Electronic Preprint for Journal of Information Processing Vol.26

Regular Paper

Separating the Direct and Global Components of a Single Image

Art Subpa-asa[1,a)] Ying Fu[2] Yinqiang Zheng[3] Toshiyuki Amano[4] Imari Sato[1,3]

Received: May 11, 2018, Accepted: July 10, 2018

Abstract: Direct and global component separation is an approach to study the light transport that provides a basic understanding of the property of a scene. The conventional technique for separation relies on multiple images or an approximation which results in loss of spatial resolution. In this article, we propose a novel single image separation technique by introducing a linear basis equation with full resolution. We evaluate the data independent Fourier basis and learning-based PCA basis to locate the better basis representation of direct and global components. We carefully analyze the importance of high spatial frequency pattern to the effectiveness of our technique. Moreover, we propose the performance enhancement technique to reduce memory usage and computation time for practical implementation. The experimental results confirm that our proposed method delivers higher separation accuracy and better image quality than the previous methods and is applicable to challenging video sequences.

Keywords: light transport, layer separation, basis representation, projector-camera system

1. Introduction

The appearance of a scene is a result of light paths traveling through the scene, ranging from simple direct reflectance to complex light phenomena such as subsurface scattering, interreflection, and volumetric scattering, all of which are instances of light transport. The understanding of light transport is beneficial to various computer vision tasks, such as shape reconstruction and image descattering. The direct and global components are introduced to reduce the complexity of light transport. Direct component represents direct reflection between surfaces and observant observed directly at the point of incidence without any interference from other phenomena. Global component is a combination of complex phenomena such as scattering and interreflection which occur further away from an observation point. An illustration of light transport is shown in **Fig. 1**.

Global components can be separated according to the differences between lit and unlit pixels under different illumination patterns by projecting a series of shifted high frequency patterns [18]. **Figure 2** (a) shows an image under high frequency pattern, Fig. 2 (b) and 2 (c) show results of direct and global components separation using multiple shifted high frequency pattern.

However, this technique requires multiple images of a scene under the spatial shifted high frequency pattern and cannot apply directly to a dynamic scene without motion compensation. While the single shot technique has been proposed by assuming a similarity between adjacent pixels, this results in loss of spatial resolution and tends to lower the separation quality. Figure 2 (d) and Fig. 2 (e) show results from using single-shot technique with pixelated upsampling results. Therefore, the challenging objective of this study is to separate direct and global components using a single image without sacrificing the spatial resolution.

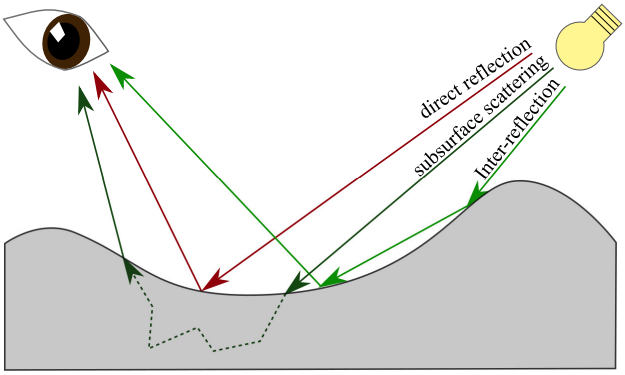
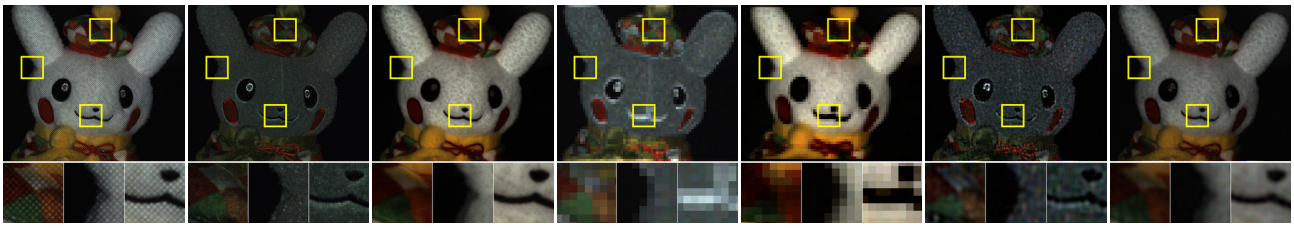


Fig. 1 Light transport can be categorized as direct component (red line) and global component (green line). Examples of the global component include inter-reflection and subsurface scattering.

By leveraging the linear basis representation and the modulation effect of high frequency illumination, we introduce a novel technique for separating direct and global components from a single frame, without loss of spatial resolution (shown in Fig. 2 (f) and Fig. 2 (g)). Differences between the two bases have been observed and used to design a proper illumination pattern to reduce the dependency on two sets of bases. We further separate direct and global components by designing a linear system. The effectiveness of our approach is demonstrated on simulated data, and real images/videos captured by a standard off-the-shelf camera and a projector mounted in a coaxial system.

A comprehensive dataset of light transport is not available at the current time since capturing the whole light transport requires a huge effort. To train a valid basis, we prepare a simplified light

[1] Tokyo Institute of Technology, Meguro, Tokyo 152–8550, Japan
[2] Beijing Institute of Technology, Haidian, Beijing China
[3] National Institute of Informatics, Chiyoda, Tokyo 101–0003, Japan
[4] Wakayama University, Wakayama 640–8510, Japan
[a)] art.s.aa@m.titech.ac.jp



(a) (b) (c) (d) (e) (f) (g)

Fig. 2 Result of direct and global components separation. (a) show input images with high frequency illumination. (b) and (c) show ground-truth of direct and global components from multiple images using the fast separation technique [18]. (d) and (e) show low resolution direct and global components from single image using the fast separation technique, while (f) and (g) show full resolution direct and global components recovered from our technique.

transport dataset consisting of 91 tuples of raw, direct and global images. It contains varieties of objects, with diverse surface texture and material, captured in a controlled environment. This dataset also allows us to analyze properties of direct and global components, and to find differences between different bases. By integrating the basis of direct and global components from our database with our single image separation technique, we achieve an improvement over other bases without having to take into account the differences between components.

In summary, our major contributions are that we

- Propose to represent direct and global components based on linear basis representation for full resolution recovery of these components;
- Carefully examine dependency of direct and global components and show that a high spatial frequency illumination pattern contributes to resolving the ambiguity and separating these two components robustly;
- Introduce a new direct and global components dataset, analyze the dependency and the differences between the two components to improve the separation quality compared to the use of other bases.
- Propose a performance enhancement method to reduce memory usage and computation time of our separation process.
- Set up a coaxial system by using a standard off-the-shelf camera and a projector to capture real images of still and moving scenes, and demonstrate the effectiveness of our separation method.

The rest of this article is organized as follows. Section 2 provides literature reviews of direct and global component separation. Section 3 discusses the ambiguity of direct and global components observed in a single image and describes our approach for robust separation. Section 4 describes the practical implementation of our technique to reduce memory usage and computation time. Our experiment setup and results are shown in Section 5. Finally, conclusions are drawn, and future directions of our research are discussed in Section 6.

2. Related Works

Direct and global component separation was first introduced using the fast separation technique [18]. Series of shifted high frequency patterns were used to separate the pixel intensity into two components by observing differences between the intensity of lit and unlit pixels. However, this technique was applicable only to static scenes. A workaround approach for single image separation was later introduced by assuming similarity among the neighboring pixels. Unfortunately, the overall resolution of a result image was severely reduced, and artifacts were found in the areas of different materials.

Effort to improve the quality of separation techniques had been proposed in various studies. Modulation between two 1 d signals was introduced to replace a constant high frequency pattern to improve accuracy in scenes with high glossiness material [6]. A multiple projectors based approach had been developed by using multiplex sinusoid patterns to reduce the number of images to three [9]. The interreflection from different sources decomposed from a single image was used to improve the performance of image recoloring [5]. Diffuse structured light generated by placing the diffuser in front of a projector was used to improve the efficiency in shadow areas of image [17]. Motion compensation technique was proposed to improve quality of scenes with tiny movement such as human body [2]. Global components which are considered a multi focus of illumination were also used for scene reconstruction [1]. Extension of microscopy imaging was also introduced to separate transmitted and scattering apart at microscale [25].

Researchers also proposed techniques to separate additional components aside from direct and global. The global component could be further divided into near-range and far-range components depending on the distance between an incident and the observed points. Separation of these components is viable by using a combination of logical coding and patterns with different sizes [12], or large series of frequency-varying sinusoid patterns for diffuse scenes [23]. Iterative technique for inverse light transport was utilized for separating inter-reflection bounces with different numbers [4]. The depth layer within a translucent material was separated by projecting series of different frequency pattern [29]. Independent component analysis was also applied to multi-frequency direct components to separate direct component, sub-surface scattering and inter-reflection [27]. Tailored features had been used for specific domains such as face [8], [30], skin [24], hair [33], fur [32], volumetric fluid [10], translucent object [14], [15], glaring [28] and outdoor scene [13].

