This CVPR paper is the Open Access version, provided by the Computer Vision Foundation.
Except for this watermark, it is identical to the accepted version;
the final published version of the proceedings is available on IEEE Xplore.

Light Source Separation and Intrinsic Image Decomposition under AC Illumination

Yusaku Yoshida, Ryo Kawahara, and Takahiro Okabe
Department of Artificial Intelligence, Kyushu Institute of Technology
680-4 Kawazu, Iizuka, Fukuoka 820-8502, Japan
{rkawahara, okabe}@ai.kyutech.ac.jp

Abstract

Artificial light sources are often powered by an electric grid, and then their intensities rapidly oscillate in response to the grid's alternating current (AC). Interestingly, the flickers of scene radiance values due to AC illumination are useful for extracting rich information on a scene of interest. In this paper, we show that the flickers due to AC illumination is useful for intrinsic image decomposition (IID). Our proposed method conducts the light source separation (LSS) followed by the IID under AC illumination. In particular, we reveal the ambiguity in the blind LSS via matrix factorization and the ambiguity in the IID assuming the diffuse reflection model, and then show why and how those ambiguities can be resolved via a physics-based approach. We experimentally confirmed that our method can recover the colors of the light sources, the diffuse reflectance values, and the diffuse and specular intensities (shadings) under each of the light sources, and that the IID under AC illumination is effective for application to auto white balancing.

1. Introduction

Artificial light sources in our surroundings are often powered by an electric grid, and therefore their intensities rapidly oscillate in response to the grid's alternating current (AC). Such intensity oscillations cause flickers in the radiance values of a scene illuminated by artificial light sources. The flickers are usually too fast to notice with our naked eyes, but can be captured by using cameras with short exposure time settings [32]. It is known that the flickers could make auto white balance unnatural [15].

Interestingly, the flickers are useful for extracting rich information on a scene of interest. Sheinin *et al*. [28] propose a method for light source separation (LSS) under AC illumination. Their method decomposes an image sequence of a scene illuminated by multiple AC light sources into the basis images of the scene, each of which is illuminated by only one of the light sources, and the temporal intensity profiles of the light sources. They make use of their self-built coded-exposure camera synchronized to AC and the dataset of temporal intensity profiles of various light sources, and then achieve LSS even for dark scenes such as a city-scale scene at night.

In this paper, we show that the flickers due to AC illumination is useful also for intrinsic image decomposition (IID). Originally, IID recovers the shading and reflectance images of a scene of interest from a single input image on the basis of the Retinex theory [2, 19]. Those intrinsic properties of a scene is useful for computer vision applications such as image segmentation [6], object recognition [25], and shape from shading [14].

Our proposed method assumes a scene illuminated by multiple AC light sources, and recovers the intrinsic properties of the scene and the light sources from an image sequence captured by using a consumer high speed camera. In contrast to the conventional methods for IID, our method assumes the dichromatic reflection model [27], and then recovers the intrinsic properties more than the reflectance and shading images: the colors of the light sources, the diffuse reflectance values, and the diffuse and specular intensities (shadings) under each of the light sources.

Specifically, our proposed method conducts the blind LSS via matrix factorization followed by the IID assuming the dichromatic reflection model. In particular, we reveal the ambiguity in the blind LSS under AC illumination via matrix factorization [26], and then resolve the ambiguity by integrating the LSS and the IID assuming the diffuse reflection model. Furthermore, we reveal the ambiguity in the IID assuming the diffuse reflection model under AC illumination, and then resolve the ambiguity on the basis of the dichromatic reflection model by taking specular highlights into consideration [1].

[1] It is analogous to uncalibrated photometric stereo; the GBR ambiguity [3] can be resolved from specularity [8]

To show the effectiveness of our proposed method, we conducted a number of experiments using both synthetic and real images. We confirmed that our method works well, *i.e.* can resolve the ambiguities in the LSS and the IID on real images as well as synthetic images. In addition, we show that the IID under AC illumination is effective for application to auto white balancing.

The main contributions of this study are threefold. First, we tackle a novel problem of the IID under AC illumination. We conduct the blind LSS via matrix factorization followed by the IID assuming the dichromatic reflection model, and show that the flickers due to AC illumination are useful not only for LSS but also for IID. Second, we reveal the ambiguity in the blind LSS via matrix factorization and the ambiguity in the IID assuming the diffuse reflection model. Then, we show why and how those ambiguities can be resolved via a physics-based approach. Third, we experimentally confirmed that our method can recover the colors of the light sources, the diffuse reflectance values, and the diffuse and specular intensities (shadings) under each of the light sources, and that the IID under AC illumination is effective for application to auto white balancing.

2. Related Work

2.1. Light Source Separation

Light source separation (LSS) is the problem of separating an image/images of a scene illuminated by multiple light sources into the basis images of the scene, each of which is illuminated by only one of the light sources. In this study, we address the LSS under AC illumination.

Sheinin *et al*. [28] find that the flickers due to AC illumination can be used for LSS, and propose a method for decomposing an image sequence of a scene illuminated by multiple AC light sources into the basis images and temporal intensity profiles of the light sources. Although their method has clear advantage that it can be applied even to dark scenes such as a city-scale scene at night, it requires the self-build coded-exposure camera synchronized to AC and the dataset of temporal intensity profiles of various light sources. Later, Sheinin *et al*. [29] achieve the LSS under AC illumination by using a consumer rolling-shutter camera, but still require the dataset of temporal intensity profiles of various light sources.

