data along the scanned (vertical) dimension. Future optimized systems can overcome these limitations.

2 Related Work

Ultrafast Devices The fastest 2D continuous, real-time monochromatic camera operates at hundreds of nanoseconds per frame [Goda et al. 2009] (about 6 · 10^9 frames per second), with a spatial resolution of 200 × 200 pixels, less than one third of what we achieve. Avalanche photodetector (APD) arrays can reach temporal resolutions of several tens of picoseconds if they are used in a photon starved regime where only a single photon hits a detector within a time window of tens of nanoseconds [Charbon 2007]. Repetitive illumination techniques used in incoherent LiDAR [Tou 1995; Gelbart et al. 2002] use cameras with typical exposure times on the order of hundreds of picoseconds [Busck and Heiselberg 2004; Colaço et al. 2012], two orders of magnitude slower than our system. Liquid nonlinear shutters actuated with powerful laser pulses have been used to capture single analog frames imaging light pulses at picosecond time resolution [Duguay and Mattick 2004; Colacó et al. 2012; Gelbart et al. 2002; Kodama et al. 1999; Qu et al. 2006]. They have also been used as line scanning devices for image transmission through highly scattering turbid media, by recording the ballistic photons, which travel a straight path through the scatterer and thus arrive first on the sensor [Hebden 1993].
principles that we develop in this paper for the purpose of transient imaging were first demonstrated by Velten et al. [2012c]. Recently, photonic mixer devices, along with nonlinear optimization, have also been used in this context [Heide et al. 2013].

Our system can record and reconstruct space-time world information of incoherent light propagation in free-space, table-top scenes, at a resolution of up to 672 × 1000 pixels and under 2 picoseconds per frame. The varied range and complexity of the scenes we capture allow us to visualize the dynamics of global illumination effects, such as scattering, specular reflections, interreflections, subsurface scattering, caustics, and diffraction.

Time-Resolved Imaging Recent advances in time-resolved imaging have been exploited to recover geometry and motion around corners [Raskar and Davis 2008; Kirmani et al. 2011; Velten et al. 2012b; Velten et al. 2012a; Gupta et al. 2012; Pandharkar et al. 2011] and albedo of from single view point [Naik et al. 2011]. But, none of them explored the idea of capturing videos of light in motion in direct view and have some fundamental limitations (such as capturing only third-bounce light) that make them unsuitable for the present purpose. Wu et al. [2012a] separate direct and global illumination components from time-resolved data captured with the system we describe in this paper, by analyzing the time profile of each pixel. In a recent publication [Wu et al. 2012b], the authors present an analysis on transient light transport in frequency space, and show how it can be applied to bare-sensor imaging.

3 Capturing Space-Time Planes

We capture time scales orders of magnitude faster than the exposure times of conventional cameras, in which photons reaching the sensor at different times are integrated into a single value, making it impossible to observe ultrafast optical phenomena. The system described in this paper has an effective exposure time down to 1.85 ps; since light travels at 0.3 mm/ps, light travels approximately 0.5 mm between frames in our reconstructed movies.

System: An ultrafast setup must overcome several difficulties in order to accurately measure a high-resolution (both in space and time) image. First, for an unamplified laser pulse, a single exposure time of less than 2 ps would not collect enough light, so the SNR would be unworkably low. As an example, for a table-top scene illuminated by a 100 W bulb, only about 1 photon on average would reach the sensor during a 2 ps open-shutter period. Second, because of the time scales involved, synchronization of the sensor and the illumination must be executed within picosecond precision. Third, standalone streak sensors sacrifice the vertical spatial dimension in order to code the time dimension, thus producing x-t images. As a consequence, their field of view is reduced to a single horizontal line of view of the scene.

We solve these problems with our ultrafast imaging system, outlined in Figure 2. (A photograph of the actual setup is shown in Figure 3 (left)). The light source is a femtosecond (fs) Kerr lens mode-locked Ti:Sapphire laser, which emits 50-fs with a center wavelength of 795 nm, at a repetition rate of 75 MHz and average power of 500 mW. In order to see ultrafast events in a scene with macro-scaled objects, we focus the light with a lens onto a Lambertian diffuser, which then acts as a point light source and illuminates the entire scene with a spherically-shaped pulse (see Figure 3 (right)). Alternatively, if we want to observe pulse propagation itself, rather than the interactions with large objects, we direct the laser beam across the field of view of the camera through a scattering medium (see the bottle scene in Figure 1).