Specialized equipment had been introduced to separate precise light transport matrix and their components. A high-speed camera was used to capture temporal dithering between frames of

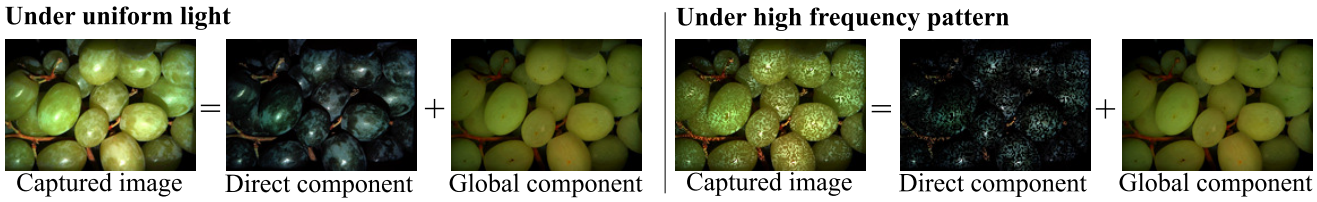


Fig. 3 Direct and global components under different illuminations. Under high spatial frequency illumination (second column), direct component is modulated by an illumination pattern spatially, while global component keeps constant up to a scaling factor.

DLP projector in real-time separation [16]. A light field camera and projector were utilized to transfer global illumination between images [7]. Time-of-flight camera was used to capture light transverse through translucent objects and scenes by picoseconds and capable of separating the light component by projecting short pulse of light [31] or phase signal [11]. Photometric mixer device (PMD) was used to acquire light transport matrix by sampling different temporal frequency of sensor and light source [20]. Moreover, dual-coding between camera pixels modulated by a co-axial set up of LCD panel and a projector was introduced to acquire the whole light transport matrix [22]. The energy efficient technique with digital micromirror device (DMD) was also utilized to separate direct component, near-range global component and far-range global component in real time [19], [21].

In this work, we propose a technique to separate direct and global components from a single image without sacrificing their spatial resolution, using only a standard off-the-shelf camera and a projector. A preliminary version of this article appeared in Ref. [26]. Compared with Ref. [26], we have prepared a database to verify the effectiveness of the basis expression and the effect of illumination patterns on the separation quality. We have also proposed practical solutions to reduce memory usage and computational time dramatically.

3. Direct and Global Component Separation

In this section, we describe our proposed technique to separate direct and global components and explore the viability of our approach.

3.1 Direct and Global Component Model

We define image i at a pixel p under uniform illumination to consist of direct component i_d and global component i_g in an equation as follows:

$$i(p) = i_d(p) + i_g(p). \tag{1}$$

We have learned that direct and global components interact differently to the projection of illumination with a pattern. Specifically, the direct component is modulated by the light projected on observation point, while the global component is a combination of light interaction further away from the observation point such as interreflection and subsurface scattering. The tangling complex phenomena among sources of the global component makes the separation process difficult. Fortunately, according to Ref. [18], the global component is assumed to be independent under a high enough spatial frequency illumination. The global component in the presence of high spatial frequency illumination is approximated to k-percent proportional to that under uniform illumination, where k is the average intensity of an illumination pattern. The relation between the direct and global components in high spatial frequency illumination is illustrated in **Fig. 3**.

We first vectorize the image of N pixels into $\boldsymbol{i} = [i(p_1), i(p_2), \cdots, i(p_N)]^T$, whose corresponding illumination is denoted by $\boldsymbol{l} = [l(p_1), l(p_2), \cdots, l(p_N)]^T$. The spatial relationship between the illumination pattern and the image can be described as:

$$\boldsymbol{i} = L\boldsymbol{i_d} + \left(\frac{\int \boldsymbol{l}(p)\,dp}{N}\right)\boldsymbol{i_g} = L\boldsymbol{i_d} + k\boldsymbol{i_g} = L\boldsymbol{i_d} + \hat{\boldsymbol{i}}_{\boldsymbol{g}}, \tag{2}$$

where $L = diag(\boldsymbol{l})$. $\boldsymbol{i_d} = [i_d(p_1), i_d(p_2), \cdots, i_d(p_N)]^T$ and $\boldsymbol{i_g} = [i_g(p_1), i_g(p_2), \cdots, i_g(p_N)]^T$ represent direct and global components under the uniform illumination in vectorized form.

3.2 Variable Reduction using Linear Basis Representation

It is clear that we cannot directly solve $\boldsymbol{i_d}$ and $\hat{\boldsymbol{i}}_{\boldsymbol{g}}$ in Eq. (2) because the number of unknown variables is double that of the number of measurements. We suggest the use of linear basis representation to reduce the number of unknown variables.

Assuming a smoothness constraint on direct and global components over 2D image space, these components can be defined as linear combination of predefined bases and unknown coefficients as follows

$$\boldsymbol{i_d} = D\boldsymbol{\alpha}, \quad \hat{\boldsymbol{i}}_{\boldsymbol{g}} = G\boldsymbol{\beta}, \tag{3}$$

where $D_{N\times n}$ and $G_{N\times m}$ are the bases for the direct and global components. $\boldsymbol{\alpha}_{n\times 1}$ and $\boldsymbol{\beta}_{m\times 1}$ are the coefficient vectors.

From Eq. (2) and Eq. (3), we obtain the relation between image in high spatial frequency illumination and basis representation as:

$$\boldsymbol{i} = L\boldsymbol{i_d} + \hat{\boldsymbol{i}}_{\boldsymbol{g}} = LD\boldsymbol{\alpha} + G\boldsymbol{\beta} = \begin{bmatrix} LD & G \end{bmatrix}\begin{bmatrix} \boldsymbol{\alpha}^T & \boldsymbol{\beta}^T \end{bmatrix}^T, \tag{4}$$

in which the number of unknown variables has been reduced to the number of bases $m+n$, which are usually less than the number of image pixels N.

In this article, we introduce two basis representations: Fourier and principal component analysis (PCA) bases.

Fourier basis has been widely used for image compensation due to a proven ability to represent general spatially smooth images and is independent of input data. In this work, we use 2D Fourier basis to describe direct and global components as baseline basis which is independent of the data.

Principle Component Analysis (PCA) basis is a statistical method to find an uncorrelated basis from correlated observations. PCA basis represents the order of the variance, which

means that higher order basis is more significant than the lower order. In this article, we use training data based on specific raw (image under a uniform illumination), direct and global component dataset. Assume that an image in a general environment is a mixture of direct and global components, the basis of the general image can be used to represent both direct and global components. On the other hand, a dataset of the direct and global component is used to specifically target the domain problem. Further discussion about PCA direct and global component basis is in Section 3.3. The results show that frequencies of Fourier basis and PCA basis gradually increase in the order. The samples of Fourier and PCA bases used in our experiment are shown in **Fig. 4**.

3.3 PCA Basis of Direct and Global Components

To study differences between direct and global components, we directly analyze the dependency between direct component, global component, and raw image in uniform illumination. In the following subsection, we introduce our light transport dataset and analyze the relation and dependency between each component aiming to improve the quality of separation.

3.3.1 Dataset

Firstly, We create a dataset by collecting light transport components using the fast separation technique with multiple images [18]. The dataset consists of 91 3-tuples of the image, and each tuple consists of raw (image in an uniform illumination pattern), direct and global images. Our selected objects contain various material from dense material with low sub-surface scattering like wood and paper, to material with high sub-surface scattering like fur, food and rubber. The objects also consist of different shapes from primitive shapes, i.e. sphere and cylinder, to complex shape with a high amount of inter-reflection. The images have varying scales from a focused region of the surface to the whole object. **Figure 5** illustrates some examples of our dataset.

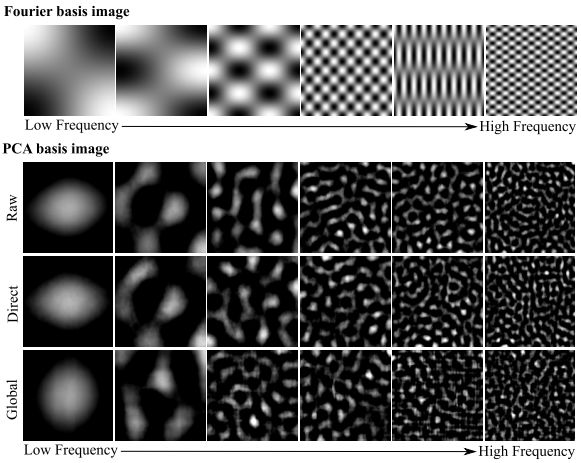


Fig. 4 Samples of Fourier basis and PCA basis; PCA basis are learned from three different sources, including raw image (image under uniform illumination), direct component and global component.