Oya *et al*. [26] solve the LSS under AC illumination as a problem of blind source separation (BSS). They show that the LSS under AC illumination results in the problem of matrix factorization, and experimentally confirmed that non-negative matrix factorization (NMF) [4] performs better than independent component analysis (ICA) [16]. Their method does not require the self-build camera synchronized to AC nor the dataset of temporal intensity profiles of various light sources. However, as shown in Section 3.1, their method based on matrix factorization has an ambiguity, and therefore its accuracy is limited.

In contrast to the former method [28], our proposed method for the LSS under AC illumination is based on BSS, and therefore it also does not require the self-built camera synchronized to AC nor the dataset of temporal intensity profiles of various light sources. Compared with the latter method [26], our method resolves the ambiguity of the blind LSS via matrix factorization by integrating the LSS and the IID assuming the diffuse reflection model.

2.2. Intrinsic Image Decomposition

Intrinsic image decomposition (IID) [2] is the problem of decomposing an image/images of a scene into the shading and reflectance images of the scene. Please refer Grosse *et al*. [13] and Garces *et al*. [12] to comprehensive survey and experimental comparison.

The IID from a single image is an under-determined problem because it decomposes a single image into two images: the reflectance and shading images. Therefore, various priors are studied in order to solve an under-determined problem of the IID from a single image. Conventionally, the Retinex theory [19] is often used for the IID from a single image. It considers the large/small gradients in an image of a scene are caused by the reflectance/shading of the scene. Later, Funt *et al*. [11] extend the above method from a gray-scale image to a color image. Tappen *et al*. [31] show that a classifier trained to recognize gray-scale patterns is useful as the prior for the IID from a single image.

To make the under-determined problem of IID tractable, we can make use of multiple images of a scene. Weiss [33] proposes the IID from an image sequence of a scene, *i.e.* time-lapse images under varying illumination conditions. He estimates the single reflectance image and the multiple shading images of the scene by using the statistics of natural images. Matsuoka *et al*. [23] use a pair of flash and no-flash images of a scene, and then estimate the reflectance and shading images on the basis of the sparseness of the reflectance image.

In contrast to the above methods, our proposed method exploits the flickers due to AC illumination; we show that IID is tractable just by observing flickers with a short exposure-time camera. More importantly, our method is based on the dichromatic reflection model [27], and estimates not only the diffuse reflectance values (reflectance image) and the diffuse intensity (shading image) under each of the light sources but also the colors of the light sources and the specular intensity (shading image) under each of the light sources. In particular, we reveal that the IID assuming the diffuse reflection model has an ambiguity, and show that the ambiguity can be resolved by taking specular highlights into consideration.

In addition to the IID from a single image or an image

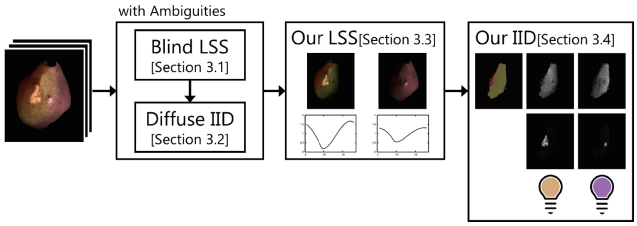

Figure 1. The flowchart of our proposed method for LSS and IID under AC illumination; LSS followed by IID.

sequence under varying illumination conditions, intrinsic video with camera motion [17, 34] and the IID from multiview images [18] are proposed. In this study, we assume a static scene and a static camera, and therefore such camera motion and multiview images are out of our focus.

Recently, deep learning-based methods [5, 9, 20–22, 24, 30, 35] show impressive results also for IID, but they rely on supervised learning and require large datasets in general. In contrast, our proposed method exploits the flickers due to AC light sources, and achieves the LSS and IID from a single image sequence of a scene. The theoretical insights that reveal the ambiguities in the blind LSS and the diffuse IID under AC light sources and show why and how those ambiguities can be resolved are the advantages of our method over the learning-based methods.

3. Proposed Method

In this section, we propose a method for conducting light source separation (LSS) followed by intrinsic image decomposition (IID) from an image sequence taken under AC illumination. Specifically, we reveal the ambiguities in the LSS and IID under AC illumination, and then explain why and how those ambiguities can be resolved. Figure 1 is the flowchart of our proposed method for LSS and IID under AC illumination.

3.1. Blind LSS and Its Ambiguity

According to the superposition principle of light, an image of a scene taken under multiple light sources is represented by a convex combination of the *basis images*, each of which is the image taken under one of the light sources. Specifically, the pixel value i_{pcf} of an input image sequence at the p-the pixel $(p = 1, 2, 3, ..., P)$ of the c-th channel $(c = 1, 2, 3)$ and in the f-th frame $(f = 1, 2, 3, ..., F)$ is represented as

$$i_{pcf} = \sum_{n=1}^{N} b_{pcn} a_{nf}. \tag{1}$$

Here, N is the number of the light sources (and the number of the basis images), b_{pcn} is the pixel value at the p-th pixel of the c-th channel and in the n-th basis image, and a_{nf} is the intensity of the n-th light source in the f-th frame.

