Eurographics Symposium on Rendering 2008
Steve Marschner and Michael Wimmer
(Guest Editors)

Volume 27 (2008), Number 4

Combining Confocal Imaging and Descattering

Christian Fuchs[†] Michael Heinz[†] Marc Levoy[‡] Hans-Peter Seidel[†] Hendrik P. A. Lensch[†]

[†]MPI Informatik [‡] Stanford University

(a) (b) (c) (d) (e)

Figure 1: *Reconstructed volume of a 3D scene in a fish tank with diluted milk. While in the floodlit image (b) only the front objects are recognizable, confocal imaging (c) partially removes the haze. Using our novel descattering algorithm combined with confocal imaging more global scattering is removed and objects at even larger distances become visible (d). The image is sharper and features more saturated colors. In (e), the recovered 3D structure is visualized for a different view. Images (c)-(e) have been created by computing a confocal or descattered confocal volume followed by maximum intensity projection.*

Abstract

In translucent objects, light paths are affected by multiple scattering, which is polluting any observation. Confocal imaging reduces the influence of such global illumination effects by carefully focusing illumination and viewing rays from a large aperture to a specific location within the object volume. The selected light paths still contain some global scattering contributions, though. Descattering based on high frequency illumination serves the same purpose. It removes the global component from observed light paths. We demonstrate that confocal imaging and descattering are orthogonal and propose a novel descattering protocol that analyzes the light transport in a neighborhood of light transport paths. In combination with confocal imaging, our descattering method achieves optical sectioning in translucent media with higher contrast and better resolution.

1. Introduction

In this work we propose a technique for computing cross-sectional images of translucent objects or scenes for visible wavelengths by combining *confocal imaging* and algorithmic *descattering*.

Confocal imaging [CK96, WJNK96, NJW97] allows for computing optical sections of partially *transparent* volumes. Arbitrary slices through the volume are assembled by focusing the illumination and the observation rays from a large aperture onto individual voxels. Due to the large aperture, all contributions of points off the selected plane are significantly blurred and darkened, and their influence on the confocal image is drastically reduced. In our approach, we perform synthetic aperture confocal imaging [LCV*04], where the aperture is sampled by a small set of cameras and projectors. Individual voxels are extracted by combining the viewing and illumination rays that intersect at the voxel.

As can be seen in Figure 1(c), confocal imaging works to some extent even in translucent media where the light transport is governed by multiple scattering. Confocal imaging can be seen as extracting a particular region, a hypersurface, from the scene's reflectance field. Hereby, most of the scattering that degrades the floodlit image (Figure 1(b)) is avoided. The individual confocal samples however still contain global illumination effects due to in- and out-scattering along the selected paths. The recorded photons, instead of following the direct path given by the illumination and the viewing ray, travel through a larger volume, the so-called *photon banana* [FZC93] (see Section 3).

The second component, algorithmic *descattering* aims at separating the reflections of an object observed for individual rays into the component, which is due to direct reflection and the contribution due to global illumination effects. In algorithmic descattering, high frequency illumination patterns are often applied [NKGR06].

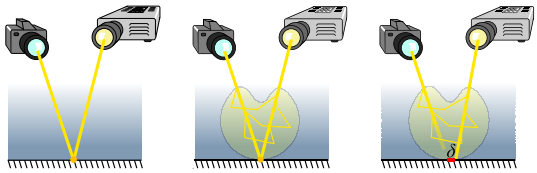
One of our contributions is a novel framework, where separation is performed on similar input data, by computing the difference of reflectance samples measured for adjacent illu-

mination samples or light paths. A qualitative analysis based on Monte-Carlo photon simulation indicates that our approach corresponds to reducing the original photon banana to a small channel around the direct path.
Furthermore, based on the proposed framework, we show that confocal imaging and descattering are similar but orthogonal concepts. Our novel descattering algorithm can be easily combined with confocal imaging, exploiting exactly the same illumination patterns for both approaches. We demonstrate the performance of our combined approach by computing significantly improved views through scattering fluids and into translucent objects.

2. Related Work

2.1. Separation of Reflection Components

The separation of measurements into different reflection components such as specular, diffuse, subsurface, or interreflections has so far been addressed by a number of different techniques. Using images captured with a polarization filter at different orientations, one can for example separate diffuse from specular reflections [RJ95, TRPE96, NFB97] or remove global illumination effects in participating media [SNN03, SK05, TS06] since multiple scattering tends to depolarize the incoming light.
Separation can also be performed using high frequency illumination patterns, based on the observation that only direct reflections will propagate high frequencies while global effects drastically dampen them. Nayar et al. [NKGR06] proposed computing the difference of measurements obtained from shifted high frequency patterns. By subtracting two measurements, the global component is removed. We compare our results to this method in detail later in this paper.
Narasimhan and Nayar [NNSK05] used swept line patterns for 3D-scanning through participating media. Based on the estimated distance they compute a "clear-air" view. They estimate the scattering parameters of a *single* scattering model and correct only for extinction along the illumination and viewing ray given the estimated depth. We make use of the geometric calibration technique proposed in this paper. In addition, we will also use swept line patterns for acquisition. In contrast to their method, we combine descattering with confocal imaging recovering a volume rather than single surfaces. Furthermore, we do not assume an explicit scattering model. Rather than just accounting for extinction along a single path, our descattering method corrects for multiple scattering in a photon banana by analyzing a local neighborhood of paths.
In our approach we will analyze the task of descattering in the context of reflectance fields [DHT*00]. Reflectance fields can be captured by scanning over all illumination rays as in [GLL*04] or structured light patterns [SY92]. Sen et al. [SCG*05] presented an adaptive algorithm that can parallelize the acquisition of multiple rays if they do not affect the same camera pixels. Since translucent objects result in rather dense reflectance fields almost no parallelization will be achieved. Garg et al. [GTLL06] developed a technique



