

## Compensating for Motion During Direct-Global Separation

Supreeth Achar, Stephen T. Nuske, and Srinivasa G. Narasimhan

Robotics Institute, Carnegie Mellon University

### Abstract

*Separating the direct and global components of radiance can aid shape recovery algorithms and can provide useful information about materials in a scene. Practical methods for finding the direct and global components use multiple images captured under varying illumination patterns and require the scene, light source and camera to remain stationary during the image acquisition process. In this paper, we develop a motion compensation method that relaxes this condition and allows direct-global separation to be performed on video sequences of dynamic scenes captured by moving projector-camera systems. Key to our method is being able to register frames in a video sequence to each other in the presence of time varying, high frequency active illumination patterns. We compare our motion compensated method to alternatives such as single shot separation and frame interleaving as well as ground truth. We present results on challenging video sequences that include various types of motions and deformations in scenes that contain complex materials like fabric, skin, leaves and wax.*

### 1. Introduction

The radiance of a scene point illuminated by a light source is the sum of the direct and global components. The direct component is the light from the source that undergoes a single reflection in the scene before reaching the observer. The global component is due to indirect lighting from inter reflections, subsurface scattering, volumetric scattering and diffusion. Separating the direct and global components of illumination provides valuable insights into how light interacts with a scene. Being able to extract the direct component of illumination can improve the performance of classical photometry based algorithms like shape from shading as well as structured light reconstruction which typically do not account for global effects. Having access to the two components separately instead of their sum can also provide useful information about the physical properties of the materials present in a scene.

An efficient method for finding the global and direct components was first proposed in [14]. The method involves capturing images while illuminating the scene with a sequence of high frequency patterns. Theoretically, only two images are needed but practical implementations use at least three images, with more images leading to better results. The light source, camera and scene need to remain stationary during image acquisition. In the same work, a single image variant of the method that could be used for moving scenes was also proposed, but it generates low spatial resolution results.

A similar trade-off exists in active illumination to recover geometry. Single shot structured light methods [9] can be used on dynamic scenes but have low spatial resolution while multi-image methods [17] produce high quality depth estimates but require the scene to remain stationary. There has been work on developing motion compensation schemes to allow multi-image structured light algorithms to be applied to dynamic scenes [10, 20]. One approach is interleaving the projector patterns for structure estimation with uniformly lighting for motion tracking. Most structured light algorithms do not account for global illumination and those that do [3, 5] require many additional images.

In this work, we address motion compensation in the context of direct-global separation. We develop a method that relaxes the requirement that the scene and camera remain static during direct-global separation. This allows separation to be performed on video sequences in which the projector-camera system and/or the scene are moving. We assume that the underlying global and direct components of a scene point vary only slightly over small motions. This means that if the frames in a temporal window can be aligned, the separation technique in [14] can be applied to the aligned frames. Optical flow techniques can not be used directly because of the time varying patterns illuminating the scene. Instead, we use a simple image formation model to approximate scene appearance under uniform lighting. We use these relit images to aid alignment and then estimate the global and direct components from the aligned images.

Compared to single image separation, our method pro-

duces more detailed, higher resolution results. We use all the frames in a temporal window for estimating the global and direct components. No frames are used exclusively for tracking, so our method can handle faster motions than interleaving at a given frame rate. We use a colocated projector-camera system which allows us to avoid the difficult projector-camera pixel correspondence and 3D reconstruction problem.

We show that our method compensates for motion effectively and generates separation results close to ground truth. We show that not compensating for motion introduces significant artifacts in the separation and compare our method to alternatives such as single shot separation and interleaving. The scenes we demonstrate our method on contain materials such as wood, plastics, fabric, wax, leaves and human skin that interact with light in complex ways. We show that our method is able to compensate for both rigid motions and non-rigid deformations in the scene.

### 1.1. Related Work

The original work on direct-global separation [14] describes methods for separation using active illumination and source occluders. With active illumination, the separation can be performed using three sinusoid patterns, but the best results with practical projector-camera systems require around 20 high frequency pattern images. A method that uses a single image was also presented, but it generates results at a fraction of the projector’s resolution which is undesirable since most projector-camera systems are projector resolution limited. In [15] an optical processing method that can be used to directly acquire the global component of illumination is presented. Global illumination and projector defocus were modeled jointly in [6] for depth recovery in scenes with significant global light transport effects. In [4], the separation technique was extended to scenes illuminated by multiple controllable light sources. Their goal was to extract the direct component for each light source to aid structure recovery techniques where global illumination is often a severe source of systematic error.

The need for motion compensation also arises in structured light for 3D estimation. Taguchi et al. [19] developed structured light patterns that can be decoded both spatially and temporally which allows for motion adaptation. Stationary and slow moving regions are decoded temporally yielding high quality depth estimates and the algorithm falls back onto spatial decoding in fast moving regions.