We can rewrite eq.(1) in a matrix form as

$$\boldsymbol{I} = \boldsymbol{B}\boldsymbol{A}, \tag{2}$$

where $\boldsymbol{I}$ is the $3P \times F$ matrix consisting of the pixel values of the input image sequence, $\boldsymbol{B}$ is the $3P \times N$ matrix consisting of the pixel values of the N basis images, and $\boldsymbol{A}$ is the $N \times F$ matrix consisting of the intensities of the N light sources. Because those pixel values and light source intensities are all non-negative, the blind LSS results in the problem of non-negative matrix factorization (NMF) [4]; factorizing the matrix $\boldsymbol{I}$ of the input image sequence into the non-negative matrix $\boldsymbol{B}$ of the basis images and the non-negative matrix $\boldsymbol{A}$ of the light source intensities. Specifically, LSS results in the minimization:

$$\min_{\{\boldsymbol{B},\boldsymbol{A}\}} ||\boldsymbol{I} - \boldsymbol{B}\boldsymbol{A}||_{\mathrm{F}}^2 \tag{3}$$

subject to the conditions that all the elements of the matrices $\boldsymbol{B}$ and $\boldsymbol{A}$ are non-negative [26]. Here, $||\ \ ||_{\mathrm{F}}$ stands for the Frobenius norm of a matrix.

Unfortunately, however, the blind LSS via matrix factorization has an ambiguity since the left-hand side in eq.(2) is invariant if we put an $N \times N$ regular matrix $\boldsymbol{X}$ and its inverse $\boldsymbol{X}^{-1}$ such that $\boldsymbol{B}\boldsymbol{X}$ and $\boldsymbol{X}^{-1}\boldsymbol{A}$ are non-negative between $\boldsymbol{B}$ and $\boldsymbol{A}$ as

$$\boldsymbol{I} = \boldsymbol{B}\boldsymbol{X}\boldsymbol{X}^{-1}\boldsymbol{A}. \tag{4}$$

Therefore, the basis images $\boldsymbol{B}^{(e)}$ and light source intensities $\boldsymbol{A}^{(e)}$ estimated via matrix factorization are represented by using the unknown matrix $\boldsymbol{X}$ as $\boldsymbol{B}^{(e)} = \boldsymbol{B}\boldsymbol{X}$ and $\boldsymbol{A}^{(e)} = \boldsymbol{X}^{-1}\boldsymbol{A}$ respectively.

3.2. Diffuse IID and Its Ambiguity

Let us assume that the reflectance properties of a scene of interest obey the diffuse reflection model [2]. Then, the pixel values, *i.e.* the RGB values $\boldsymbol{b}_{pn} = (b_{p1n}, b_{p2n}, b_{p3n})^\top$ at the p-th pixel in the n-th basis image is represented as

$$\begin{aligned} \boldsymbol{b}_{pn} &= d_{pn} \begin{pmatrix} r_{p1} & 0 & 0 \\ 0 & r_{p2} & 0 \\ 0 & 0 & r_{p3} \end{pmatrix} \begin{pmatrix} l_{n1} \\ l_{n2} \\ l_{n3} \end{pmatrix} \\ &= d_{pn} \boldsymbol{R}_p \boldsymbol{l}_n. \end{aligned} \tag{5}$$

Here, d_{pn}, $\boldsymbol{r}_p = (r_{p1}, r_{p2}, r_{p3})^\top$, and $\boldsymbol{l}_n = (l_{n1}, l_{n2}, l_{n3})^\top$ are the intensity of the diffuse reflection, the diffuse reflectance values at the p-th pixel, and the color of the n-th light source respectively. Therefore, the IID assuming the diffuse reflection model is the problem of estimating the

[2] As explained in Appendix A, we can conduct diffuse-specular separation if necessary.

intensity of the diffuse reflection d_{pn} per basis image and pixel, the diffuse reflectance values $\boldsymbol{r}_p$ per pixel, and the color of the light source $\boldsymbol{l}_n$ per basis image. We term the 3×3 matrix $\boldsymbol{R}_p$, whose diagonal elements are the diffuse reflectance values, the diffuse reflectance matrix.

Unfortunately, however, the IID assuming the diffuse reflection model also has an ambiguity in a similar manner to the blind LSS via matrix factorization. This is because the left-hand side in eq.(5) is invariant if we put a 3×3 positive diagonal matrix $\boldsymbol{Y}$ and its inverse $\boldsymbol{Y}^{-1}$ between $\boldsymbol{R}_p$ and $\boldsymbol{l}_n$ as

$$\boldsymbol{b}_{pn} = d_{pn} \boldsymbol{R}_p \boldsymbol{Y} \boldsymbol{Y}^{-1} \boldsymbol{l}_n. \tag{6}$$

Therefore, when we conduct the IID assuming the diffuse reflection model from the given basis images as described below, the estimated diffuse reflectance matrix $\boldsymbol{R}_p^{(e)}$ and the light source color $\boldsymbol{l}_n^{(e)}$ are represented by using the unknown diagonal matrix $\boldsymbol{Y}$ as $\boldsymbol{R}_p^{(e)} = \boldsymbol{R}_p \boldsymbol{Y}$ and $\boldsymbol{l}_n^{(e)} = \boldsymbol{Y}^{-1} \boldsymbol{l}_n$ respectively.

We can analytically estimate the intensity of the diffuse reflection d_{pn} per basis image and pixel, the diffuse reflectance values $\boldsymbol{r}_p$ per pixel, and the color of the light source $\boldsymbol{l}_n$ per basis image up to the ambiguity of an unknown diagonal matrix as follows. First, we take the ambiguity in the light source colors into consideration, and fix the color of the 1st light source as $\boldsymbol{l}_1 = (1/3, 1/3, 1/3)^\top$ without loss of generality. Second, we estimate the diffuse reflectance values $\boldsymbol{r}_p$ in eq.(5) from the pixel values $\boldsymbol{b}_{p1}$ in the 1st basis image and the light source color $\boldsymbol{l}_1$ by element-wise division. Third, we estimate the light source color $\boldsymbol{l}_n$ in eq.(5) for $n \neq 1$ from the pixel values $\boldsymbol{b}_{pn}$ in the n-th basis image and the estimated diffuse reflectance values $\boldsymbol{r}_p$ by element-wise division. In order to robustly estimate the light source color against specular highlights and shadows, we compute the median of the light source colors obtained from all the pixels if necessary. Here, we normalize $\boldsymbol{r}_p$ and $\boldsymbol{l}_n$ so that $r_{p1} + r_{p2} + r_{p3} = 1$ and $l_{n1} + l_{n2} + l_{n3} = 1$, and then put the remaining scale in the intensity of the diffuse reflection d_{pn}. Note that we could resolve the ambiguity in scale among d_{pn}, $\boldsymbol{R}_p$, and $\boldsymbol{l}_n$ by using prior knowledge, but it is out of our focus.