(a) (b) (c)

Figure 2: *In translucent objects a reflectance sample for one pair of intersecting rays will always contain a direct contribution from the scattering at the intersection point (a) plus some contribution due to multiple scattering (b), the so-called photon-banana [FZC93]. (c) For a pair of adjacent but non-intersecting rays offset by δ, there will be no direct component but approximately the same global component. Combining both measurements, the direct component in (b) can be estimated.*

that determines reflectance samples of multiple illumination rays in parallel even for dense reflectance fields. The acquisition of a reflectance field in itself does not yet provide any means for descattering.
Seitz et al. [SMK05] presented a theory for inverse light transport computation based on measured reflectance fields. They propose a cancellation operator to remove the multiple scattering events from the reflectance fields, leaving the desired direct reflectance. This operator is computed using the inverse of the reflectance field, which does not always exist. In our framework we consider descattering as a local operation applied to a small neighborhood around each entry of the reflectance field; no inversion is necessary in our approach.

2.2. Volumetric Reconstruction

Confocal imaging, which we will explain in more detail in the next section, is related to other volumetric reconstruction techniques operating in the visible light range. Techniques for tomographic reconstruction of transparent objects have been proposed by Sharpe et al. and Trifonov et al. [SAP*02, TBH06]. In the first step, they determine the geometry of rays passing through the object by placing the possibly refracting object into some liquid with the same index of refraction. They measure the absorption along rays through the object from different directions and then perform algorithmic reconstruction techniques [GBH70]. Their technique only works for transparent objects since it is not resistant to multiple scattering.
Diffuse optical tomography [Arr99] tries to perform volumetric reconstruction of translucent objects but considers only the diffuse light transport. It must solve an ill-posed inversion problem to obtain any localized information.

3. Confocal Imaging

With confocal imaging one can perform optical sectioning, i.e. capturing images of slices through a given volumetric object, or one can capture the full confocal volume and then apply volume rendering techniques. In confocal imaging the

incident light and the observation rays from a large aperture are focused to a specific location within the object [CK96]. The shallow depth of field of a large aperture blurs the contribution of any point not in focus, and at the same time darkens it relative to the illuminated focused point, since any point out of focus will receive light only from a small fraction of the aperture. In mostly transparent scenes the influence of any point out of focus in the final image decays with r^4 where r is the distance to the focal plane.

Sectioned slices can be generated by illuminating individual points (for example by scanning or by means of a spinning Nipkow disk [EAH02]), or by illuminating and measuring multiple points at once with varying patterns. In the latter case the confocal image is computed after decorrelating the measurements for individual voxels, which can be done using pseudo random noise patterns [WJNK96, HHAJJ01] or periodic patterns [WNJ98, NJW97].

In the scanning configuration the illumination is directed to a single point v from an illumination aperture Ω_i. The computed confocal irradiance $L_{conf}(v)$ observed with aperture Ω_o is the integral over all rays intersecting at v:

$$L_{conf}(v) = \int_{\Omega_o} \int_{\Omega_i} L(\omega_i, v, \omega_o) d\omega_i d\omega_o. \tag{1}$$

3.1. Synthetic Aperture Confocal Imaging

Levoy et al. [LCV*04] performed confocal imaging with a synthetic aperture setup where the large aperture is sampled at a sparse and discrete set of directions from a set of cameras and projectors. In order to extract the information about a given voxel v, one selects the pair of rays $(\omega_o(v), \omega_i(v))$ that intersects at v. Given a set of cameras C and projectors P the confocal radiance is computed as:

$$L_{conf}(v) = \sum_{c \in C} \sum_{p \in P} L(\omega^p(v), \omega^c(v)). \tag{2}$$

Note, that the pair of rays $(\omega^p(v), \omega^c(v))$ that intersect at v will be different for each camera/projector pair due to the geometric setup. They are determined using the calibration data described in Section 6.1.

3.2. Separation Due to Confocal Imaging

The configuration for one pair of rays is depicted in Figure 2. In a purely *transparent* medium, the light transport will be governed by the reflection at v and the extinction along the incident light and the viewing ray (Figure 2(a)). Since each camera/projector pair observes v along a different path, the influence of the extinction along the paths is averaged in the confocal image, minimizing the influence of voxels other than v.

In the case of *translucent* media (Figure 2(b)), the light transport will include in- and outscattering along the path. The measurements will be affected by multiple scattering, effectively incorporating contributions from voxels within the resulting photon banana [FZC93]. The photon banana is formed by all photon traces that finally reach the camera pixel. It is clearly visible that the photon banana integrates the contributions of a much larger volume than the original direct path. At the same time, only a small fraction of the observed intensity will depend on the reflection at v.

When the observations of different aperture samples are combined, there is a large overlap in the photon bananas. The voxel v is no longer singled out but instead the net effect of the intersection of photon bananas will contribute to the confocal image. The influence of defocused voxels will no longer decay with r^4.