In [20] a motion compensation method for the phase shift structured light algorithm is presented. Motion in the scene during image acquisition causes high frequency ripples in the phase estimates which are corrected by modeling the true phase as a locally linear function.

Motion estimation and compensation in image sequences with projected patterns is often done by interleav-

ing the patterns with uniform lighting [21]. A similar approach is used in the structured light motion compensation scheme in [10] where patterns for structure estimation are interleaved with patterns optimized for estimating motion. Interleaving is a valid approach for our problem, but it increases the number of images that are needed and can introduce registration artifacts if the motion is not smooth.

Most techniques for optical flow are based on brightness and gradient constancy assumptions. Because we illuminate the scene with time varying, high frequency patterns, these assumptions are not valid. Illumination robust optical flow methods have been designed based on photometric invariants [13] and physical models [8]. Computationally, most optical flow methods are based on local linearization of images. Since our images are dominated by the pattern, this linearization causes problems. An alternative optical flow formulation was derived in [18] that uses a direct search to compute optical flow and which can accommodate arbitrary data loss terms. We use a variant of this direct search method to refine our alignments.

### 1.2. Limitations

We do not model changes in the underlying direct and global components at a scene point within a small temporal window. This is generally valid for global effects, diffuse reflections and smooth gloss reflections but not for sharp specularities. We assume that there is no blurring of the images due to motion and that the scene lies completely within the depth of field of the projector-camera system.

## 2. Image Formation Model

The brightness  $I^t(x)$  of a pixel  $x$  at time  $t$  is a combination of the direct component  $I_d^t$  and global component  $I_g^t$ . When a binary pattern illuminates the scene, the direct component is modulated by the pattern. If the pattern has an equal number of bright and dark pixels and has high spatial frequency compared to  $I_g^t$ , the contribution of the global illumination to the brightness is  $\frac{1}{2}I_g^t$  [14]. Thus we have

$$I^t(x) = \frac{1}{2}I_g^t(x) + s^t(x)I_d^t(x) \quad (1)$$

where,  $s^t$  is the value of the projected pattern at camera pixel  $x$ . We colocate our projector and camera so the mapping between projector and camera pixels is fixed and independent of scene geometry. Even though the patterns are binary, the value of  $s^t$  at a pixel can be continuous because real projectors do not have ideal step responses and the projector and camera pixels need not be aligned. An example of the model is illustrated in Figure 1.

For convenience and without loss of generality, we fix  $t = 0$  to be the current frame, at the center of a temporal sliding window. As a scene point moves with respect to the projector and camera, its global and direct components

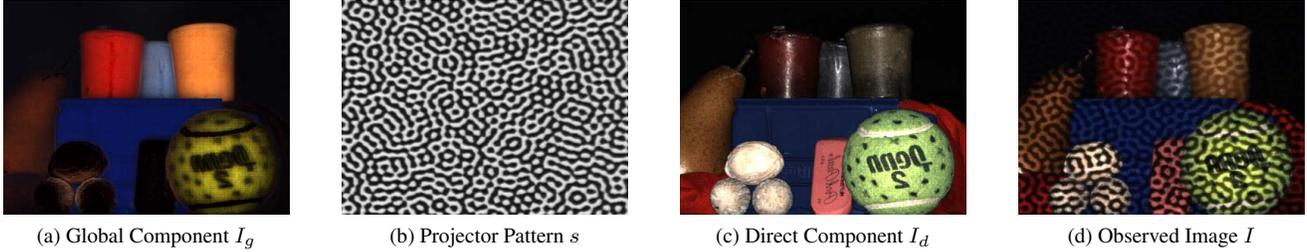


Figure 1: Image Formation Model: The observed image  $I$  is modeled as a linear combination of the global component ( $I_g$ ) and a projector pattern ( $s$ ) modulated version of the direct component ( $I_d$ ).  $I = \frac{1}{2}I_g + sI_d$ . The specularities on the candles appear in the direct image and most of the color is due to subsurface scattering in the wax and appears in the global image.

will change. We assume that the motion within a sliding window is small enough for these changes to be negligible. This allows us to relate the global and direct components at time instant  $t$  in the sliding window to time 0

$$I_g^0(x) \approx I_g^t(W^t(x)) \quad I_d^0(x) \approx I_d^t(W^t(x))$$

where,  $W^t$  is an (unknown) warping function that aligns the view at time 0 to the view at time  $t$ . At  $t = 0$ ,  $W$  is just an identity mapping.  $W^t$  depends on the geometry of the scene and the motion of the scene and projector-camera system. We use the notation  $W \circ I$  to denote the image that results from applying warp  $W$  to image  $I$ .