3.3. Resolving Ambiguity in Blind LSS

As shown in Section 3.1, the basis images estimated by using NMF has the ambiguity; $\boldsymbol{B}^{(e)} = \boldsymbol{B}\boldsymbol{X}$. Therefore, the pixel values $\boldsymbol{b}_{pn}^{(e)}$ at the p-th pixel in the n-th estimated basis image is represented as

$$\boldsymbol{b}_{pn}^{(e)} = \sum_{m=1}^{N} \boldsymbol{b}_{pm} x_{mn}, \tag{7}$$

where x_{mn} is the element of $\boldsymbol{X}$ at the m-th row and n-th column. Substituting eq.(5) into eq.(7), we obtain

$$\boldsymbol{b}_{pn}^{(e)} = \begin{pmatrix} r_{p1} & 0 & 0 \\ 0 & r_{p2} & 0 \\ 0 & 0 & r_{p3} \end{pmatrix} \left(\sum_{m=1}^{N} d_{pm} \boldsymbol{l}_m x_{mn} \right). \tag{8}$$

Comparing eq.(5) and eq.(8), we can see that $\sum d_{pm} \boldsymbol{l}_m x_{mn}$ corresponds to the light source color.

If the blind LSS via matrix factorization has no ambiguity, *i.e.* $\boldsymbol{X}$ is equal to the $N \times N$ identity matrix $\boldsymbol{E}_N$, we can show that eq.(8) results in eq.(5) since $x_{mn} = \delta_{mn}$[3]. On the other hand, if the blind LSS has the ambiguity, *i.e.* $\boldsymbol{X} \neq \boldsymbol{E}_N$, the light source color $\sum d_{pm} \boldsymbol{l}_m x_{mn}$ depends on the pixel p. It means that the light source color is non-uniform across the basis image, and contradicts to the assumption of our proposed method that the basis image is taken under one of the light sources and as a result the light source color is uniform. This result suggests that we can resolve the ambiguity in the blind LSS by integrating the LSS with the IID assuming the diffuse reflection model.

In order to resolve the ambiguity, we estimate $\boldsymbol{X}$ from the estimated basis images $\boldsymbol{B}^{(e)}$ and the reflection model in eq.(5) by the nonlinear minimization:

$$\min_{\{\boldsymbol{X}^{-1}, \boldsymbol{B}^{(r)}\}} \frac{1}{P} ||\boldsymbol{B}^{(e)} \boldsymbol{X}^{-1} - \boldsymbol{B}^{(r)}||_{\mathrm{F}}^2 + w \frac{1}{[\det(\boldsymbol{X}^{-1})]^2} \tag{9}$$

subject to $\boldsymbol{B}^{(e)} \boldsymbol{X}^{-1} \geq 0$ and $\boldsymbol{A}^{(e)} \boldsymbol{X}^{-1} \geq 0$. Here, P is the number of pixels and w is the weight for balancing the first and second terms. The matrix $\boldsymbol{B}^{(r)}$ is the basis images reconstructed from eq.(5), and we actually optimize the unknowns in the right-hand side. We use the results of the diffuse IID described in Section 3.2 as the initial values for those unknowns. The second term is required for preventing $\boldsymbol{X}$ from falling into a singular matrix. The initial value for $\boldsymbol{X}^{-1}$ is the identity matrix.

3.4. Resolving Ambiguity in Diffuse IID

In Section 3.2, we assume that the reflectance properties of a scene of interest obey the diffuse reflection model. Since not only diffuse reflection components but also specular reflection components are observed on object surfaces in general, we assume the dichromatic reflection model [27] here. According to the dichromatic reflection model, the color of a specular reflection component is independent of an object surface and is equal to the color of a light source, while the color of a diffuse reflection component depends not only on the light source color but also on the spectral reflectance of the object surface. Then, the pixel values at

[3]The Kronecker delta δ_{mn} is 1 if $m = n$ and 0 if $m \neq n$.

the p-th pixel in the n-th basis image is represented as

$$\begin{aligned} \boldsymbol{b}_{pn} &= d_{pn} \begin{pmatrix} r_{p1} & 0 & 0 \\ 0 & r_{p2} & 0 \\ 0 & 0 & r_{p3} \end{pmatrix} \begin{pmatrix} l_{n1} \\ l_{n2} \\ l_{n3} \end{pmatrix} + s_{pn} \begin{pmatrix} l_{n1} \\ l_{n2} \\ l_{n3} \end{pmatrix} \\ &= d_{pn} \boldsymbol{R}_p \boldsymbol{l}_n + s_{pn} \boldsymbol{l}_n. \end{aligned} \tag{10}$$

Here, s_{pn} is the intensity of the specular reflection at the p-th pixel under the n-th light source.