3.3. Descattering for Individual Rays

Besides the direct observation of the reflection at v and the extinction along the path, the reconstructed confocal image will contain multiple scattering events. We search for a separation into a non-local, or global, term L_g and the remaining local, or direct component L_d, which would correspond to the observation in transparent media:

$$L(\omega_o, \omega_i) = L_d(\omega_o, \omega_i) + L_g(\omega_o, \omega_i) \tag{3}$$

The key observation by Nayar et al. [NKGR06] is that the global component acts as a low-pass filter on high-frequencies in the incident illumination, while only the direct component will keep them. Posed differently, the global component L_g will be approximately the same in the vicinity of paths around (ω_o, ω_i) while the direct component L_d will be different.

Let us assume two neighboring rays ω_i and $\omega_i + \delta$ with the following two properties:

1. ω_o and ω_i intersect at v while ω_o and $\omega_i + \delta$ do not. It follows that $L_d(\omega_o, \omega_i + \delta) = 0$ (Figure 2(c)).
2. If the scene is homogeneous in the δ-neighborhood around the path (ω_o, ω_i) one can state that their global component is the same, $L_g(\omega_o, \omega_i) = L_g(\omega_o, \omega_i + \delta)$.

The direct component can in theory be determined as:

$$L_d(\omega_o, \omega_i) = L(\omega_o, \omega_i) - L(\omega_o, \omega_i + \delta), \text{ or} \tag{4}$$

$$= L(\omega_o, \omega_i) - L(\omega_o + \delta, \omega_i). \tag{5}$$

Due to the duality of the light transport [vH56, SCG*05] the role of viewing and illumination rays can be interchanged such that the difference can be computed for neighboring viewing rays as well.

The assumption of a homogeneous medium is often violated, especially at the places of interest, where there are volumetric features. In practice, we estimate the global component more robustly by incorporating a larger neighborhood. In the algorithm in Section 5 we construct a smooth neighborhood by weighted averaging.

4. Simulation

We use a Monte-Carlo photon simulator to qualitatively and quantitatively analyze the proposed descattering approach of Equation 4. The simulated scene represents the fish tank setup for which real measurements are demonstrated in Section 7.1. Photons enter a homogeneous scattering medium in the direction of a diffuse reflector which is embedded at

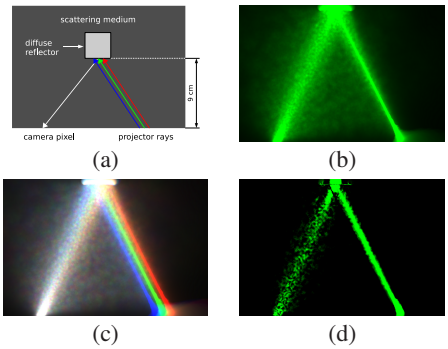


Figure 3: *(a) Simulation of photon traces from projector rays to a camera pixel in scattering media ($\sigma_t = 0.1$, $g = 0.9$, $\alpha = 0.9$). (b) Multiple scattering widens the direct path into the shape of a photon banana. (c) Neighboring illumination rays (red and blue) result in similar photon distributions. (d) Subtracting the averaged neighboring distributions extracts a photon channel which closely resembles the direct path.*

the center of the volume (Figure 3). The volume extends over $20 \times 20 \times 20$ cm^3 , while the diffuse target has a size of $2 \times 2 \times 2$ cm^3. The distance from the front plane of the volume to the target is 9 cm. We only consider photons that leave the volume at a predefined exit point, the camera pixel. Any other photon is discarded since it does not contribute to the measurement. For the selected photons, we record every scattering location and the energy along its trace.

4.1. From Photon Bananas to Photon Channels

In the first experiment (Figure 3), we simulated three neighboring incident light rays, marked red, green, and blue (c). The photon banana of the central illumination ray is shown in (b). By subtracting the averaged contribution of red and blue from the green photon banana, we reduce the green photon banana to a photon channel that encloses the original direct path (Figure 3(d)). Instead of the large volume of the full photon banana, only the much smaller volume covered by the photon channel will contribute to the final descattered image. In the confocal setting, the intersection of the tighter photon channels is significantly smaller than that of photon bananas.

While all photons within the channel might undergo multiple scattering events, they have a similar net effect as simple extinction in a transparent medium and therefore closely resemble the direct component.

4.2. Dependence on Scattering Parameters

In a second experiment, we investigate the dependence of the signal strength on varying scattering parameters of the medium: extinction coefficient (σ_t $[cm^{-1}]$) and the average scattering cosine (g) represented by the Henyey-Greenstein [HG41] model. The range of values covers most of the materials measured in [NGD*06]. As we focus at a point on the diffuse reflector we distinguish the total energy L_{signal} of photons arriving at the camera after interacting with the

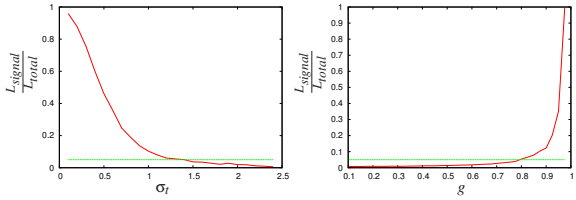


Figure 4: *Simulation results for varying scattering parameters. (a) L_{signal}/L_{total} falls off exponentially with increasing absorption ($\alpha = 0.9, g = 0.9$). A meaningful signal (above 5%) can be obtained up to an extinction coefficient of about 1.4 cm^{-1}. (b) SNR for varying forward scattering coefficient ($\sigma_t = 0.1$ cm^{-1}).*

diffuse reflector from the total energy carried by all photons L_{total} †.