### 3. Motion Estimation and Compensation

We compute the direct-global separation at a frame in the video sequence using a small temporal sliding window centered at that frame. We seek to compensate for the motion that occurs inside a temporal sliding window so that the frames can be aligned to each other. With the help of the image formation model, we estimate how the scene would have appeared at each time instant under uniform lighting instead of the patterned illumination. We use these fully lit versions of the images to find warps that align the images and then refine the warps. Once the images are aligned we can compute the global and direct components robustly.

#### 3.1. Estimating Scene Appearance Under Uniform Illumination

Finding the warps that align frames is difficult because of the time varying pattern being projected onto the scene. The patterns violate the brightness and contrast constancy assumptions most optical flow methods rely on. To aid alignment, we compute an approximation of how the scene would have appeared ( $\tilde{I}_f^t$ ) under uniform illumination from the frame  $I^t$  and the pattern  $s^t$  used to illuminate the scene. These fully lit estimates are better suited for image alignment than the original patterned frames.

Under uniform illumination, the brightness at a pixel is the sum of two unknowns, the direct component and the global component  $I_f^t(x) = I_g^t(x) + I_d^t(x)$ . The two unknowns are related by equation 1. The problem is under constrained and can not be solved uniquely because there is only one equation for every two unknowns. To find an approximate solution to the problem, we introduce a regularizer that enforces piecewise spatial continuity of the estimated global and direct components ( $\tilde{I}_g^t$  and  $\tilde{I}_d^t$  respectively). The loss function minimized is

$$L(\tilde{I}_g^t, \tilde{I}_d^t) = \|I^t - \frac{1}{2}\tilde{I}_g^t - s^t\tilde{I}_d^t\|_2^2 + \lambda_g TV(\tilde{I}_g^t) + \lambda_d TV(\tilde{I}_d^t) \quad (2)$$

where,  $\lambda_g$  and  $\lambda_d$  are smoothing parameters for the global and direct components.  $TV(Z)$  is the isotropic total variation of the function  $Z(x, y)$ .

$$TV(Z) = \sum_{Domain(Z)} \sqrt{\left(\frac{\partial Z}{\partial x}\right)^2 + \left(\frac{\partial Z}{\partial y}\right)^2} \quad (3)$$

For color images, we sum the total variation of each color channel. This objective function is similar to those used in L1 regularized image deblurring and denoising [2]. It is convex in  $\tilde{I}_g$  and  $\tilde{I}_d$  and can be solved efficiently to find a globally optimal solution. An example of the illumination pattern removal is shown in Figure 2.

Any high frequency illumination pattern can be used to perform direct-global separation. We use random bandpass patterns similar to those used for structure estimation in [3]. The relit images sometimes contain minor artifacts along the edges in the patterns. These artifacts are caused by projector blur and small errors in the colocation between the projector and camera. Using random patterns prevents correlation between these artifacts across time from affecting the registration process.

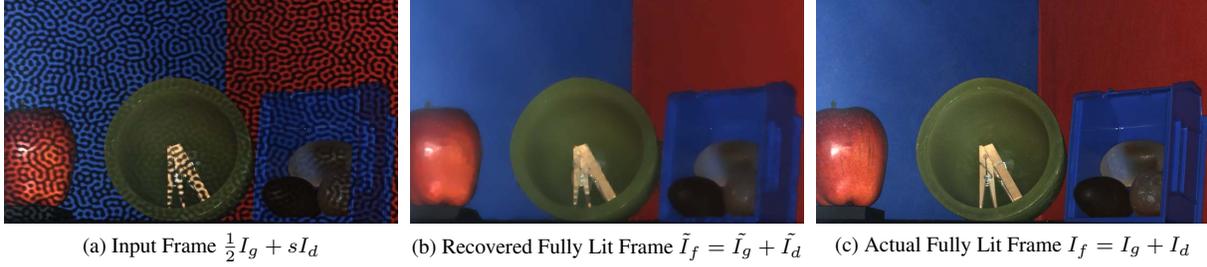


Figure 2: Removing the Illumination Pattern: An image captured with a high frequency illumination pattern (a) is relit to form an estimate (b) of how the scene would have appeared under uniform lighting. For reference, the actual appearance of the scene under uniform lighting is shown in (c). The relit estimate (b) captures the structure of the scene but smooths over some of the finer detail that is visible in (c). Relit images are used to perform image alignment on the original images