As shown in Section 3.2, the IID assuming the diffuse reflection model has the ambiguity described by using an unknown positive diagonal matrix $\boldsymbol{Y}$; $\boldsymbol{R}_p^{(e)} = \boldsymbol{R}_p \boldsymbol{Y}$ and $\boldsymbol{l}_n^{(e)} = \boldsymbol{Y}^{-1} \boldsymbol{l}_n$. Therefore, we can rewrite eq.(10) as

$$\boldsymbol{b}_{pn} = d_{pn} \boldsymbol{R}_p \boldsymbol{Y} \boldsymbol{Y}^{-1} \boldsymbol{l}_n + s_{pn} \boldsymbol{Y}^{-1} \boldsymbol{l}_n. \tag{11}$$

We can see that the first term is invariant since $\boldsymbol{Y}\boldsymbol{Y}^{-1} = \boldsymbol{E}_3$, but the second term depends on the diagonal matrix $\boldsymbol{Y}$. Therefore, if we can estimate the light source color from specular highlights, we can fix the unknown matrix $\boldsymbol{Y}$ and resolve the ambiguity.

In order to robustly estimate the light source color and resolve the ambiguity, we make use of RANdom SAmple Consensus (RANSAC) [10]. Here, we assume that the ambiguity in the blind LSS is resolved in advance as described in Section 3.3. Please see Appendix for the details on the light source color estimation.

4. Experiments

To confirm the effectiveness of our proposed method, we conducted a number of experiments using both synthetic and real images. In particular, we confirmed that our method can resolve the ambiguities in the blind LSS and the diffuse IID. For the LSS under AC illumination, we compared the following two methods:

- **LSS via NMF**: the existing method for the blind LSS via NMF [26] that has the ambiguity of a regular matrix $\boldsymbol{X}$ as shown in Section 3.1.

- **Our LSS**: our proposed LSS that resolves the ambiguity in the blind LSS as described in Section 3.3.

For the IID under AC illumination, we compared the following two methods:

- **Diffuse IID**: the IID assuming the diffuse reflection model described in Section 3.2. It has the ambiguity of a diagonal matrix $\boldsymbol{Y}$ as shown in Section 3.2.

- **Our IID**: our proposed IID assuming the dichromatic reflection model that resolves the ambiguity in the diffuse IID as described in Section 3.4.

For NMF in eq.(3), we tested multiple initial conditions, and then found the optimal solution with minimal residual. We empirically set the weight in eq.(9) as $w = 10^{-2}$, and used the interior-point algorithm [7] implemented as fmincon in MATLAB for solving the constrained nonlinear minimization problem of eq.(9). We used it also for solving the nonlinear minimization problem of eq.(12). Our proposed method assumes an arbitrary number of light sources N, but we used image sequences under two light sources in our experiments because the blind LSS [26] is somewhat prone to be sensitive to image noises as the number of light sources increases.

4.1. Synthetic Images

We conducted a number of experiments using synthetic images for which all the ground truths are available. We tested four cases: (i) a pear illuminated by two light sources, (ii) a pear illuminated by two light sources with different temporal intensity profiles from (i), (iii) a turtle illuminated by two light sources with the same colors and directions as (i), and (iv) a turtle illuminated by two light sources with different colors and directions from (i). The image sequences of those scenes were synthesized by using two datasets: one is for the shapes of the objects [1] and the other is for the temporal intensity profiles of the light sources [28]. We added random noises to the synthesized images; we assumed the zero-mean Gaussian noises with the standard deviation $\sigma = 0.01$ for pixel values normalized to [0,1].

Figure 2 shows the results for the first case, *i.e.* a pear illuminated by two light sources: some of the input images, the first and second basis images, the temporal intensity profiles, the diffuse reflection values, and the first and second diffuse and specular intensities from top to bottom. For the LSS under AC illumination, we show the ground truths in (a), and the results obtained by using the LSS via NMF and our LSS in (b) and (c) respectively. For the IID under AC illumination, we show the ground truths in (d), and the results obtained by using the diffuse IID and our IID in (e) and (f) respectively. The numerical values under the recovered basis images show the PSNRs (Peak Signal-to-Noise Ratios); the higher the better.

Comparing the results of the LSS, we can see qualitatively and quantitatively that (c) our LSS performs better than (b) the LSS via NMF. Specifically, the PSNRs of the recovered basis images are improved by using our LSS. We can also see that (f) our IID performs better than (e) the diffuse IID. In particular, we can see that the specular reflection components are accurately recovered by using our IID. Note that both the diffuse IID and our IID can recover the intrinsic properties at a point on the object surface when it is illuminated by both of the two light sources.