Our dataset acquisition system consists of a simple camera and projector setup. We project high frequency 8×8 checker pattern into a scene shifting by 1 pixel each. Each scene is also captured with uniform illumination (without a pattern) to obtain a raw image. Each image is directly cropped to eliminate background and area which is not illuminated to ensure that every pixel contains accurate separation of both direct and global components.

Later, we extract the PCA basis from our dataset. Each tuple is segmented directly into 2,000 small chunks of 128×128 pixels. The location of each chunk is selected randomly but consistently between raw, direct and global components. We separately compute the basis image for each component to derive individual basis images.

By doing so, we obtain basis images for raw, direct and global components. However, further analysis of each set of basis images is needed to investigate the correlation between each basis which directly affect the quality of separation.

3.3.2 Correlation between Direct, Global and Raw Image Bases

To solve Eq. (4), the matrix $M = \begin{bmatrix} LD & G \end{bmatrix}$ has to be invertible, which means that all columns in M should be independent. This indicates that if D and G are statistically independent, the equation is solvable without any requirements for the illuminated pattern L. Unfortunately, as will be shown later, these two bases are somehow correlated.

Firstly, we analyze the covariance between each pair of basis image in the tuple. We apply cosine similarity to measure the similarity between the two basis vectors. Independent pair of the vectors leads to low similarity. We measure cosine similarity for three combinations: direct and global, direct and raw, global and raw. **Figure 6** (a) illustrates cosine similarity of the first 64 components and the values of the most similar image. We have found common trends in each combination of the components, that is, dominant bases have high similarity. This trend indicates that dominant bases contain a low-frequency basis regardless of the type of component. On the other hand, minor bases contain

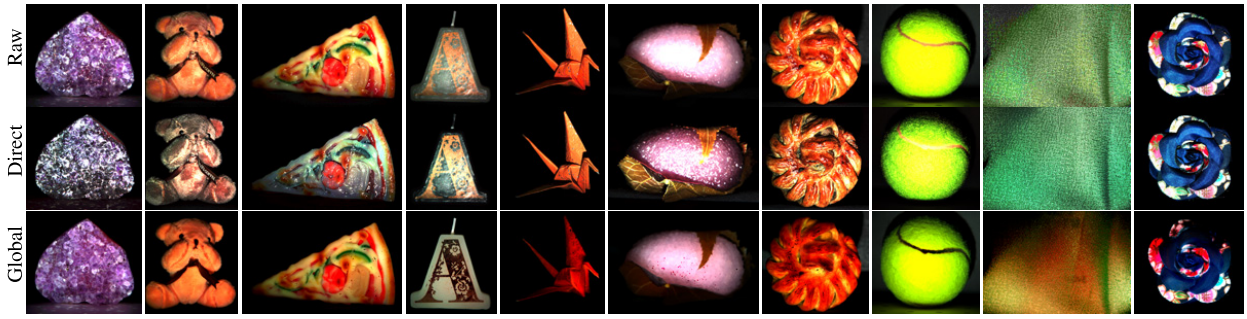


Fig. 5 Example of images in our dataset. Tuples of raw image (image in uniform light), direct component and global component.

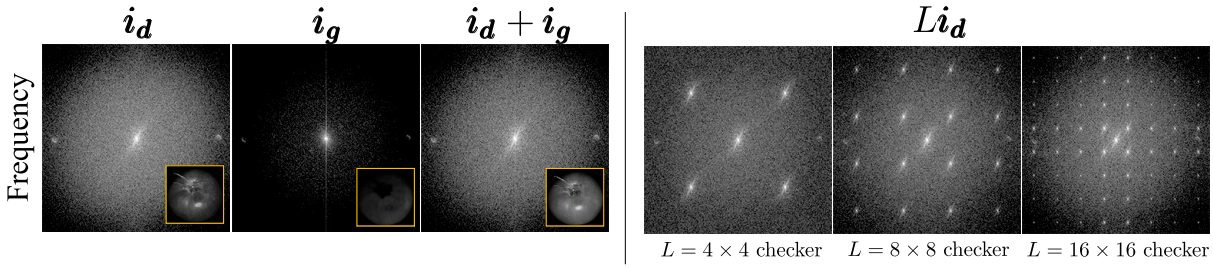
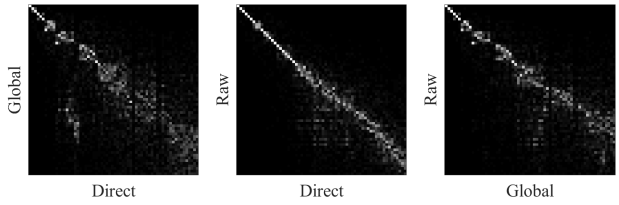
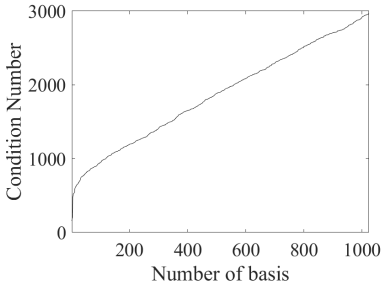


Fig. 7 The intensity of images in frequency domain while illuminating pattern of different frequencies.



(a) Cosine similarity between pairs of component bases.

(b) Separation result without high frequency illumination.



(c) Condition numbers of M_{flat}

Fig. 6 Analysis of relationship between PCA basis image. (a) Cosine similarity between the first 64 basis image, Intensity of pixel indicate the similarity between pair of basis image in respective index: (left) Direct and Global, (Center) Direct and Raw (image in uniform light), (Right) Global and Raw. (b) Poor separation result is due to the dependency between direct and global bases. (c) The condition number of direct and global components when the number of bases increases.

high frequency details such as texture, which depends on the type of component. The results demonstrate the statistical difference between light transport component (direct and global) and image in uniform illumination and the difference between direct and global components, and therefore confirm usage of the basis image learned from light transport component over generic image.

Secondly, we investigate dependency of the matrix $M_{flat} = \begin{bmatrix} D & G \end{bmatrix}$. We conduct an experiment to measure the condition number of the matrix M_{flat} by increasing the number of bases. The results of this experiment are shown in Fig. 6 (c). We have found that the condition number becomes greater when the number of basis images increases, which actually leads to an ill-conditioned matrix with inaccurate separation result. The separation result without a high frequency pattern is shown in Fig. 6 (b).

In conclusion, we have found that there is some extent of independence between direct and global bases from our dataset. However, the degree of independence is not enough to separate each of them.

3.4 Independence of Linear Equation with High frequency Illumination

As described in the previous section, direct and global components cannot be separated effectively using just a basis image. The independence of columns in the matrix $M = \begin{bmatrix} LD & G \end{bmatrix}$ is vital for an invertible capability of a linear system. In this section, we analyze projected pattern L which modulates just the direct components from our basis and discuss the importance of selecting an appropriate pattern L, to create independent columns of M.

3.4.1 Use of High Spatial frequency Illumination

The desirable properties of an illumination pattern have to meet two criteria. Firstly, spatial frequency of an illumination pattern must be high enough to maintain constant properties of the global component as described in Section 3.1. Secondly, the illumination pattern must be capable of differentiating pattern-modulated bases LD of a direct component and pattern-independence bases of a global component. In the following subsection, we analyze the effect of the high spatial frequency checker pattern and summarize our viewpoint in the frequency domain analysis and the linear algebra analysis.

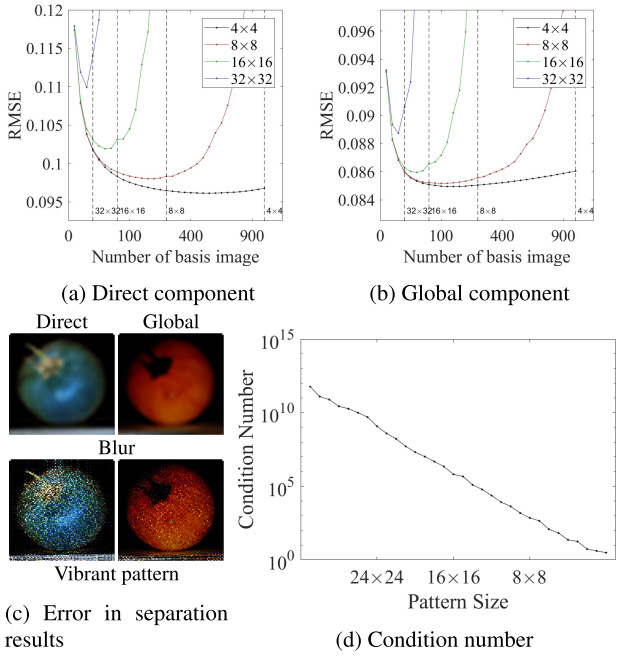
3.4.2 Frequency Analysis

High spatial frequency illumination is widely used to separate direct and global components [18] in a spatial domain. However, the effect in the high frequency domain has not been investigated.