### 3.2. Registering Images

To align a frame to the center frame, we could simply compute optical flow between the relit frames. But for scene points that are illuminated in both images, it is better for the warp to match the original image pixel values  $I^0(x)$  and  $I^t(x)$  than the smoothed, relit estimates  $\tilde{I}_f^0(x)$  and  $\tilde{I}_f^t(x)$ . At scene points that are not illuminated or which are illuminated in one frame but not the other, matching the relit estimates is preferable. We implement this idea in two stages. First, we compute the warping that best aligns  $\tilde{I}_f^t$  to  $\tilde{I}_f^0$  with variational optical flow [1]. This initial warp estimate is then refined by minimizing the following cost functional:

$$C(W^t) = \sum_x (1 - \alpha(x, W^t)) |\tilde{I}_f^0(x) - \tilde{I}_f^t(W^t(x))| + \sum_x \alpha(x, W^t) |I^0(x) - I^t(W^t(x))| + \gamma TV(W^t) \quad (4)$$

where,  $\alpha(x, W^t)$  is a weight that is high when a point is lit ( $s$  close to 1) in both the center frame  $I^0$  and the current frame  $I^t$ . The total variation term is a regularizer to ensure that the computed warp is piecewise continuous. We set  $\alpha(x, W^t) = (s(x) s^t(W^t(x)))^2$  when both  $s$  values are above 0.8 and to 0 otherwise.

The data term in the cost functional does not linearize well, so we minimize it approximately using the direct search algorithm proposed in [18]. Because we are seeking to correct small errors in an existing optical flow estimate we search for a refined warp at each pixel using a small window centered around the original warp estimate.

If the motion that occurs in a sliding window is large, optical flow may fail to correctly align some frames to the center frame. We detect poorly aligned frames by thresholding the correlation between the warped frame  $W^t \circ \tilde{I}_f^t$  and center frame  $\tilde{I}_f^0$ . Poorly aligned frames are discarded from the sliding window.

### 3.3. Computing Direct-Global Separation

Once the frames in a window have been warped to align with the center frame, we in effect have a set of images of the scene captured from the same viewpoint with different illumination patterns.

If the set of patterns is designed such that it can be guaranteed that for each camera pixel there will be at least one frame where the corresponding projector pixel is fully lit and another frame where it is fully dark, the separation can be performed using simple pixel-wise min and max operations over the frames in the sliding window [14]. Alternatively, since the projector pattern values ( $s^t$ ) at each pixel are known, the global and direct components can be determined by fitting a line to the observed brightness values at a pixel using equation 1. For this line fit to make sense, each pixel needs to be observed under a range of projector pattern values.

Since scene structure and motion are not known a priori, a pattern sequence cannot be designed that is guaranteed to satisfy either of the above two criteria. As a result, there will be pixels in the image where the global and direct components can not be estimated well because the projector brightness did not change sufficiently at the corresponding scene point. We fill in these pixels by extending the idea of equation 2 to multiple images. We search for piecewise continuous global and direct components that are a good fit to the observed aligned image data by minimizing

$$L(I_g^0, I_d^0) = \sum_{t \in T} \|W^t \circ I^t - \frac{1}{2} I_g - (W^t \circ s^t) I_d^0\|_2^2 + \lambda_g TV(I_g^0) + \lambda_d TV(I_d^0) \quad (5)$$

where,  $T$  is the sliding window of frames selected about the center frame. At scene points where a variety of different  $s$  values were projected,  $I_g^0(x)$  and  $I_d^0(x)$  are estimated confidently as only a single line fits the data term. At pixels where the value of  $s$  was similar throughout the temporal

window, many separations are plausible fits to the data and the smoothness terms help resolve the ambiguity.

## 4. Results

### 4.1. Experimental Setup

In our experiments, the scenes were illuminated using a  $1024 \times 768$  DLP projector. For the experimental results presented in 4.2, the scenes were imaged with a Point Grey Grasshopper camera at 10 frames per second. For the results on deformable objects (4.3), the scenes were acquired at 60 frames per second. For all experiments, the camera was radiometrically calibrated to have a linear response curve and the camera and projector were colocated using a plate beam splitter. The experimental setup is shown in Figure 4.

An offline calibration step is used to find the image  $s$  corresponding to each illumination pattern. Each pattern is projected onto a diffuse, planar white surface and imaged by the camera. To correct for projector vignetting, all images were normalized with respect to a reference image of the same planar surface while fully lit by the projector. This calibration needs to be performed only once.

### 4.2. Comparisons on Rigidly Moving Scenes

The goal of these experiments is to compare the direct and global components generated by our algorithm on moving scenes to ground truth and to analyze the effect of temporal window size on separation accuracy.

Ground truth was acquired by first capturing 25 frames of a scene while projecting checkerboard patterns at different offsets. These frames were captured while the scene and camera were stationary. The direct and global components calculated on these 25 frames are used as ground truth (RMS Error 0). We then captured a video sequence with the scene in motion while patterns were being projected. The pose of the first frame in the video matches the pose from which the ground truth frames were acquired.