Then, we show the results for the second, third, and the

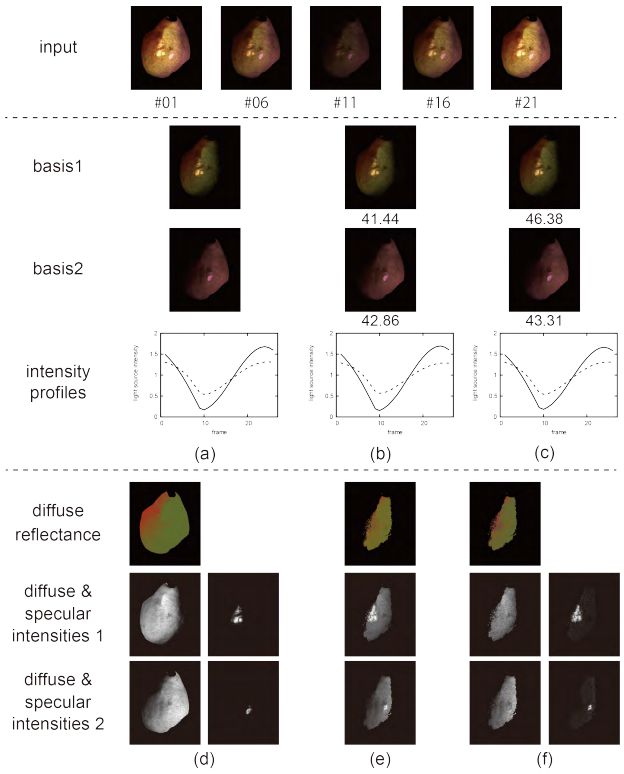

Figure 2. The results on the synthetic images of the 1st scene (a pear illuminated by two light sources): some of the input images, (a) the ground truths of LSS, (b) the LSS via NMF, (c) our LSS, (d) the ground truths of IID, (e) the diffuse IID, and (f) our IID. The numerical values under the recovered basis images are PSNRs.

Table 1. The ambiguity matrix $\boldsymbol{X}$: the ground truth and the matrix estimated by using our LSS.

case	ground truth	our LSS
(i)	$\begin{pmatrix} 0.928 & 0.036 \\ 0.072 & 0.964 \end{pmatrix}$	$\begin{pmatrix} 0.935 & 0.065 \\ 0.065 & 0.935 \end{pmatrix}$
(ii)	$\begin{pmatrix} 0.914 & 0.022 \\ 0.086 & 0.978 \end{pmatrix}$	$\begin{pmatrix} 0.931 & 0.091 \\ 0.069 & 0.909 \end{pmatrix}$
(iii)	$\begin{pmatrix} 0.925 & 0.090 \\ 0.075 & 0.910 \end{pmatrix}$	$\begin{pmatrix} 0.958 & 0.094 \\ 0.042 & 0.906 \end{pmatrix}$
(iv)	$\begin{pmatrix} 0.899 & 0.047 \\ 0.101 & 0.953 \end{pmatrix}$	$\begin{pmatrix} 0.951 & 0.065 \\ 0.049 & 0.935 \end{pmatrix}$

fourth cases in Figure A in the supplementary material, Figure 3, and Figure B in the supplementary material respectively. We can see qualitatively and quantitatively that (c) our LSS performs better than (b) the LSS via NMF and that (f) our IID performs better than (e) the diffuse IID, for light sources with different temporal intensity profiles (Figure A in the supplementary material), for an object with a complex shape (Figure 3), and for light sources with different colors

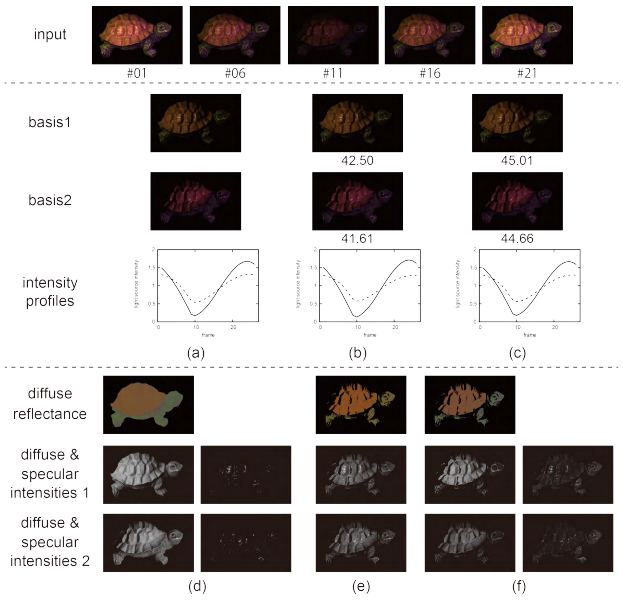

Figure 3. The results on the synthetic images of the 3rd scene (a turtle illuminated by two light sources with the same colors and directions as (i)). Please see the caption for Figure 2.

Table 2. The light source colors: the ground truths, the colors estimated by using the diffuse IID and our IID.

light	ground truth	diffuse IID	our IID
(i) 1	(0.42, 0.35, 0.23)	(0.33, 0.33, 0.33)	(0.42, 0.37, 0.21)
2	(0.38, 0.19, 0.43)	(0.28, 0.16, 0.56)	(0.39, 0.20, 0.41)
(ii) 1	(0.42, 0.35, 0.23)	(0.33, 0.33, 0.33)	(0.41, 0.38, 0.21)
2	(0.38, 0.19, 0.43)	(0.28, 0.16, 0.57)	(0.38, 0.20, 0.43)
(iii) 1	(0.42, 0.35, 0.23)	(0.33, 0.33, 0.33)	(0.42, 0.36, 0.22)
2	(0.38, 0.19, 0.43)	(0.28, 0.17, 0.56)	(0.39, 0.20, 0.41)
(iv) 1	(0.34, 0.40, 0.27)	(0.33, 0.33, 0.33)	(0.34, 0.39, 0.26)
2	(0.24, 0.33, 0.43)	(0.23, 0.27, 0.43)	(0.25, 0.34, 0.42)

and directions (Figure B in the supplementary material).

In Table 1, we show the 2×2 ambiguity matrix $\boldsymbol{X}$: the ground truth and the matrix estimated by using our LSS. In Table 2, we show the light source color: the ground truth and the colors estimated by using the diffuse IID and our IID. Those quantitative results also show that our proposed method works well for resolving the ambiguities in the blind LSS and the diffuse IID.