Figure 4(a) shows how the ratio L_{signal}/L_{total} decreases with increasing σ_t. For an albedo of $\alpha = 0.9$ and a scattering angle of $g = 0.9$, the signal to noise ratio drops exponentially with σ_t. Around $\sigma_t = 1.4$ cm^{-1} the signal will become indistinguishable from the typical noise floor of the cameras (5%). A similar curve will be observed for the relative drop in signal with regard to the distance of the object to the projector and camera.

The signal-to-noise ratio furthermore depends significantly on the selected average scattering cosine g. Given the number of scattering events, photons are likely to reach the diffuse reflector only if the material is mostly forward scattering.

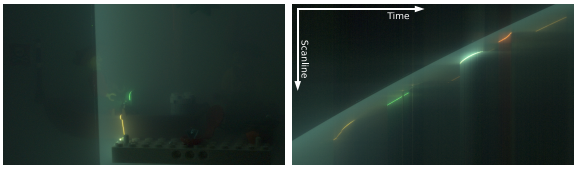
5. Analysis in the Context of Reflectance Fields

We will now analyze confocal imaging and descattering in the context of reflectance fields between pairs of cameras and projectors. The reflectance field T [DHT*00, MPDW03, SCG*05] describes how an incident light field L_i will be reflected by the scene, forming an outgoing light field $L_o = TL_i$. While the reflectance field for arbitrary light fields is an 8D function, the reflectance field between a single camera and one projector is only four-dimensional. Representing T as a tensor, each of its entries $T_{(\omega_o,\omega_i)}$ is the transport coefficient for one pair of camera and projector pixels or rays.

Figure 5(b) shows a 2D slice of the reflectance field of the island scene (Figure 1). The rows show always the same scan line of the camera while a vertical projector line (a) sweeps through the scene. The various object surfaces light up for different projector planes, in different rows. At the same time, multiple scattering in the surrounding fluid is clearly visible as it affects all entries in the reflectance field.

Under floodlit illumination (see Figure 1(b)), the contribution of all projector pixels, i.e. every line, integrates to the

† This is a slightly different definition than used before, where L_d incorporates only photons whose traces stays within the photon channel.



(a) (b)

Figure 5: *(a) A single vertical line projected into the island scene of Figure 1. (b) A horizontal slice of the recorded reflectance field. The opaque surfaces are clearly visible as well as the haze due to the scattering fluid.*

final image. Besides the actual surface, all entries that represent reflections by the fluid do contribute to the floodlit image. The interesting signal will be masked by the multiple scattering.

5.1. Confocal Imaging in Reflectance Fields

In confocal imaging, focusing to a specific voxel v corresponds to selecting the single entry in the reflectance field that is due to the particular selection of the view and the illumination ray. Extracting a volume slice means extracting all entries on a specific hyperplane whose geometry is defined by the camera/projector configuration and the selected volume slice.

The improved contrast in confocal images is mainly achieved by focusing on the entries of interest, where the viewing and the light ray do intersect, while discarding all those entries in the reflectance field that are just due to multiple scattering. Thus, compared to a floodlit image, in a confocal slice most of the global effects are separated out. However, the net reflectance recorded for a single intersecting pair of viewing and illumination rays still contains some global portion.

5.2. Descattered Confocal Imaging

As explained in the previous section, the global component L_g will be similar for neighboring non-intersecting paths. In reflectance fields, neighboring paths correspond to neighboring entries. Thus, one can interpret descattering as a local difference operator applied to the reflectance field of the scene.

Since in practice the homogeneity assumption (condition (2), Section 3.3) will be violated, a good approximation is found by averaging the difference over multiple samples in the neighborhood of $T_{(\omega_o,\omega_i)}$ as long as condition (1) (non-intersecting rays) is met. We compute the averaged difference by applying a Laplacian of Gaussian (LoG) in the neighborhood.

For simplicity, we only apply a 1D LoG kernel in the dimension of the illumination, i.e. for the camera pixels' recordings at different illumination.

6. Acquisition Approach

As outlined in Section 5, we first record a reflectance field T^{cp} for every camera projector pair (c, p) and then compute the descattered confocal volume by applying a local LoG kernel. In order to produce the result images we either directly render the recovered volume or extract a single slice.

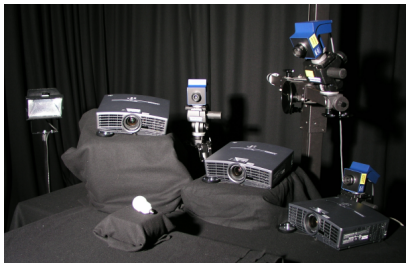


Figure 6: *Our synthetic aperture confocal imaging setup consists of three Jenoptik CFcool cameras and three Mitsubishi 490 DLP projectors.*

6.1. Setup

Our measurement setup (Figure 6 is similar to the one proposed by Levoy et al. [LCV*04]. We employ three cameras (Jenoptic CFcool) with about 1.3 megapixel resolution and three Mitsubishi 490 DLP projectors. The devices are placed at a distance of 60 *cm* away from the scene. The largest angle between the cameras and the projectors is about 65 degrees. In order to obtain a pixel-precise alignment between a camera and a projector we perform calibration as described by Narsimhan and Nayar [NNSK05]. A planar calibration target with a printed checker board is placed at three different known distances, moving it perpendicular to the calibration plane. The recorded images allow for precise recovery of the view-ray-to-voxel mapping $(c2v)$. Similarly, the three projectors are calibrated using projected checkerboard patterns, resulting in the illumination-ray-to-voxel mapping $(p2v)$. Assuming a rather smooth mapping we can easily compute the inverse mappings $(v2c)$ and $(v2p)$.