This experiment was performed from two different poses on two scenes (see Figs. 2 and 5) yielding a total of four trials. RMS errors against ground truth for different separation methods averaged over the four trials are shown in Figure 3.

In the static case, the results improve in quality as the number of frames used increases ('Static' in Fig. 3). We also used the regularization method described in 3.3 on the static sequence ('Static regularized' in Fig. 3). The regularization improves performance when the number of frames is small and many pixels have not seen enough different projector pattern values. However, it smooths over some of the fine details in the scene and so it does not perform as well the unregularized technique when more frames are used.

For the video sequence corresponding to each trial, we tested our motion compensation method with different sliding window sizes using the first frame as the window cen-

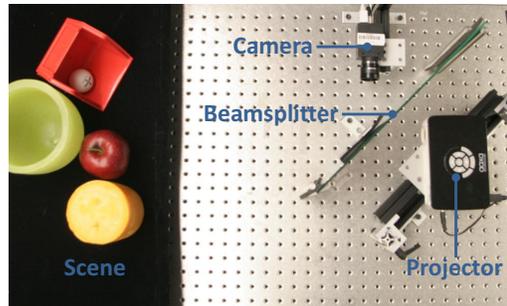


Figure 4: Experimental Setup: Our experiments were performed using a camera and projector colocated with a plate beamsplitter. We also have a portable version of the setup.

ter. We evaluated the motion compensation with the warp refinement described in 3.2 ('Moving Refined' in Fig. 3) and with the unrefined warps ('Moving' in Fig. 3). We also tested an interleaved approach where the projector alternates between patterns and uniform illumination ('Interleaved' in Fig. 3). Warps computed between the fully lit images are interpolated to align the patterned images. These aligned images are then separated as described in 3.3.

When the number of frames used is small, the regularized static method and the proposed motion compensated methods perform similarly. As the number of frames increases, the improvement in the motion compensated output reduces and then stops. When the window size is large, the frames near the edges of the sliding window can not be aligned to the center frame because the viewpoint changes are too large and the global and direct components of the scene points change appreciably. The motion compensation algorithm automatically discards these frames and they yield no improvement in the results.

Interleaving does not perform as well as our proposed method. The temporal window available for performing separation on dynamic scenes is small. With interleaving, only half the frames in this window can be used for computing separation. Additionally, if the motion is not smooth, the interpolated warps do not align images accurately. These problems could be solved by using higher frame rates, but at a given capture rate our method can handle faster, more complex motion than interleaving. Fig. 5 shows results from our motion compensation algorithm and interleaving with different temporal window sizes in an example scene.

Direct-global separation using a single image creates blurred, low resolution results. Examples are shown in the supplementary material.

### 4.3. Deformable Motions

In addition to rigid motion, our algorithm can also compensate for deformations and articulated motions. Fig. 6 shows some examples of separations obtained from videos

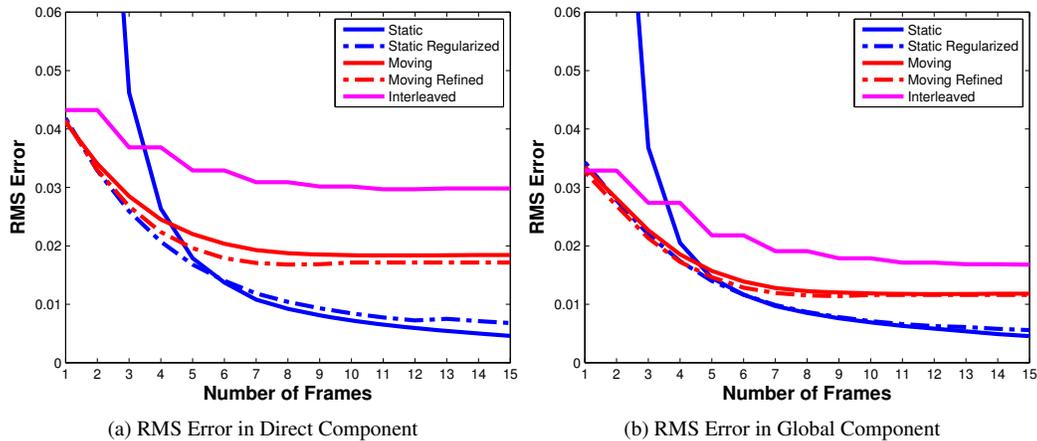


Figure 3: Comparison of Separation Techniques: The RMS errors in the direct (a) and global (b) components for different separation techniques as the number of frames used is varied. The blue ‘static’ curves are from direct-global separation on stationary scenes and represent the best possible performance a method could achieve for a given number of frames. The red ‘moving’ curves are from using our motion compensation algorithm on moving scenes. When the number of frames is small, the motion compensation method performs just as well on the moving sequences as normal separation on an equal number of static frames. When the window size increases, frames far away from the window center are discarded because alignment fails and so performance of the motion compensated algorithm levels off.