4.2. Real Images

We conducted a number of experiments using real images of two scenes (A) and (B). The images of those scenes were captured by using a high-speed camera FASTCAM Mini UX50 from Photoron with the frame rate of 2,500 fps with the exposure time of 0.4 ms. The ground truths of the basis images were captured by turning only one of the light

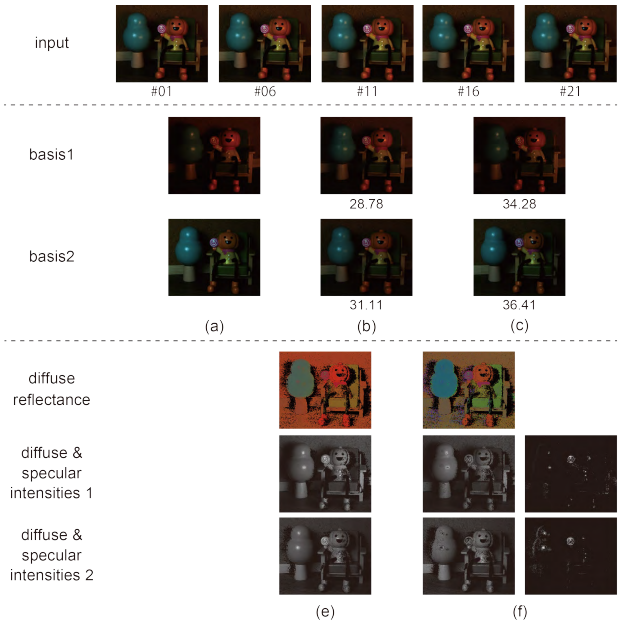

Figure 4. The results on the real images of the scene (A): some of the input images, (a) the ground truths of LSS, (b) the LSS via NMF, (c) our LSS, (e) the diffuse IID, and (f) our IID. The numerical values under the recovered basis images are PSNRs.

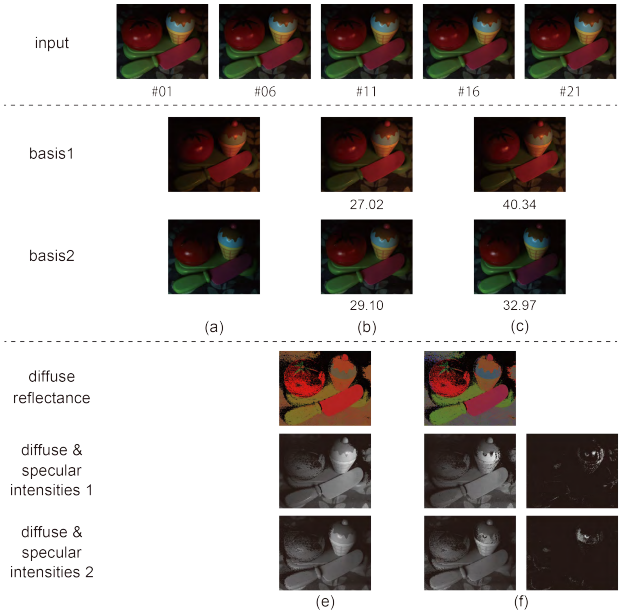

Figure 5. The results on the real images of the scene (B). Please see the caption for Figure 4.

sources on. The light sources were a halogen lamp and an LED, and the ground truths of those light source colors were measured by using a white target for evaluation.

First, Figure 4, shows the results for the scene (A): some of the input images, the first and second basis images, the diffuse reflectance values, and the first and second diffuse and specular intensities from top to bottom. We can see qualitatively that our proposed method works well for real images even though the image quality of the high-speed camera is relatively low and the amplitudes of the light source intensities are small. In particular, we can see that the specular highlights observed on the object surfaces under each of the light sources appear in the first and second specular intensities as expected. Comparing the PSNRs of the recovered basis images in (b) and (c), we can see quantitatively that our LSS performs better than the LSS via NMF. As shown in Figure 5, we obtained similar results for the scene (B).

Second, in Table 3, we show the light source colors: the ground truths and the light source colors estimated by using our proposed method. We can see quantitatively that our method can accurately estimate the light source colors. Those quantitative results show that our method can resolve the ambiguities in the LSS and IID under AC illumination.

Limitations: There are two limitations in our proposed method. First, as is often the case with color-based reflectance separation, our proposed IID does not work well for gray objects since the colors of specular and diffuse components are the same for such objects. Second, as we mentioned in Section 4.1, our IID can recover the intrinsic properties at a point on an object surface if it is illuminated by multiple light sources. Otherwise, our IID without any priors becomes an under-determined problem.

Table 3. The light source colors: the ground truths, the colors estimated by using the diffuse IID and our IID.

light	ground truth	diffuse IID	our IID
(A) 1	(0.56, 0.30, 0.14)	(0.33, 0.33, 0.33)	(0.58, 0.29, 0.13)
2	(0.24, 0.36, 0.40)	(0.16, 0.39, 0.45)	(0.24, 0.33, 0.44)
(B) 1	(0.56, 0.30, 0.14)	(0.33, 0.33, 0.33)	(0.57, 0.28, 0.15)
2	(0.24, 0.36, 0.40)	(0.16, 0.39, 0.45)	(0.20, 0.38, 0.42)

4.3. Application to White Balancing

White balancing is a technique for converting an image taken under unknown light sources as if it were taken under white light sources. In particular, white balancing under multiple light sources with different colors is a challenging problem. This is because the illumination color is non-uniform across the image, in other words, the mixture ratio of the light source colors spatially varies.

Our proposed method for LSS recovers the basis images, each of which is illuminated by only one of the light sources, and our proposed method for IID estimates the light source color per basis image without any ambiguity.