6.2. Planes of Light

Instead of recording a reflectance field with individual light rays, we sweep a plane of light through the volume generated by a line of one pixel width. Compared to the ray-based scanning, this reduces the number of required images from $O(N^2)$ to $O(N)$ for N projector rows. As Masselus et al. [MPDW03] demonstrated, one can approximate a reflectance field from observations of swept horizontal and vertical lines.

6.3. Image Acquisition

In order to maximize contrast, every projector is turned on individually and its sweep is recorded while the two other projectors are turned off. Because every projector emits some black-level even for off pixels, we first record a black frame for every camera c and every projector p. We then sweep a single pixel line horizontally and vertically and record images I_x^{cp} and subtract the black frames. Here, x is the projector coordinate of the line.

In order to produce a confocal volume, we apply Equation 2

(a) (b)

Figure 7: *(a) A bright illumination plane in front of the considered confocal plane will darken the image region if a symmetric descattering kernel is applied. (b) The one-sided kernel neglects those planes, rendering a slightly brighter image of the internal structures and resulting in more recovered detail.*

for every voxel, extracting the reflectance field samples for intersecting pairs of rays. Let $(s,t)^c$ be the projection location of voxel v in camera c determined by the calibrated $v2c$ map, and x be the x-coordinate of v being projected back to the projector image using $v2p$. We determine the confocal irradiance by averaging over the tri-linear interpolated reflectance field entries, first interpolating in the camera image and then between the selected slices with weight $\xi = x - \lfloor x \rfloor$:

$$\begin{aligned} L_{conf}(v) &= \sum_{c\in C}\sum_{p\in P}(1-\xi)\cdot I^{cp}_{(\lfloor x^p \rfloor)}(s,t)^c + \xi \cdot I^{cp}_{(\lceil x^p \rceil)}(s,t)^c \\ &= \sum_{c\in C}\sum_{p\in P} T^{cp}((s,t)^c, x^p). \end{aligned} \quad (6)$$

In order to compute a descattered confocal image this step is augmented by weighting the local illumination neighborhood with a LoG kernel of size $2K$:

$$L_d(v) = \sum_{c\in C}\sum_{p\in P}\sum_{k=-K}^{K} w(k) T^{cp}((s,t)^c, x^p + k). \quad (7)$$

We chose $K = 3$ and $w[-3,3] = \{-3,-10,1,24,1,-10,-3\}$. The results for horizontal and vertical sweeps are simply averaged.

Within a translucent object, for a non-coaxial camera and projector setup there will typically be one intersection of the ray (s,t) with every plane in the neighborhood. If there would be a strong scattering event in front of voxel v along (s,t) the computed $L_d(v)$ will be too small. Without loss of generality, let $x+k$, with $k<0$, be the planes where the intersection is in front of the voxel v. In this case, the direct component is better approximated by $L_d(v) = \sum_{c\in C}\sum_{p\in P}\sum_{k=0}^{K} w'(k) T^{cp}((s,t)^c, x^p + k)$ with a one-sided kernel $w'[0,k] = \{24,2,-20,-6\}$. The difference is shown in Figure 7.

7. Results

Our proposed combined descattered confocal imaging method is demonstrated in two different kinds of experiments. First, we demonstrate the superior result of our method in imaging through murky water, after which we acquire a partially translucent object.

7.1. Looking through Murky Water

Computing clear images in participating media has been addressed in quite a number of previous applications, e.g. [SNN03]. In this section we demonstrated the effect of performing descattering on top of confocal imaging in a fish tank filled with diluted milk and compare our results against those obtained with descattering or confocal imaging alone. We try to estimate clear-air views into the tank for varying milk concentration. The tank has a dimension of $39 \times 25 \times 24$ cm^3. We have acquired a resolution test chart placed at a distance of 9 cm from the front plane (see Figure 8) and show a scene with higher depth complexity in Figure 1.

We swept about 200 horizontal and 240 vertical lines from three projectors as shown in Figure 6 and recorded HDR images of the scene with the three cameras. The acquisition took roughly 1 hour followed by 15 minutes of further processing, most of which is due to I/O.

In Figure 8, the performance of various algorithms is compared for two different concentrations. The first column (a) shows the tank in floodlit illumination. In the next two columns, we applied the descattering algorithm proposed by Nayar et al. [NKGR06]. As input the technique requires a set of images recorded with periodic illumination pattern shifted in phase over time. Similar global patterns have been used even in the context of confocal imaging [WNJ98, NJW97]. A simple calculation approximates the direct component by $L_d = \frac{1}{1-b}(I_{\max} - I_{\min})$, where $I_{\min}$, $I_{\max}$ are the minimum/maximum of each pixel in the sequence, and b is the relative black-level of the projector. We have calculated the direct component on two different input sets. In column (b), the illumination pattern showed alternating on/off stripes of five pixels width, repeated over the entire projector pattern. The shifted sequence has been captured five times to reduce camera noise. In column (c), the same algorithm is applied to a synthetic input sequence generated from the reflectance field recorded by sweeping a single pixel line. The input patterns are computed by adding the captured images according to the previously described periodic patterns.

The next column (d) shows a single slice of the captured synthetic aperture confocal volume [LCV*04] as computed using Equation 6. Finally, we show our descattered confocal image in column (e) using three cameras and three projectors and a single camera/projector pair in (f) (c.f. Equation 7).