of moving hands, deforming faces and a plant being blown in the wind. For comparison, we have included the results without motion compensation. When the motion estimation was switched off, we selected the window size for each example that gave the best result to compare against our method. Without motion compensation, the results are blurred and edges in the scene (around the fingers for example) are corrupted. Subsurface scattering occurs in the skin so most of the color appears in the global component. The supplementary material includes these videos.

## 5. Discussion

Although we do not model the changes in global and direct components that occur within a small temporal window, our method is still able to handle broad specular lobes like shiny surfaces on wax and highlights on skin. Sharp specularities and specular inter reflections such as those from polished metal surfaces would cause both the image alignment and component separation steps to break down. The fast direct-global separation algorithm for static scenes can handle sharp specularities but not specular inter reflections. One solution would be to use crossed polarization filters to remove specular reflections. Methods for removing specular components from images include [11] and [12].

The experiments presented were designed so that motion blur and defocus would not cause additional errors. To use our method in less controlled settings, we would need to consider the trade-off between acquisition time and accu-

racy. Using shorter exposure times and smaller apertures to avoid motion blur and defocus means that less light reaches the camera and image noise becomes more of a problem. We would need to consider how computational photography methods like coded aperture for motion deblurring [16] and light efficient photography [7] could be applied.

**Acknowledgements:** This research was supported in parts by NSF grant IIS-0964562, ONR grant N00014-11-1-0295, a University Transportation Center T-SET grant and a National Grape and Wine Initiative grant.

## References

- [1] T. Brox, N. Papenberg, and J. Weickert. High Accuracy Optical Flow Estimation Based on a Theory for Warping. *European Conf. on Computer Vision*, 4(May):25–36, 2004. 4
- [2] A. Chambolle. An Algorithm for Total Variation Minimization and Applications. *Journal of Mathematical Imaging and Vision*, 20(1-2):89–97, 2004. 3
- [3] V. Couture, N. Martin, and S. Roy. Unstructured light scanning to overcome interreflections. *ICCV*, 2011. 1, 3
- [4] J. Gu, T. Kobayashi, M. Gupta, and S. K. Nayar. Multiplexed illumination for scene recovery in the presence of global illumination. *Intl. Conf. on Computer Vision*, 2011. 2
- [5] M. Gupta, A. Agrawal, A. Veeraraghavan, and S. G. Narasimhan. A Practical Approach to 3D Scanning in the Presence of Interreflections, Subsurface Scattering and Defocus. *IJCV*, 102(1-3):33–55, 2012. 1
- [6] M. Gupta, Y. Tian, S. G. Narasimhan, and L. Zhang. A Combined Theory of Defocused Illumination and Global Light Transport. *IJCV*, 98(2):146–167, Oct. 2011. 2