As shown in **Fig. 7**, we use the fast separation method [18] to separate direct component i_d and global component i_g. The frequencies of the two components are densely located in low frequency. Similar result is obtained from the image in uniform illumination $i_d + i_g$. Frequency-wise direct and global components are indifferent under uniform illumination and, therefore inseparable. However, after we modulate direct component i_d with high spatial frequency pattern L, the frequency of i_d is shifted into high frequency, while the unmodulated global component is still in the low frequency. It can be concluded that by modulating scene with high spatial frequency illumination, we create differences in frequencies of the direct and global components. We also observe that the similarity in frequency between direct and global components becomes larger (harder to separate) when lowering the frequency of an illumination pattern (i.e., larger pattern size), and causes the shifted frequency from modulated bases Li_d to be closer to the frequency of unmodulated global bases. In other words, the quality of separation decreases when using lower frequency illumination.

3.4.3 Linear Algebra Analysis

We measure the independence of our proposed linear system in terms of condition number. A lower condition number indicates a higher degree of independence and provide a highly accurate result. We experiment by creating matrix M with varied size of



(a) Direct component (b) Global component

(c) Error in separation results (d) Condition number

Fig. 8 Analysis on the spatial frequency of an illumination pattern and the number of basis. (a) and (b) Comparison of RMSE between different spatial frequency of an illumination pattern when changing the number of basis. *The vertical dotted lines* indicate frequency of basis that equals the frequency of the illumination pattern. (c) Two types of error: (Top) - Blur due to lack of bases, (Bottom) - Vibration due to overfitting. (d) Condition number of the linear equation under a different frequency illumination pattern in logarithm scale; an illumination pattern size varies from 32×32 to 1×1.

a checker illumination pattern from the 128×128 pixels of first 256 of Fourier basis. **Figure 8** (d) shows that the condition number decreases exponentially as the size of a checker illumination pattern reduces, indicating a well-posed linear system with high spatial frequency illumination.

In summary, both experiments demonstrate that by applying a high spatial frequency pattern, our proposed linear system can be solved, and direct and global components can be differentiated.

3.5 Connection between Optimal Number of Basis and Spatial Frequency of Illumination

Furthermore, we investigate the connection between the basis representation and the frequency of an illumination pattern. From Section 3.4, we find that the effectiveness of our linear equation depends on the frequency of an illumination pattern. In practice, the projection of a high frequency pattern is constrained by the specification of equipment. Such limitation creates misalignment between color channels of a projector and false color artifact due to the abrupt change in intensity in a high frequency pattern in the demosaicing process. As a result, it causes lower quality in the captured image. In this part, we propose a methodology to find an optimal number of bases in different frequencies of the illumination pattern, aiming to achieve the highest separation accuracy. We measure Root Mean Square Error (RMSE) on direct and global components between results from the fast separation method with multiple images and our proposed method. The input image of our method is simulated by using Eq. (4) from the actual results. The illumination pattern with checker pattern size (size of each checker block) varied from 4×4, 8×8, 16×16 and 32×32. 1,400 segments of size 128×128 are randomly selected from the test set and used in this simulation. Each image in the test set is captured in the same setup to ensure the consistency of frequency in direct and global components. The analysis is as follows.

Firstly, we consider the optimal basis of uniformly distributed Fourier basis. The relation between separation accuracy and a number of Fourier basis in different spatial frequency illumination is shown in Fig. 8 (a) and 8 (b). The result indicates two critical aspects. The result confirms our hypothesis in Section 3.4 that separation error decreases when the spatial frequency of illumination (smaller pattern size) increases under any number of basis. We have found that by increasing number of basis, the separation accuracy improves until the number of basis image passes through a certain point. The result confirms that the optimal number of basis exists. Two major properties are discovered at such optimal point: the number of optimal basis image increases as the spatial frequency of illumination increases, and the frequency of basis at optimal number is always lower than the frequency of spatial frequency of the illumination pattern. In conclusion, the quality of separation is limited by the spatial frequency of the illumination.

Secondly, we investigate the sources of error from the experiment. These errors result in two contrasting outcomes, blur and vibrant artifact, as shown in Fig. 8 (c). Blur occurs when the basis image is insufficient to represent the source. In contrast, vibrant artifact occurs when basis with higher frequency than spatial frequency of illumination is used. The noise uniformly occurs in the block of the checker pattern which indicates the mixing of direct and global components in the high frequency basis.

Lastly, we compare and evaluate the optimal number of basis and its robustness by and between the use of Fourier, PCA raw and PCA direct-global bases (**Fig. 9**). Basis with a lower number indicates better representative of the component. From the experiment, we find the similarity of the optimal number for every basis image. However, error tendency indicates differences in the robustness of each basis representation. We measure the growth rate of RMSE after an optimal point and illustrate the result in the following order from high to low as follows: Fourier basis, PCA raw basis, and PCA direct-global basis. The direct-global basis from our dataset has the highest robustness to compensate errors and provides more stable separation accuracy in the actual environment where an optimal number of bases is uncertain.

4. Performance Enhancement of the Solving System

The direct implementation of the proposed method requires a large memory which is impractical in a current computer system. The total size of matrix M is a multiplication of a number of pixel and a number of the optimal basis. Both factors are linearly increasing when the image is larger and consume an excessive amount of memory.

Despite of this compactness and simplicity, a naive implementation of the proposed linear system for the whole image is impractical due to excessive memory usage. The length of each basis vector equals to the number of the image pixels, and the

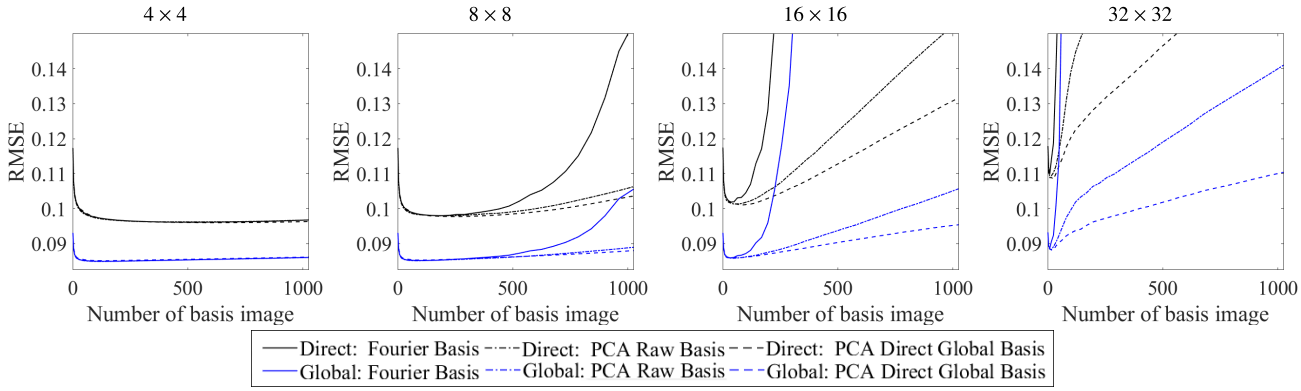


Fig. 9 RMSE comparison on Fourier and PCA bases varied by the frequency of the illumination pattern and the number of bases.

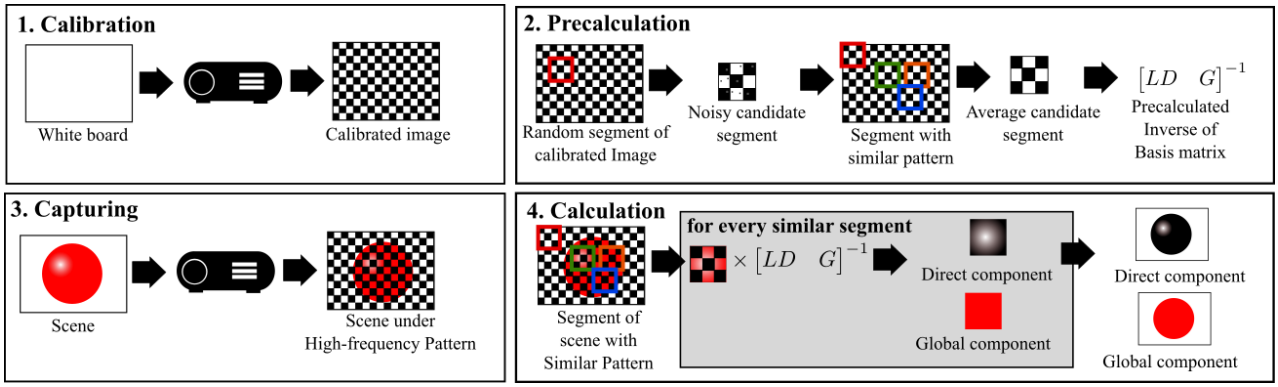


Fig. 10 Overall process of direct and global separation with precomputation for performance enhancement.

number of optimal basis vector increases depending on the frequency of pattern and the image size (larger image size required larger number of basis).