The pattern is barely visible in the floodlit configuration (a). In (b), descattering using images acquired with a periodic illumination pattern increases the contrast very little but increases the noise significantly. It produces even partially inverted patterns. This can be explained by the fact that half the projector pixels are on at any time. The tiny direct component, which is due to individual light rays, can hardly be separated from the camera noise because the added global component of half a million illumination rays is so much stronger in each input image. In the input images for column (c), the direct component is observed together with the

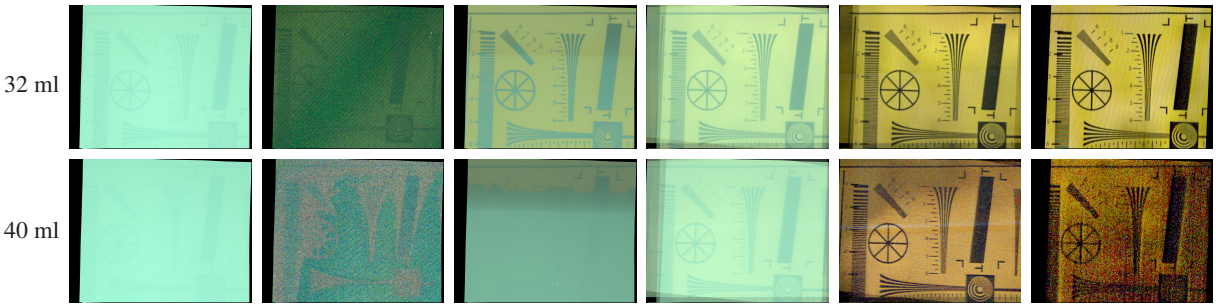


Figure 8: *Comparison for different milk concentrations in a* 23 *l water tank: (a) white frame, (b) fast separation [NKGR06] with periodic full-frame patterns, (c) fast separation with patterns synthesized from line sweeps, (d) a single confocal slice, (e) a descattered confocal slice incorporating all cameras and projectors, (f) a descattered slice from a single camera/projector pair. Please zoom in for full resolution.*

global component of one line only. The signal is recorded much stronger. However, during the synthesis of the input sequence, half of all reflectance field entries are again combined. Recovering the direct component works better for the 32 *ml* case but breaks down at 40 *ml* milk in 23 *l* water, due to the exponential fall-off of signal-to-noise ratio with increasing concentration (cf. Figure 4).

Combining the recordings from three cameras and three projectors, confocal imaging (d) succeeds in producing a much clearer image of the resolution chart since it considers only those entries of the reflectance field that contain a direct contribution, while the remainder of the entries is neglected. The pattern is visible in both concentrations indicating a precise calibration. Close to the image boundary the overlap of the swept projector volumes ends, resulting in intensity differences. While of good quality, the confocal images are still affected by multiple scattering. The pattern gets blurred more with increasing milk concentration. Since the target has a constant depth the method proposed by Narasimhan et al. [NNSK05] would just produce a scaled version of the confocal image.

Applying our proposed descattering algorithm (column (e)) removes the remaining global effects yielding higher contrast. It furthermore extracts sharper pictures of the direct component with higher resolution than any other method. Since only a small neighborhood is used for computing the direct component, the contrast is significantly increased compared to methods using global periodic kernels.

Note, that columns (c), (d), and (e) are computed from the same set of camera images. Only the processing is different. In Figure 8(e), one can furthermore observe a color shift from 32 to 40 *ml* in our results. This indicates that we are indeed extracting the direct component. Increasing the concentration increases the out-scattering along the path. Since milk scatters blue more strongly than red, the blue part of the direct component is filtered out faster than the red one. The difference between our descattered confocal imaging approach and only applying local descattering becomes obvious when comparing Figure 8(e) with (f) which is much more noisy.

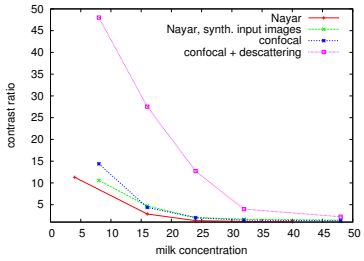


Figure 9: *The maximum contrast between a white and a dark patch in the result images of the approaches in Figure 8 decreases with the scattering density (ml milk in* 23*l water).*

The achieved contrast of the various methods for different milk concentrations is plotted in Figure 9. As expected, the signal drops exponentially with increasing concentration, as also predicted by the simulations (c.f. Figure 4). Descattering with large kernels and pure confocal imaging consistently produce a smaller contrast, rendering the image unrecognizable much earlier than with our descattered confocal imaging technique.

Island Scene

While the previous scene consisted of a single plane at fixed depth, the second experiment captures a scene with some depth complexity (Figure 1). It touches the front of the fish tank and has a depth of approximately 9.5 *cm*.

We have captured the full confocal and descattered confocal volume to recover the volumetric 3D structure of the scene as can be seen in the offset view in Figure 1(e). We render orthographic views of the scene using maximum intensity projection and increase the reconstructed intensity linearly with depth in order to counteract the extinction.

Compared to pure confocal imaging our descattered confocal imaging technique can look much deeper into the volume. The pirate on the right and the palm tree in the back are recovered. As for the resolution chart, our results are sharper and the colors are more saturated.