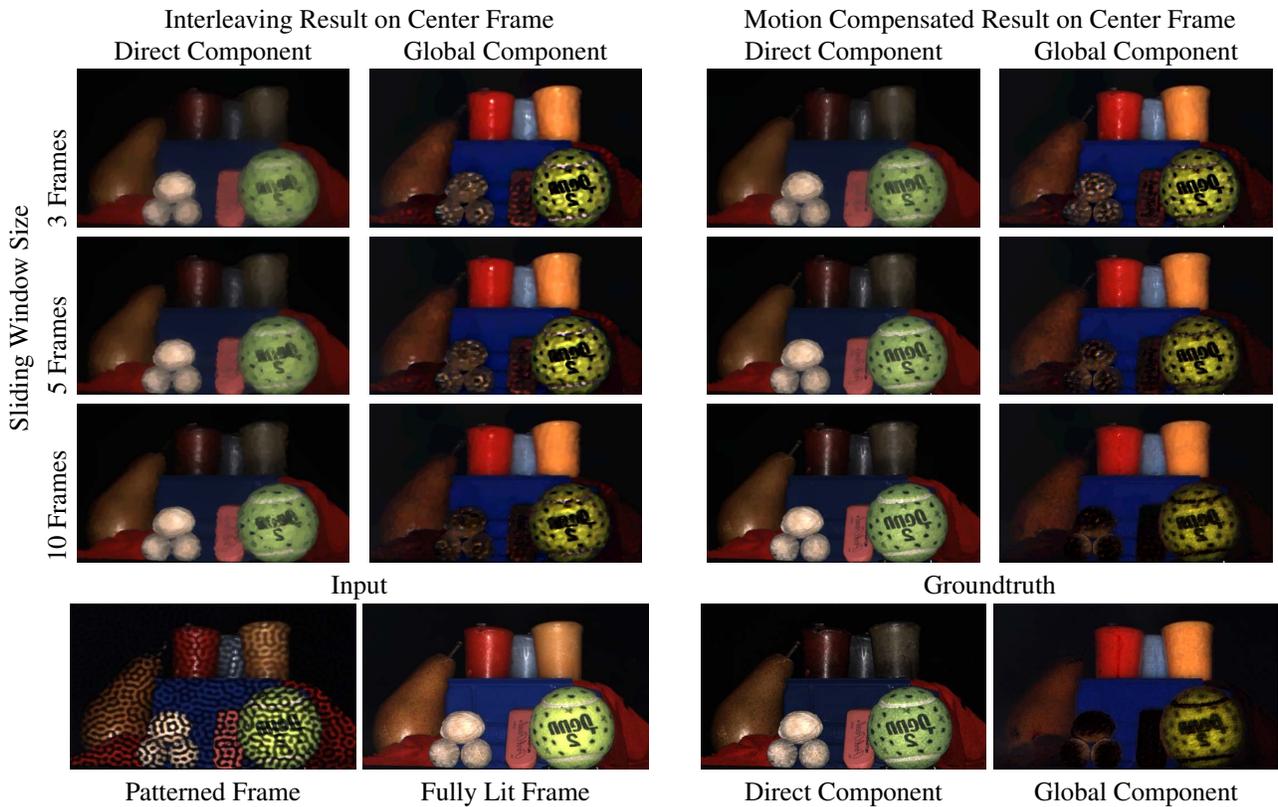


Figure 5: Comparison to Interleaving: The separation from interleaving and our algorithm as the number of frames in the sliding window is changed. The camera is panning across the scene. The first three rows show the results generated by both methods for various sliding window sizes (global images are shown at 2 times actual brightness). Our method makes more efficient use of images than interleaving because no frames are needed exclusively for tracking. Separations that resolve a given level of detail can be obtained with a smaller temporal sliding window than an interleaving approach. For instance, our method resolves the text on the tennis ball much earlier and more clearly than interleaving.

- [7] S. W. Hasinoff and K. N. Kutulakos. Light-efficient photography. *IEEE PAMI*, 33(11):2203–14, Nov. 2011. 6
- [8] H. Haussecker and D. Fleet. Computing optical flow with physical models of brightness variation. *IEEE PAMI*, 23(6), June 2001. 2
- [9] H. Kawasaki, R. Furukawa, R. Sagawa, and Y. Yagi. Dynamic scene shape reconstruction using a single structured light pattern. *CVPR*, 2008. 1
- [10] S. Konig and S. Gumhold. Image-Based Motion Compensation for Structured Light Scanning of Dynamic Surfaces. *Dynamic 3D Imaging Workshop*, 2007. 1, 2
- [11] B. Lamond and P. Debevec. Fast Image-based Separation of Diffuse and Specular Reflections. *SIGGRAPH Sketches*, 2007. 6
- [12] S. P. Mallick, T. Zickler, P. N. Belhumeur, and D. J. Kriegman. Specularity Removal in Images and Videos : A PDE Approach. *European Conf. on Computer Vision*, 2006. 6
- [13] Y. Mileva. Illumination-Robust Variational Optical Flow with Photometric Invariants. *DAGM Conference on Pattern Recognition*, pages 152–162, 2007. 2
- [14] S. K. Nayar, G. Krishnan, M. D. Grossberg, and R. Raskar. Fast separation of direct and global components of a scene using high frequency illumination. *ACM Transactions on Graphics*, 25(3):935, July 2006. 1, 2, 4
- [15] M. O’Toole, R. Raskar, and K. N. Kutulakos. Primal-dual coding to probe light transport. *ACM Transactions on Graphics*, 31(4):1–11, July 2012. 2
- [16] R. Raskar, A. Agarwal, and J. Tumblin. Coded Exposure Photography : Motion Deblurring using Fluttered Shutter. *ACM Transactions on Graphics*, 25(3):795–804, 2006. 6
- [17] V. Srinivasan, H. C. Liu, and M. Halioua. Automated phase-measuring profilometry. *Applied Optics*, 24(2), 1985. 1
- [18] F. Steinbrücker and D. Cremers. Large Displacement Optical Flow Computation without Warping. *ICCV*, 2009. 2, 4
- [19] Y. Taguchi, A. Agrawal, and O. Tuzel. Motion-Aware Structured Light Using Spatio-Temporal Decodable Patterns. *ECCV*, 2012. 2
- [20] T. Weise, B. Leibe, and L. Van Gool. Fast 3D Scanning with Automatic Motion Compensation. *CVPR*, 2007. 1, 2
- [21] L. Zhang, N. Snavely, B. Curless, and S. Seitz. Spacetime Faces: High-Resolution Capture for Modeling and Animation. *SIGGRAPH*, 2004. 2