The most straightforward approach to solve such memory issue is to divide an image into small segments and solve each segment separately. The memory usage will be limited to the size of a single segment. However, continuity around the transition of two segments has been left out, which usually creates an artifact line between the border of the two segments. To reduce this artifact, we propose the use of the average between overlapped segments instead of using an exact grid. As a consequence, the computation time will become substantially longer, as the number of segments increases.

To reduce computation time, we propose a pre-computation process. Equation (4) clearly shows that coefficient α and β can be solved by multiplication of the image $\boldsymbol{i}$ with pseudo-inverse of the matrix $M = \begin{bmatrix} LD & G \end{bmatrix}$. The pseudo-inverse matrix M^{-1} depends on light L and does not depend on the scene; thus it can be precomputed. Moreover, considering checker pattern as the high frequency pattern, repetitiveness between segments of the pattern can be used along the image. Such calculation of the inverse matrix is done only once in a small segment which reduces computation time for the whole separation process.

In an actual experiment, the intensity of captured projection illumination is slightly different from ideal projection due to misalignment in the setup, misalignment between projector color channels and noise from a camera. Calibration is required to ensure similarity of pattern in each segment. We calibrate an exact projected illumination for every pixel by capturing a whiteboard in high frequency illumination. Initial segments are randomly selected, then brute-force scanned through the calibrated pattern to find a set of candidate segments containing a similar pattern. We compute an average pattern from candidate segment to reduce noise, then precompute matrix M^{-1}. We apply our separation technique to every candidate segment and average the results to compose the full resolution of direct and global components image. By applying this technique, we substantially reduce the computation time. This acceleration in turn allows us to increase the number of segments, which results in better separation quality. The overall process of our separation technique is shown in **Fig. 10**.

5. Experimental Results

In this section, we conducted experiments to assess the quality of separation and its computational performance. Firstly, we compared the quality of separation between our novel technique and the conventional fast separation technique by measuring the separation error. Then we evaluate the quality of our method on the performance of the dynamic scene. Secondly, we apply our proposed enhancement technique to enable the implementation in a real environment and evaluate its performance.

We set up a real experiment using an off-the-shelf camera and a projector in a coaxial system. We implemented the projection center calibration technique [3] to ensure the alignment. Our coaxial system consisted of 1,920 × 1,080 projector and 1,600 × 1,080 camera. Such alignment is crucial to obtain pixel

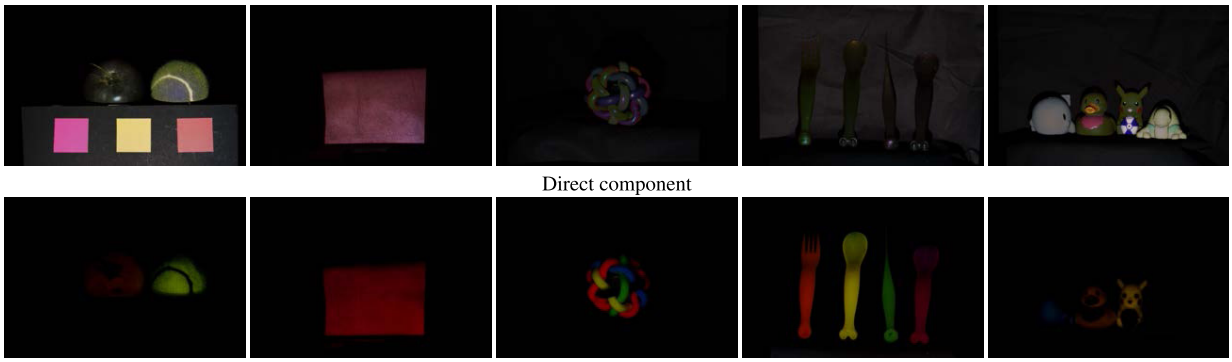


Fig. 11 Samples test image of direct and global components used in our experiment. Each image is captured by using the multiple image technique to create ground-truth data.

Table 1 Average RMSE between baseline and simulated results from each methods. Our method is simulated under $N \times N$ checker patterns. Fast separation single image method is simulated under $N \times 1$ horizontal stripe patterns.

	Direct component				Global component			
Method	$N = 4$	$N = 8$	$N = 16$	$N = 32$	$N = 4$	$N = 8$	$N = 16$	$N = 32$
Fast separation (Single) [18]	0.1955	0.3282	0.4061	0.4897	0.0976	0.1703	0.1950	0.2242
Our (Fourier basis)	**0.0950**	0.0967	0.0996	0.1050	0.0800	0.0801	0.0808	0.0828
Our (PCA raw basis)	0.0952	0.0965	0.0988	0.1030	0.0800	0.0801	0.0805	0.0816
Our (PCA direct-global basis)	0.0951	**0.0962**	**0.0981**	**0.1010**	**0.0800**	**0.0801**	**0.0802**	**0.0804**

correspondence between camera and projector regardless of the geometry of the scene.

5.1 Quality of Separation

We conducted two experiments: a simulated experiment to evaluate the separation accuracy on a perfect projection varied by spatial frequency of the illumination pattern, and a real experiment to evaluate the quality of our technique in a real environment under challenging dynamic scenarios. The results were discussed in the following subsections.

5.1.1 Simulated images

Firstly, we evaluated the separation accuracy on simulated images. The ground-truth of the simulation was obtained by using multiple images method [18]. The example of the test set is shown in **Fig. 11**. The input of a single image separation method was synthesized from ground-truth data assuming that the global component was constant under a different spatial frequency illumination pattern. Our separation technique and the single image fast separation were performed on these synthesized data. The checkerboard pattern was used as an illumination pattern for our technique varied by pattern size from 4×4, 8×8, 16×16 and 32×32 pixels. The number of basis of our method was selected from the optimal point with the lowest error. As for the single image variant of the fast separation technique, we used vertical stripe pattern with pattern width varied from 4, 8, 16 and 32 pixels and prompting window of $2 \times$ pattern size + 1 pixels.

The average RMSE of direct and global component separation was shown in **Table 1** and image-wise RMSE comparison was shown in **Fig. 13**. From these results, we have observed the following aspects.

Firstly, spatial frequency of an illumination pattern affects separation quality. In every method, RMSE in a high spatial frequency illumination pattern was always lower than the result of a lower frequency. The result confirmed our analysis of condition number in our linear system.

Secondly, the accuracy of our technique outperforms the single image fast separation method. With the same pattern frequency, RMSE of our direct and global components was two times lower than the single image fast separation technique. The image-wise plot in Fig. 13 (a) also indicated that our technique provided better separation quality in most of the test images.

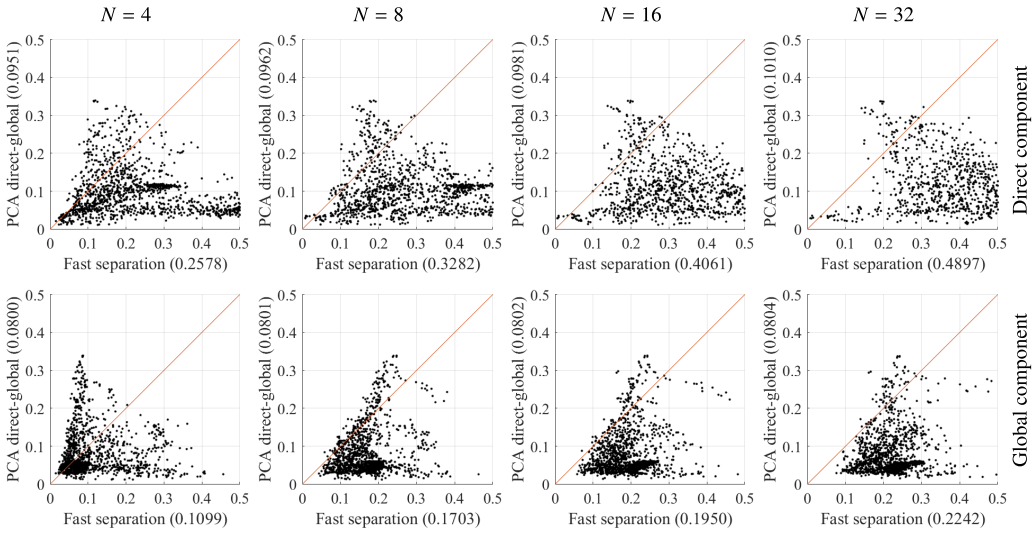
Lastly, in most cases, the accuracy of direct-global basis is better than other bases with identical direct and global components. The improvement indicated that bases separately learned from real separation were the appropriate representatives of the direct and global components. This disclosed the statistical differences between the two components.