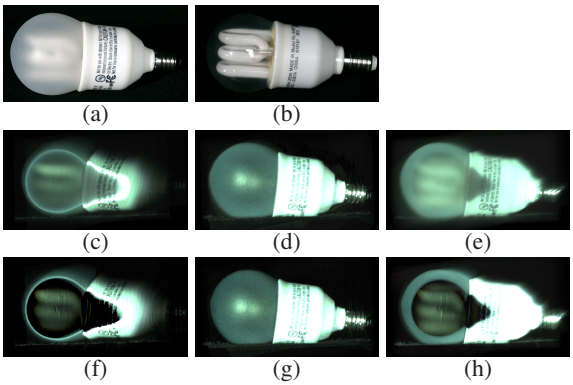


Figure 10: *Confocal (middle row) and descattered confocal images (bottom row) of an energy saving light bulb with frosted coating (a). In (b), the coating has been removed for demonstration. The depth discrimination for a single confocal slice (c) increases using descattering (f). (d) and (g) show the full integrated volume, from which in (e) and (h) the front slices have been removed.*

In the offset view one can observe some artifacts around depth discontinuities where the assumption of a homogeneous medium around the path is violated. For rather bright objects such as the flag, the silhouette is partially extruded in depth due to small errors in the camera calibration and the low aperture sampling rate.
The ground plane is not rendered white in the confocal images because a couple of volume slices have been removed from the front in order to exclude any contribution from the fish tank's front glass.

7.2. Looking into Translucent Objects

We further investigate the internal structure of a light bulb with frosted coating (Figure 10). Here, partially translucent layers are separated by air.
The first column shows one slice of the reconstructed volumes. While one sees a blurry image of the internal tubes in the confocal image (middle row), applying descattering (bottom row) reveals sharper detail. In addition, the considered plane is much more focused in the descattered version. The second column shows the integrated volume, i.e. all slices added up. The outer surface is not affected by our descattering algorithm. In the third column, the front most volume slices have been clipped away, revealing the internal structure of the fluorescent tubes. While being slightly more noisy the descattered image provides a significantly higher contrast. Upon further inspection (see video), one can even make out the traces of a tiny, completely transparent glass pipe in front of the white fluorescent tubes. It is hardly visible in the pure confocal volume.

7.3. Limitations

As indicated in Figure 4 the performance of active light volumetric acquisition methods, such as ours, with regard to contrast or penetration depth will be limited in principle by the scattering density σ_t, since the ratio of the direct to the global contribution will decrease exponentially. A low noise camera and a high-contrast projector can help increasing the maximum recoverable depth.
Increasing the number of aperture samples can further help both increasing the contrast, as well as minimizing the aliasing artifacts which are currently visible in the island scene when solid surfaces come out of focus.
So far, our calibration assumes a constant index of refraction in order to predict the intersection of viewing and illumination rays. As can be seen in the bulb example, the internal structures appear slightly distorted due to refraction.

8. Conclusion

The performance of confocal imaging in translucent media can be significantly improved by combining it with our novel descattering procedure based on a local descattering kernel. The same input data is used for both steps. While confocal imaging reduces the effect of multiple scattering by selecting individual light paths, our descattering operator further removes global effects from the selected paths by analyzing the vicinity of paths: The photon banana corresponding to a confocal sample is effectively reduced to photon channels around the direct path.
For descattering, we have so far applied a local LoG kernel in the illumination domain only. In the future, one might obtain even better descattering by computing spatially varying deconvolution of the volume around the confocal entries in the reflectance field, or by further investigating the duality of light transport.
Our approach for combined descattering and confocal imaging currently operates at macroscopic scales. Using a confocal microscope, or by augmenting a light field microscope [LNA*06] by a light field projector, migration to microscopic scales should be straight forward.

Acknowledgements

We would like to thank our reviewers for their suggestions and comments and Christian Schmaltz for his photon simulation. This work has been partially funded by the DFG Emmy Noether fellowship (Le 1341/1-1) and the Max Planck Center for Visual Computing and Communication (BMBF-FKZ01IMC01).

References

[Arr99] ARRIDGE S.: Optical tomography in medical imaging. *Inverse Problems 15* (1999), R41–R93.

[CK96] CORLE T. R., KINO G. S.: *Confocal Scanning Optical Microscopy and Related Imaging Systems*. Academic Press, 1996.

[DHT*00] DEBEVEC P., HAWKINS T., TCHOU C., DUIKER H.-P., SAROKIN W., SAGAR M.: Acquiring the reflectance field of a human face. In *Proceedings of ACM SIGGRAPH 2000* (July 2000), Computer Graphics Proceedings, Annual Conference Series, pp. 145–156.

[EAH02] EGNER A., ANDRESEN V., HELL S. W.: Comparison of the axial resolution of practical nipkow-disk confocal fluorescence microscopy with that of multifocal multiphoton microscopy: theor y and experiment. *Journal of Microscopy 206*, 1 (2002), 24–32.

[FZC93] FENG S., ZENG F., CHANCE B.: Monte Carlo Simulations of Photon Migration Path Distributions in Multiple Scattering Media. In *Photon Migration and Imaging in Random Media and Tissues, Proc. of SPIE Vol. 1888* (1993), pp. 78–89.

[GBH70] GORDON R., BENDER R., HERMAN G.: Algebraic Reconstruction Techniques (ART) for Three-dimensional Electron Microscopy and X-ray Phtography. *Journal of Theoretical Biology 29* (1970), 471–481.