Because all the pulses are statistically identical, we can record the scattered light from many of them and integrate the measurements to average out any noise. The result is a signal with a high SNR. To synchronize this illumination with the streak sensor (Hamamatsu C5680 [Hamamatsu 2012]), we split off a portion of the beam with a glass slide and direct it onto a fast photodetector connected to the sensor, so that, now, both detector and illumination operate synchronously (see Figure 2 (a)).

Capturing space-time planes: The streak sensor then captures an x-t image of a certain scanline (i.e. a line of pixels in the horizontal dimension) of the scene with a space-time resolution of 672 × 512. The exact time resolution depends on the amplification of an internal sweep voltage signal applied to the streak sensor. With our hardware, it can be adjusted from 0.30 ps to 5.07 ps. Practically, we choose the fastest resolution that still allows for capture of the entire duration of the event. In the streak sensor, a photomultiode converts incoming photons, arriving from each spatial location in the scanline, into electrons. The streak sensor generates the x-t image by deflecting these electrons, according to the time of their arrival, to different positions along the t-dimension of the sensor (see Figure 2(b) and 2(c)). This is achieved by means of rapidly changing the sweep voltage between the electrodes in the sensor. For each horizontal scanline, the camera records a scene illuminated by the pulse and averages the light scattered by 4.5 × 10^5 pulses (see Figure 2(d) and 2(e)).

Performance Validation To characterize the streak sensor, we compare sensor measurements with known geometry and verify the linearity, reproducibility, and calibration of the time measurements. To do this, we first capture a streak image of a scanline of a simple scene: a plane being illuminated by the laser after hitting the diffuser (see Figure 4 (left)). Then, by using a Faro digitizer arm [Faro 2012], we obtain the ground truth geometry of the points along that plane and of the point of the diffuser hit by the laser; this allows us to compute the total travel time per path (diffuser-plane-streak sensor) for each pixel in the scanline. We then compare the travel time captured by our streak sensor with the real travel time computed...
from the known geometry. The graph in Figure 4 (right) shows agreement between the measurement and calculation.

4 Capturing Space-Time Volumes

Although the synchronized, pulsed measurements overcome SNR issues, the streak sensor still provides only a one-dimensional movie. Extension to two dimensions requires unfeasible bandwidths: a typical dimension is roughly $10^{12}$ pixels, so a three-dimensional data cube has $10^9$ elements. Recording such a large quantity in a $10^{-9}$ second (1 ns) time window requires a bandwidth of $10^{15}$ bytes/s, far beyond typical available bandwidths.

We solve this acquisition problem by again utilizing the synchronized repeatability of the hardware: A mirror-scanning system (two 9 cm × 13 cm mirrors, see Figure 3 (left)) rotates the camera’s center of projection, so that it records horizontal slices of a scene sequentially. We use a computer-controlled, one-rpm servo motor to rotate one of the mirrors and consequently scan the field of view vertically. The scenes are about 25 cm wide and placed about 1 meter from the camera. With high gear ratios (up to 1:1000), the continuous rotation of the mirror is slow enough to allow the camera to record each line for about six seconds, requiring about one hour for 600 lines (our video resolution). We generally capture extra lines, above and below the scene (up to 1000 lines), and then crop them to match the aspect ratio of the physical scenes before the movie was reconstructed.

These resulting images are combined into one matrix, $M_{ijk}$, where $i = 1...672$ and $k = 1...512$ are the dimensions of the individual x-t streak images, and $j = 1...1000$ addresses the second spatial dimension $y$. For a given time instant $k$, the submatrix $N_{ij}$ contains a two-dimensional image of the scene with a resolution of $672 \times 1000$ pixels, exposed for as short to 1.85 ps. Combining the x-t slices of the scene for each scanline yields a 3D x-y-t data volume, as shown in Figure 5 (left). An x-y slice represents one frame of the final movie, as shown in Figure 5 (right).