5.1.2 Real images

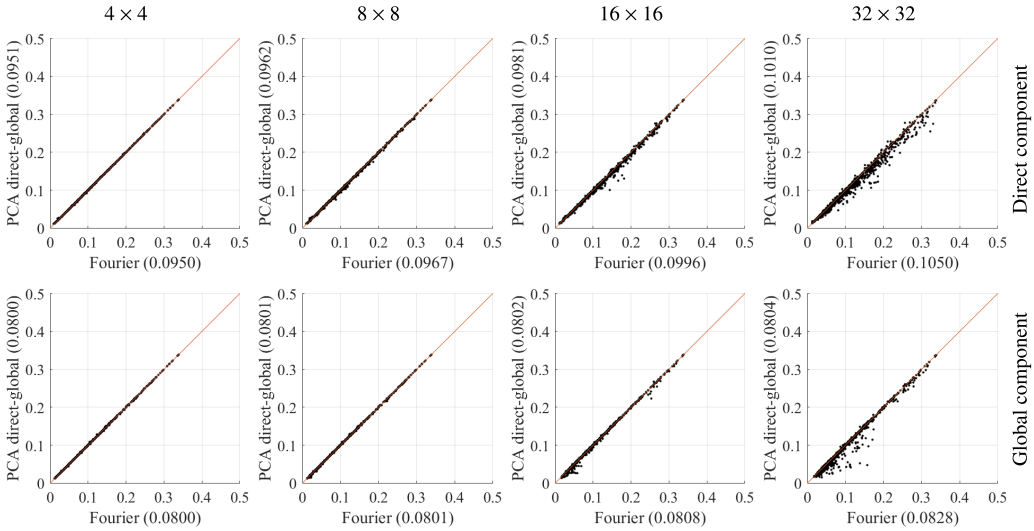
In an actual experiment, we used an 8×8 checker illumination pattern for our technique and multiple images methods, and an 8-pixel width horizontal stripe pattern for single image fast separation method. We captured the performance on both a still image and a video sequence of moving objects. The qualitative comparison of the appearance of the direct and global components from real images between the two methods was shown in **Fig. 14**. The quantitative analysis was discussed as follows.

Firstly, we compared the separation result on a still object using our proposed method and a single image fast separation method. Moreover, its separation accuracy was found to outperform the conventional single image method, by for example, observing the transition between objects in the area of the mouth and eyes of the doll in Fig. 14 (b).

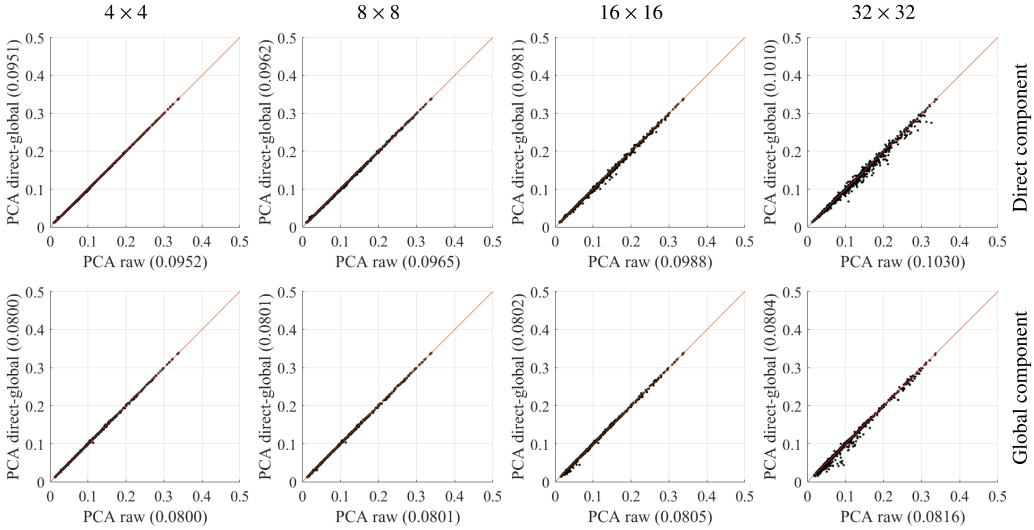
Secondly, we compared the separation result on moving objects using our proposed method in a single frame and the multiple images fast separation method. The result on the image of a human hand as shown Fig. 14 (c) illustrated a smooth result from



(a) Comparison of RMSE for the fast separation single image method and our method under different spatial frequencies of illumination patterns ($N \times N$ checker pattern for our Fourier basis method and $N \times 1$ horizontal stripe pattern for conventional single image method).



(b) Comparison of RMSE for Fourier and PCA direct-global bases using different spatial frequencies of illumination patterns.



(c) Comparison of RMSE for PCA raw and PCA direct-global bases using different spatial frequencies of illumination patterns.

Fig. 13 Experimental evaluation on simulated images.

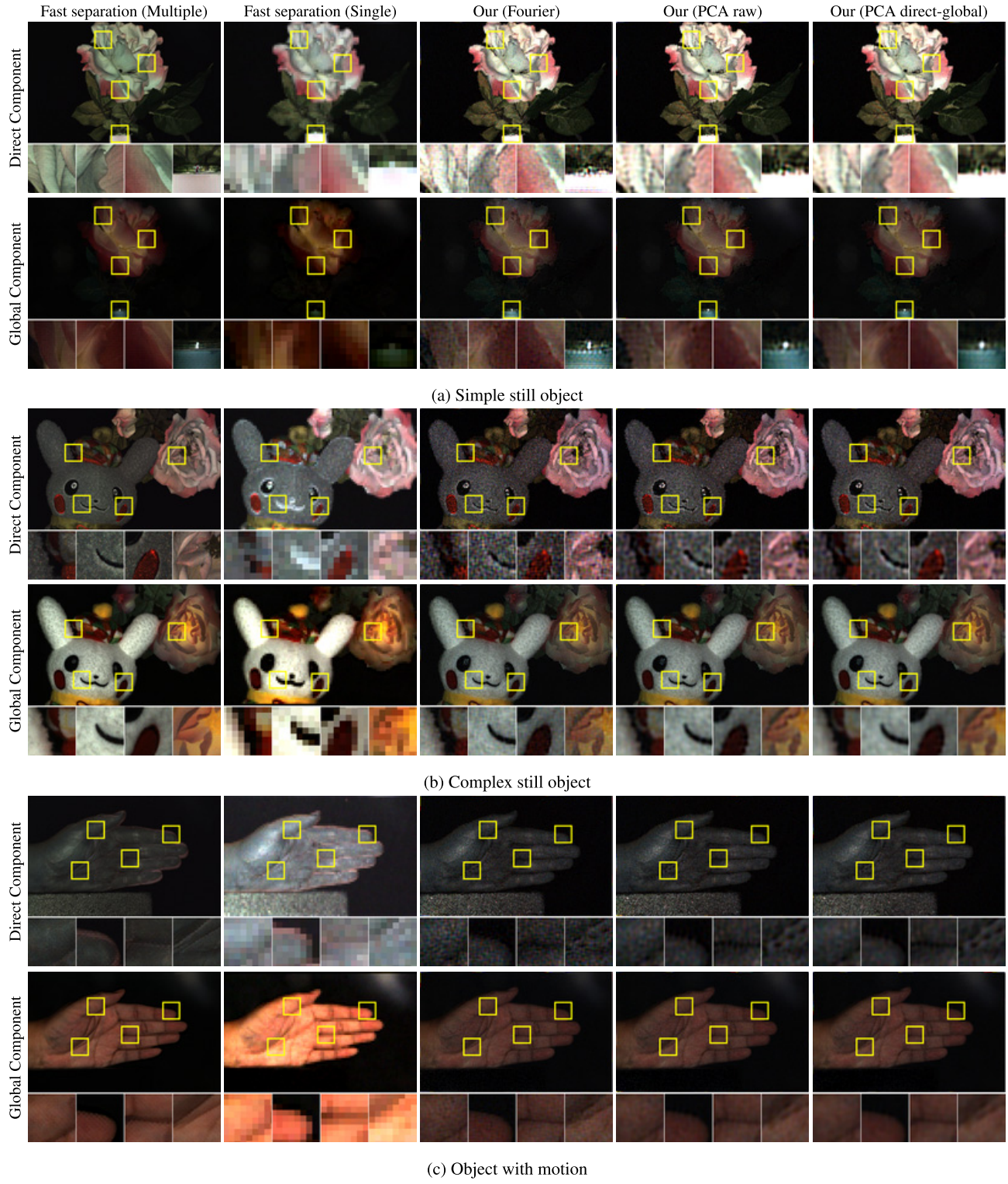


Fig. 14 Experimental result comparison between the fast separation method (multiple images or single image) and our method (Fourier, PCA raw, PCA direct-global basis).

our method while an artifact border around fingertips caused by jiggling of hand was found in the conventional multiple images method. Such result confirmed the capability of our proposed method in handling scene with a moving object.