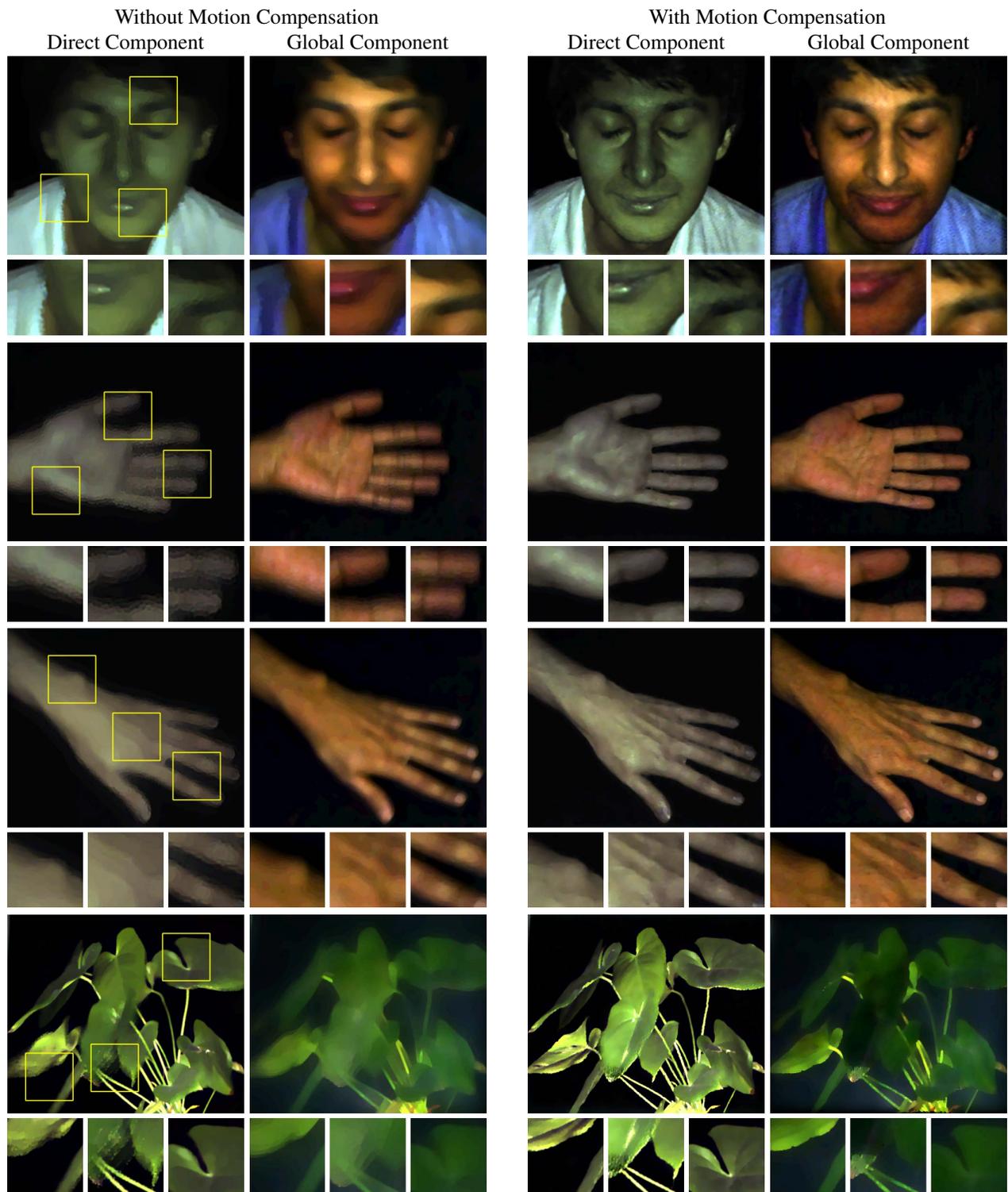


Figure 6: Direct-Global Separation on Deformable Objects: On the left are the global and direct components estimated without any motion compensation for faces changing expression, articulating hands and a plant moving in the breeze. Many details get blurred away like the hair, and the lines on the palm. Other motion artifacts are clearly visible on the fingers and leaves. The two columns on the right show the component estimates on the same frames using our motion compensation method. With motion compensation, many of the artifacts are corrected and a lot more of the original scene detail is recovered.