5 Depicting Ultrafast Videos in 2D

We have explored several ways to visualize the information contained in the captured x-y-t data cube in an intuitive way. First, contiguous $N_{ij}$ slices can be played as the frames of a movie. Figure 1 (bottom row) shows a captured scene (bottle) along with several representative $N_{ij}$ frames. (Effects are described for various scenes in Section 7.) However, understanding all the phenomena shown in a video is not a trivial task, and movies composed of x-y frames such as the ones shown in Figure 10 may be hard to interpret. Merging a static photograph of the scene from approximately the same point of view with the $N_{ij}$ slices aids in the understanding of light transport in the scenes (see movies within the supplementary video). Although straightforward to implement, the high dynamic range of the streak data requires a nonlinear intensity transformation to extract subtle optical effects in the presence of high intensity reflections. We employ a logarithmic transformation to this end.

We have also explored single-image methods for intuitive visualization of full space-time propagation, such as the color-coding in Figure 1 (right), which we describe in the following paragraphs.

Integral Photo Fusion By integrating all the frames in novel ways, we can visualize and highlight different aspects of the light flow in one photo. Our photo fusion results are calculated as $N_{ij} = \sum w_k M_{ijk}, \{k = 1..512\}$, where $w_k$ is a weighting factor determined by the particular fusion method. We have tested several different methods, of which two were found to yield the most intuitive results: the first one is full fusion, where $w_k = 1$ for all $k$. Summing all frames of the movie provides something resembling a black and white photograph of the scene illuminated by the laser, while showing time-resolved light transport effects. An example is shown in Figure 6 (left) for the alien scene. (More information about the scene is given in Section 7.) A second technique, rainbow fusion, takes the fusion result and assigns a different RGB color to each frame, effectively color-coding the temporal dimension. An example is shown in Figure 6 (middle).

Peak Time Images The inherent integration in fusion methods, though often useful, can fail to reveal the most complex or subtle behavior of light. As an alternative, we propose peak time images, which illustrate the time evolution of the maximum intensity in each frame. For each spatial position $(i, j)$ in the x-y-t volume, we find the peak intensity along the time dimension, and keep information within two time units to each side of the peak. All other values in the streak image are set to zero, yielding a more sparse space-time volume. We then color-code time and sum up the x-y frames in this new sparse volume, in the same manner as in the rainbow fusion case but use only every 20th frame in the sum to create black lines between the equi-time paths, or isochrones. This results in a map of the propagation of maximum intensity contours, which we term peak time image. These color-coded isochronous lines can be thought of intuitively as propagating energy fronts. Figure 6 (right) shows the peak time image for the alien scene, and Figure 1 (top, middle) shows the captured data for the bottle scene depicted using this visualization method. As explained in the next section, this visualization of the bottle scene reveals significant light transport phenomena that could not be seen with the rainbow fusion visualization.
In order to visualize all light transport events as they have occurred (not as the camera captured them), we transform the captured data from camera time to world time, a transformation which we term time unwarping. Mathematically, for a scene point \( P = (i, j) \), we apply the following transformation:

\[
\begin{align*}
    t'_{ij} &= t_{ij} + \frac{z_{ij}}{c/\eta},
\end{align*}
\]

where \( t'_{ij} \) and \( t_{ij} \) represent camera and world times respectively, \( c \) is the speed of light in vacuum, \( \eta \) the index of refraction of the medium, and \( z_{ij} \) is the distance from point \( P \) to the camera. For our table-top scenes, we measure this distance with a Faro digitizer arm, although it could be obtained from the data and the known position of the diffuser, as the problem is analogous to that of bistatic LiDAR. We can thus define light travel time from each point \((i, j)\) in the scene to the camera as \( \Delta t_{ij} = t'_{ij} - t_{ij} = z_{ij}/(c/\eta) \). Then, time unwarping effectively corresponds to offsetting data in the x-y-t volume along the time dimension, according to the value of \( \Delta t_{ij} \) for each of the \((i, j)\) points, as shown in Figure 8.

In most of the scenes, we only have propagation of light through air, for which we take \( \eta \approx 1 \). For the bottle scene, we assume that the laser pulse travels along its longitudinal axis at the speed of light, and that only a single scattering event occurs in the liquid inside. We take \( \eta = 1.33 \) as the index of refraction of the liquid and ignore refraction at the bottle’s surface. A step-by-step unwarping process is shown in Figure 9 for a frame (i.e. x-y image) of the bottle scene. Our unoptimized Matlab code runs at about 0.1 seconds per frame. A time-unwarped peak-time visualization of the whole of this scene is shown in Figure 1 (right). Notice how now the caustics originate from the bottle and propagate outward, energy fronts along the label are correctly depicted as straight lines, and the pulse precedes related phenomena, as expected.