Thirdly, we evaluated the effect of various bases. The separation results from our direct-global bases showed less noise than other bases with no prior knowledge of direct and global components. This result is consistent with the calculated RMSE in our simulation.

Lastly, we performed our technique on the video sequences which demonstrated the capability of single frame separation and its compatibility with a fully dynamic scene. The comparison results in a video sequence between our method and fast separation are shown in **Fig. 15**.

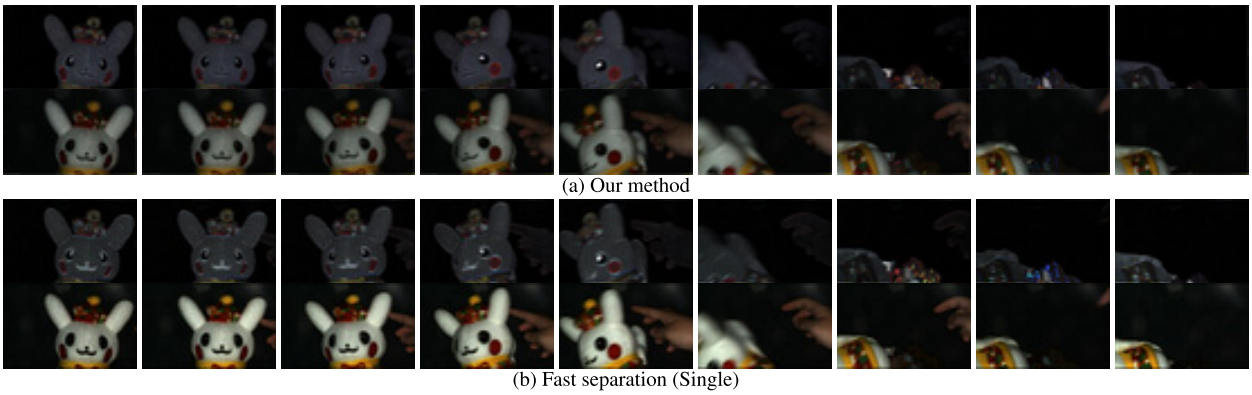
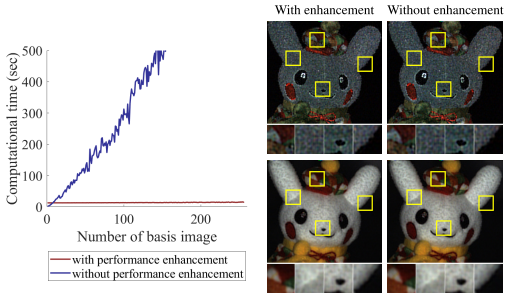


Fig. 15 Separation results from video. (a) Our method - direct and global. (b) Fast separation single image - direct and global.



(a) Computational time (b) Separation result

Fig. 12 Performance enhancement of our separation technique without sacrificing the quality of separation. (a) Comparison of computation time with/without enhancement. (b) Comparison of separation results with/without performance enhancement.

Table 2 Computation time (sec) between implementation of our method with/without performance enhancement technique with different basis numbers N.

Step	N=4	N=64	N=256
Precalculation	12.1182	12.3634	14.1599
Separation	0.8572	0.9143	1.9609
With enhancement	12.9755	**13.2777**	**16.1207**
Without enhancement	**5.1041**	195.3822	1491.5955

5.2 Performance Enhancement

We evaluated the quality and performance of our enhancement technique under two critical criteria: the acceleration of computation time and the stability of separation quality. We evaluated the computation time on 400×400 images from the previous experiment. Our separation system was implemented by using Matlab R2018a on Ubuntu 14.04LTS with an Intel Xeon(R) CPU E5-1650 v4 @ 3.60 GHz. To improve computation time, we performed matrix-matrix multiplication to compute multiple segments simultaneously.

We measured computation time including the precomputation time for the entire process proposed in Section 4. **Figure 12** (a) illustrates that the overhead of the precomputation process was quickly compensated when the number of bases increased. The results in **Table 2** also confirmed that our performance enhancement technique was significantly advantageous. Noted that reduction in computation time from performance enhancement technique is substantial for larger image with an increased number of segments.

Moreover, Fig. 12 (b) showed that by applying performance enhancement technique to the approximate pattern, we obtained almost identical separation result. Such result was predictable due to the similarity of the exact pattern for each segment. Noted that if the pattern was distorted due to the geometry of an object, it indicated the misalignment between camera and projector system and caused failure in our system.

6. Conclusion

In this article, we propose a novel technique for direct and global components separation from a single image with full spatial resolution. We introduce a simple linear basis model and a light transport dataset to learn the PCA basis representation. We explore the dependency between basis images and illustrate how the high spatial frequency illumination pattern contributes to solving the ambiguity between these bases as well as solving the linear solution. We also suggest a performance enhancement technique to reduce computer memory usage and computation time to a practical level in the real environment.

Our system requires the use of only a coaxial standard off-the-shelf camera and a projector. The study shows that our data-driven bases are better in representing direct and global components. The experiment on both simulated and real images confirm that our method delivers better separation quality and creates lower error compared to other approaches. Moreover, our method was proven capable of accurately separating scene with moving objects.

A limitation on our technique is the noise sensitivity. In our model, a perfect model of both direct and global component has been assumed without the representation of the noise term. The input image with excessive noise results in high RMSE in both direct and global components.

In this study, we have assumed coaxial alignment between a camera and a projector in the system, which would require much time and trouble for system alignment and calibration. We are exploring the possibility of using data-driven basis representation for accurate separation in a real environment with no prior knowledge of the correspondence between the camera and the projection pattern.

Acknowledgments This work was supported in part by JSPS KAKENHI Grant Number JP15H05918.

References

[1] Achar, S. and Narasimhan, S.G.: Multi Focus Structured Light for Recovering Scene Shape and Global Illumination, *Computer Vision – ECCV 2014*, Fleet, D., Pajdla, T., Schiele, B. and Tuytelaars, T. (Eds.), Cham, Springer International Publishing, pp.205–219 (2014).

[2] Achar, S., Nuske, S.T. and Narasimhan, S.G.: Compensating for motion during direct-global separation, *Proc. IEEE International Conference on Computer Vision*, pp.1481–1488, IEEE (2013) (online), available from ⟨http://ieeexplore.ieee.org/document/6751294/⟩.

[3] Amano, T.: Projection center calibration for a co-located projector camera system, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, pp.449–454 (2014).

[4] Bai, J., Chandraker, M., Ng, T.T. and Ramamoorthi, R.: A dual theory of inverse and forward light transport, *Lecture Notes in Computer Science* (*including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics*), Vol.6312 LNCS, No.PART 2, Springer Berlin Heidelberg, pp.294–307 (2010) (online), available from ⟨http://graphics.cs.berkeley.edu/papers/Bai-ADT-2010-09/ http://link.springer.com/10.1007/978-3-642-15552-9_22⟩.

[5] Carroll, R., Ramamoorthi, R. and Agrawala, M.: Illumination decomposition for material recoloring with consistent interreflections, *ACM Trans. Graphics*, Vol.30, No.4, p.1 (2011).

[6] Chen, T., Seidel, H.P. and Lensch, H.P.: Modulated phase-shifting for 3D scanning, *26th IEEE Conference on Computer Vision and Pattern Recognition, CVPR*, pp.1–8 (2008).

[7] Cossairt, O., Nayar, S. and Ramamoorthi, R.: Light field transfer, *ACM Trans. Graphics*, Vol.27, No.3, p.1 (online), DOI: 10.1145/1360612.1360656 (2008).

[8] Ghosh, A., Hawkins, T., Peers, P., Frederiksen, S. and Debevec, P.: Practical modeling and acquisition of layered facial reflectance, *ACM Transactions on Graphics*, Vol.27, No.5, p.1 (2008) (online), available from ⟨http://portal.acm.org/citation.cfm?doid=1409060.1409092⟩.

[9] Gu, J., Kobayashi, T., Gupta, M. and Nayar, S.K.: Multiplexed illumination for scene recovery in the presence of global illumination, *Proc. IEEE International Conference on Computer Vision*, pp.691–698 (2011).

[10] Gupta, M., Narasimhan, S.G. and Schechner, Y.Y.: On controlling light transport in poor visibility environments, *26th IEEE Conference on Computer Vision and Pattern Recognition, CVPR*, pp.1–8 (2008).

[11] Gupta, M., Nayar, S.K., Hullin, M.B. and Martin, J.: Phasor Imaging, *ACM Trans. Graphics*, Vol.34, No.5, pp.1–18 (2015) (online), available from ⟨http://dl.acm.org/citation.cfm?doid=2843519.2735702⟩.

[12] Gupta, M., Tian, Y., Narasimhan, S.G. and Zhang, L.: A combined theory of defocused illumination and global light transport, *International Journal of Computer Vision*, Vol.98, No.2, pp.146–167 (2012) (online), available from ⟨http://dx.doi.org/10.1007/s11263-011-0500-9 http://link.springer.com/10.1007/s11263-011-0500-9⟩.