[GLL*04] GOESELE M., LENSCH H. P. A., LANG J., FUCHS C., SEIDEL H.-P.: Disco: acquisition of translucent objects. In *ACM SIGGRAPH 2004* (2004), pp. 835–844.

[GTLL06] GARG G., TALVALA E.-V., LEVOY M., LENSCH H. P. A.: Symmetric photography: Exploiting data-sparseness in reflectance fields. In *Proceedings of Eurographics Symposium on Rendering* (2006), pp. 251–262.

[HG41] HENYEY L. G., GREENSTEIN J. L.: Diffuse radiation in the Galaxy. *Astrophys. J 93* (Jan. 1941), 70–83.

[HHAJJ01] HEINTZMANN R., HANLEY Q. S., ANRDT-JOVIN D., JOVIN T. M.: A dual path programmable array microscope (pam) simultaneous acquisition of conjugate and non-conjugate images. *Journal of Microscopy 204*, 2 (2001), 119–137.

[LCV*04] LEVOY M., CHEN B., VAISH V., HOROWITZ M., MCDOWALL I., BOLAS M.: Synthetic aperture confocal imaging. *ACM TOG (Proc. ACM SIGGRAPH 2004) 23*, 3 (2004), 825–834.

[LNA*06] LEVOY M., NG R., ADAMS A., FOOTER M., HOROWITZ M.: Light field microscopy. *ACM Trans. Graph. 25*, 3 (2006), 924–934.

[MPDW03] MASSELUS V., PEERS P., DUTRÉ P., WILLEMS Y. D.: Relighting with 4D incident light fields. *ACM Transactions on Graphics 22*, 3 (2003), 613–620.

[NFB97] NAYAR S. K., FANG X. S., BOULT T.: Separation of reflection components using color and polarization. *IJCV 21*, 3 (1997), 163–186.

[NGD*06] NARASIMHAN S. G., GUPTA M., DONNER C., RAMAMOORTHI R., NAYAR S. K., JENSEN H. W.: Acquiring scattering properties of participating media by dilution. *ACM Trans. on Graphics* (2006), 1003–1012.

[NJW97] NEIL M. A. A., JUSKAITIS R., WILSON T.: Method of obtaining optical sectioning by using structured light in a conventional microscope. *Optics Letters 22*, 24 (1997).

[NKGR06] NAYAR S. K., KRISHNAN G., GROSSBERG M. D., RASKAR R.: Fast separation of direct and global components of a scene using high frequency illumination. *ACM TOG (Proc. ACM SIGGRAPH 2006) 25*, 3 (2006), 935–944.

[NNSK05] NARASIMHAN S., NAYAR S., SUN B., KOPPAL S.: Structured light in scattering media. *IEEE International Conference on Computer Vision (ICCV) I* (2005), 420–427.

[RJ95] ROWE M. P., JR. E. N. P.: Polarization-difference imaging: a biologically inspired technique for observation through scattering media. *Optics Letters 20*, 6 (1995), 608–610.

[SAP*02] SHARPE J., AHLGREN U., PERRY P., HILL B., ROSS A., HECKSHER-SORENSEN J., BALDOCK R., DAVIDSON D.: Optical Projection Tomography as a Tool for 3D Microscopy and Gene Expression Studies. *Science 296*, 19 (2002).

[SCG*05] SEN P., CHEN B., GARG G., MARSCHNER S., HOROWITZ M., LEVOY M., LENSCH H. P. A.: Dual photography. *ACM Trans. on Graph 24*, 3 (2005), 745–755.

[SK05] SCHECHNER Y. Y., KARPEL N.: Recovery of underwater visibility and structure by polarization analysis. *IEEE Journal of Oceanic Engineering 30*, 3 (2005), 570–587.

[SMK05] SEITZ S., MATSUSHITA Y., KUTULAKOS K.: A theory of inverse light transport. In *ICCV* (2005), pp. 1440–1447.

[SNN03] SCHECHNER Y. Y., NARSIMHAN S. G., NAYAR S. K.: Polarization-based vision through haze. *Applied Optics 42*, 3 (2003), 511–525.

[SY92] SEBERRY J., YAMADA M.: Hadamard matrices, sequences, and block designs. In *Dinitz, J. H. and Stinson, D. R., editors (1992), Contemporary Design Theory: A Collection of Essays* (1992), pp. 431–560.

[TBH06] TRIFONOV B., BRADLEY D., HEIDRICH W.: Tomographic reconstruction of transparent objects. In *Rendering Techniques 2006: 17th Eurographics Workshop on Rendering* (2006), pp. 51–60.

[TRPE96] TYO J. S., ROWE M. P., PUGH E. N., ENGHETA N.: Target detection in optically scattering media by polarization difference imaging. *App. Opt. 35* (1996), 639–647.

[TS06] TREIBITZ T., SCHECHNER Y. Y.: Instant 3descatter. In *Proceedings of CVPR* (2006), pp. 1861–1868.

[vH56] VON HELMHOLTZ H.: Treatise on physiological optics, 1856. The Optical Society of America. Electronic edition(2001):University of Pennsylvania http://psych.upenn.edu/backuslab/helmholtz.

[WJNK96] WILSON T., JUSKAITIS R., NEIL M., KOZUBEK M.: Confocal microscopy by aperture correlation. *Optics Letters 21*, 3 (1996).

[WNJ98] WILSON T., NEIL M. A. A., JUSKAITIS R.: Real-time three-dimensional imaging of macroscopic structures. *Journal of Microscopy 191*, 2 (1998), 113–220.