### 7 Captured Scenes

We have used our ultrafast photography setup to capture interesting light transport effects in different scenes. Figure 10 summarizes them, showing representative frames and peak time visualizations. The exposure time for our scenes is between 1.85 ps for the crystal scene, and 5.07 ps for the bottle and tank scenes, which required imaging a longer time span for better visualization. Please refer to the video in the supplementary material to watch the reconstructed movies. Overall, observing light in such slow motion reveals both subtle and key aspects of light transport. We provide here brief...
Figure 9: Time unwarping for the bottle scene, containing a scattering medium. From left to right: a frame of the video without correction, where the energy front appears curved; the same frame after time-unwarping with respect to distance to the camera $z_{ij}$; the shape of the energy front is now correct, but it still appears before the pulse; the same frame, time-unwarped taking also scattering into account.

Figure 10: More scenes captured with our setup (refer to Figure 1 for the bottle scene). For each scene, from left to right: photograph of the scene (taken with a DSLR camera), a series of representative frames of the reconstructed movie, and peak time visualization of the data. Please refer to the supplementary video for the full movies. Note that the viewpoint varies slightly between the DSLR and the streak sensor.
The system could be extended to image in color by adding additional pulsed laser sources at different colors or by using one continuously tunable optical parametric oscillator (OPO). A second color of about 400 nm could easily be added to the existing system by doubling the laser frequency with a nonlinear crystal (about $1000$). The streak tube is sensitive across the entire visible spectrum, with a peak sensitivity at about 450 nm (about five times the sensitivity at 800 nm). Scaling to bigger scenes would require less time resolution and could therefore simplify the imaging setup. Scaling should be possible without signal degradation, as long as the camera aperture and lens are scaled with the rest of the setup. If the aperture stays the same, the light intensity needs to be increased quadratically to obtain similar results.

Beyond the ability of the commercially available streak sensor, advances in optics, material science, and compressive sensing may bring further optimization of the system, which could yield increased resolution of the captured x-t streak images. Nonlinear shutters may provide an alternate path to femto-photography capture systems. However, nonlinear optical methods require exotic materials and strong light intensities that can damage the objects of interest (and must be provided by laser light). Further, they often suffer from physical instabilities.

We believe that mass production of streak sensors can lead to affordable systems. Also, future designs may overcome the current limitations of our prototype regarding optical efficiency. Future research can investigate other ultrafast phenomena such as propagation of light in anisotropic media and photonic crystals, or may be used in applications such as scientific visualization (to understand ultra-fast processes), medicine (to reconstruct subsurface elements), material engineering (to analyze material properties), or quality control (to detect faults in structures). This could provide radically new challenges in the realm of computer graphics. Graphics research can enable new insights via comprehensible simulations and new data structures to render light in motion. For instance, relativistic rendering techniques have been developed using our data, where the common assumption of constant irradiance over the surfaces does no longer hold [Jarabo et al. 2013]. It may also allow a better understanding of scattering, and may lead to new physically valid models, as well as spawn new art forms.

Acknowledgements

We would like to thank the reviewers for their insightful comments, and the entire Camera Culture group for their support. We also thank Greg Gbur for letting us use some of the images shown in Figure 2, Elisa Amoros for the 3D illustrations in Figures 3 and 4, and Paz Hernandez and Julio Marco for helping with the video. This work was funded by the Media Lab Consortium Members, MIT Lincoln Labs and the Army Research Office through the Institute for Soldier Nanotechnologies at MIT, the Spanish Ministry of Science and Innovation through the Mimesis project, and the EU-funded projects Golem and Verve. Di Wu was supported by the National Basic Research Project (No.2010CB731800) of China and the Key Project of NSFC (No. 61120106003 and 60932007). Belen Masia was additionally funded by an FPU grant from the Spanish Ministry of Education and by an NVIDIA Graduate Fellowship. Ramesh Raskar was supported by an Alfred P. Sloan Research Fellowship and a DARPA Young Faculty Award.
References