[13] Liu, Y., Qin, X., Xu, S., Nakamae, E. and Peng, Q.: Light source estimation of outdoor scenes for mixed reality, *Visual Computer*, Vol.25, No.5-7, pp.637–646 (2009) (online), available from ⟨http://dx.doi.org/10.1007/s00371-009-0342-4⟩.

[14] Mukaigawa, Y., Suzuki, K. and Yagi, Y.: Analysis of Subsurface Scattering Based on Dipole Approximation, *Information and Media Technologies*, Vol.4, No.4, pp.951–961 (2009).

[15] Munoz, A., Echevarria, J.I., Seron, F.J., Lopez-Moreno, J., Glencross, M. and Gutierrez, D.: BSSRDF estimation from single images, *Computer Graphics Forum*, Vol.30, No.2, pp.455–464 (2011) (online), available from ⟨http://dx.doi.org/10.1111/j.1467-8659.2011.01873.x⟩.

[16] Narasimhan, S.G., Koppal, S.J. and Yamazaki, S.: Temporal dithering of illumination for fast active vision, *Lecture Notes in Computer Science* (*including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics*), Vol.5305 LNCS, No.PART 4, pp.830–844 (2008).

[17] Nayar, S.K. and Gupta, M.: Diffuse structured light, *2012 IEEE International Conference on Computational Photography* (*ICCP*), IEEE (2012).

[18] Nayar, S.K., Krishnan, G., Grossberg, M.D. and Raskar, R.: Fast separation of direct and global components of a scene using high frequency illumination, *ACM Transactions on Graphics*, Vol.25, No.3, p.935 (2006) (online), available from ⟨http://portal.acm.org/citation.cfm?doid=1141911.1141977⟩.

[19] O'Toole, M., Achar, S., Narasimhan, S.G. and Kutulakos, K.N.: Homogeneous codes for energy-efficient illumination and imaging, *ACM Trans. Graphics*, Vol.34, No.4, pp.35:1–35:13 (2015) (online), available from ⟨http://dl.acm.org/citation.cfm?id=2809654.2766897⟩.

[20] O'Toole, M., Heide, F., Xiao, L., Hullin, M.B., Heidrich, W. and Kutulakos, K.N.: Temporal frequency probing for 5D transient analysis of global light transport, *ACM Trans. Graphics*, Vol.33, No.4, pp.1–11 (2014) (online), available from ⟨http://dl.acm.org/citation.cfm?id=2601097.2601103⟩.

[21] OToole, M., Mather, J. and Kutulakos, K.N.: 3D Shape and Indirect Appearance by Structured Light Transport, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.38, No.7, pp.1298–1312 (2016).

[22] O'Toole, M., Raskar, R. and Kutulakos, K.N.: Primal-dual coding to probe light transport, *ACM Trans. Graphics*, Vol.31, No.4, pp.1–11 (2012) (online), available from ⟨http://dl.acm.org/citation.cfm?doid=2185520.2185535⟩.

[23] Reddy, D., Ramamoorthi, R. and Curless, B.: Frequency-space decomposition and acquisition of light transport under spatially varying illumination, *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Vol.7577 LNCS, No.PART 6, pp.596–610, Springer Berlin Heidelberg (2012) (online), available from ⟨http://graphics.berkeley.edu/papers/Reddy-FSD-2012-10/ http://link.springer.com/10.1007/978-3-642-33783-3_43⟩.

[24] Satat, G., Barsi, C. and Raskar, R.: Skin perfusion photography, *2014 IEEE International Conference on Computational Photography, ICCP 2014*, pp.1–8, IEEE (2014) (online), available from ⟨http://ieeexplore.ieee.org/document/6831804/⟩.

[25] Shimano, M., Bise, R., Zheng, Y. and Sato, I.: Separation of Transmitted Light and Scattering Components in Transmitted Microscopy, *Medical Image Computing and Computer-Assisted Intervention − MICCAI 2017*, Descoteaux, M., Maier-Hein, L., Franz, A., Jannin, P., Collins, D.L. and Duchesne, S. (Eds.), Cham, Springer International Publishing, pp.12–20 (2017).

[26] Subpa-asa, A., Fu, Y., Zheng, Y., Amano, T. and Sato, I.: Direct and Global Component Separation from a Single Image Using Basis Representation, *Computer Vision – ACCV 2016*, pp.99–114, Springer International Publishing (2017).

[27] Subpa-asa, A., Zheng, Y., Ono, N. and Sato, I.: Light transport component decomposition using multi-frequency illumination, *2017 IEEE International Conference on Image Processing* (*ICIP*), pp.3595–3599, IEEE (2017) (online), available from ⟨http://ieeexplore.ieee.org/document/8296952/⟩.

[28] Talvala, E. and Adams, A.: Veiling glare in high dynamic range imaging, *ACM Trans. Graphics* (*TOG*), *SIGGRAPH '07*, Vol.26, No.3, pp.1–10 (2007) (online), available from ⟨http://dl.acm.org/citation.cfm?id=1276424⟩.

[29] Tanaka, K., Mukaigawa, Y., Kubo, H., Matsushita, Y. and Yagi, Y.: Recovering Inner Slices of Layered Translucent Objects by Multi-Frequency Illumination, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.39, No.4, pp.746–757 (online), DOI: 10.1109/TPAMI.2016.2631625 (2017).

[30] Tariq, S., Gardner, A., Llamas, I., Jones, A., Debevec, P. and Turk, G.: Efficient Estimation of Spatially Varying Subsurface Scattering Parameters, ICT Technical Report ICT TR 01 2006, University of Southern California Institute for Creative Technologies (2006).

[31] Wu, D.B., Velten, A.D., O'Toole, M.E., Masia, B.F., Agrawal, A., Dai, Q. and Raskar, R.: Decomposing Global Light Transport Using Time of Flight Imaging, *International Journal of Computer Vision*, pp.1–16 (2013) (online), available from ⟨http://www.scopus.com/inward/record.url?eid=2-s2.0-84886416962&partnerID=40&md5=ac70cbaf533d2340e3fbaf0f6744711f⟩.

[32] Yan, L.-Q., Sun, W., Jensen, H.W. and Ramamoorthi, R.: A BSSRDF model for efficient rendering of fur with global illumination, *ACM Trans. Graphics*, Vol.36, No.6, pp.1–13 (2017).

[33] Zinke, A., Rump, M., Lay, T., Weber, A., Andriyenko, A. and Klein, R.: A practical approach for photometric acquisition of hair color, *ACM Trans. Graphics*, Vol.28, No.5, p.1 (2009) (online), available from ⟨http://doi.acm.org/10.1145/1618452.1618511⟩.

Art Subpa-asa is a Ph.D. candidate from Tokyo Institute of Technology in Interdisciplinary School of Science and Engineering. He received B.E. and M.E. in Computer Engineering from Chulalongkorn University, Thailand. His research interests include computational photography of global illumination and acquisition system.

Ying Fu received her B.S. degree in Electronic Engineering from Xidian University in 2009, the M.S. degree in Automation from Tsinghua University in 2012, and the Ph.D. degree in Information Science and Technology from the University of Tokyo in 2015. She is currently a project research associate in the Institute of Industrial Science, the University of Tokyo. Her research interests include physics-based vision, image processing, and computational photography

Yinqiang Zheng received his Bachelor degree from the Department of Automation, Tianjin University, Tianjin, China, in 2006, Master degree of engineering from Shanghai Jiao Tong University, Shanghai, China, in 2009, and Doctoral degree of engineering from the Department of Mechanical and Control Engineering, Tokyo Institute of Technology, Tokyo, Japan in 2013. He is currently an assistant professor in the National Institute of Informatics, Japan. His research interests include image processing, computer vision, and mathematical optimization.

Toshiyuki Amano received his B.E., M.E., and Ph.D. degrees in engineering science from Osaka University in 1995, 1997, and 2000, respectively. He is a Professor at the Faculty of Systems Engineering, Wakayama University. His research interests include spatial augmented reality, light field sensing and display, and installation art applications.

Imari Sato received the B.S. degree in policy management from Keio University in 1994. After studying at the Robotics Institute of Carnegie Mellon University as a visiting scholar, she received her M.S. and Ph.D. degrees in interdisciplinary Information Studies from the University of Tokyo in 2002 and 2005, respectively. In 2005, she joined the National Institute of Informatics, where she is currently a professor. Her primary research interests are in the fields of computer vision (physics-based vision, image-based modeling) and Computer Graphics (image-based rendering, augmented reality). She has received various research awards, including IPSJ Nagao Special Researcher award (2010), The Young Scientists' Prize from The Commendation for Science and Technology by the Minister of Education, Culture, Sports, Science and Technology (2009), and Microsoft Research Japan New Faculty award (2011).