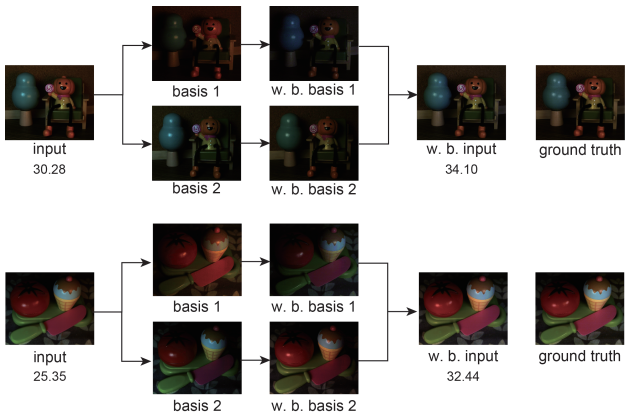

Figure 6. The process and the result of white balancing.

Therefore, our proposed LSS and IID are useful for white balancing; we can change the light source color of each recovered basis image, and then combine the white balanced basis images again. Figure 6 shows the process and the result of white balancing for the real images of the scenes (A) and (B): one of the input images, the basis images, the white-balanced basis images, and the white-balanced input image. The numerical values under the input/white balanced input images are PSNRs.

5. Conclusion and Future Work

In this paper, we proposed a method for LSS and IID under AC illumination. Our proposed method conducts the blind LSS via matrix factorization followed by the IID assuming the dichromatic reflection model, and estimates the intrinsic properties of a scene and light sources from an image sequence under multiple AC light sources. In particular, we revealed the ambiguity in the blind LSS via matrix factorization and the ambiguity in the IID assuming the diffuse reflection model, and then showed why and how those ambiguities can be resolved. We experimentally confirmed that our method works well on real images as well as synthetic images, and that the IID under AC illumination is useful for application to auto white balancing.

Our future work includes LSS and IID under more than two light sources, the improved accuracy by integrating sophisticated noise removal, and the application to image segmentation, object recognition, shape from shading and so on. We hope that this paper contributes to open a new vista of passive computer vision with short exposure time observation.

Acknowledgement: This work was supported by JSPS KAKENHI Grant Number JP20H00612.

A. Light Source Color Estimation

First, we extract the pixels where specular reflection components are observed in the first or second basis images. Specifically, we conduct the diffuse IID, and obtain the light source colors $\boldsymbol{l}_1$ and $\boldsymbol{l}_2$ with the ambiguity as described in Section 3.2. Then, we find the outlier pixels with large errors when the diffuse reflection model is fitted to the pixel values in the first and second basis images. We consider those pixels as specular pixels.

Second, we randomly select two pixels p' and p'' from the specular pixels, and then estimate the reflectance matrix $\boldsymbol{R}_{p'}$ and $\boldsymbol{R}_{p''}$. Specifically, we assume that the pixel values $\boldsymbol{b}_{p'1}$ and $\boldsymbol{b}_{p''1}$ are specular-free, and then estimate the elements of $\boldsymbol{R}_{p'}$ and $\boldsymbol{R}_{p''}$ by element-wise division since the light source color $\boldsymbol{l}_1$ with the ambiguity is computed as described above. Then, we obtain the diffuse reflection colors in the second basis image as $\boldsymbol{R}_{p'}\boldsymbol{l}_2$ and $\boldsymbol{R}_{p''}\boldsymbol{l}_2$ by using the computed light source color $\boldsymbol{l}_2$ with the ambiguity. Note that there are no ambiguities in those diffuse reflection colors as shown in eq.(11).

Third, we compute the candidate of the diagonal matrix $\boldsymbol{Y}$. Since the light source color $\boldsymbol{l}_2$ is parallel to both $(\boldsymbol{b}_{p'2} - d_{dp'}\boldsymbol{R}_{p'}\boldsymbol{l}_2)$ and $(\boldsymbol{b}_{p''2} - d_{dp''}\boldsymbol{R}_{p''}\boldsymbol{l}_2)$ and the diffuse reflection colors in the second basis image $\boldsymbol{R}_{p'}\boldsymbol{l}_2$ and $\boldsymbol{R}_{p''}\boldsymbol{l}_2$ are computed as described above, we find $d_{dp'}$ and $d_{dp''}$ by the minimization:

$$\min_{\{d_{dp'}, d_{dp''}\}} ||(\boldsymbol{b}_{p'2} - d_{dp'}\boldsymbol{R}_{p'}\boldsymbol{l}_2) \times (\boldsymbol{b}_{p''2} - d_{dp''}\boldsymbol{R}_{p''}\boldsymbol{l}_2)||^2. \tag{12}$$

Here, $\times$ stands for the outer product. Thus, we obtain the candidate of the light source color $\boldsymbol{l}_2(p', p'')$, and then compute the candidate of the diagonal matrix $\boldsymbol{Y}(p', p'')$ by element-wise division.

Forth, we evaluate the errors of eq.(11) when the candidate of the diagonal matrix $\boldsymbol{Y}(p', p'')$ is used, and then count the number of inliers. Specifically, we assume that the pixel values $\boldsymbol{b}_{p1}$ in the first basis image are specular-free, and then compute the pixel values in the second basis image and evaluate the error $e_{p2,1}$ between the observed pixel values $\boldsymbol{b}_{p2}$ and the computed pixel values and vice versa. We evaluate the errors $e_{p2,1}$ and $e_{p1,2}$ for all of the specular pixels p, and consider a pixel as an inlier if $e_{p2,1}$ or $e_{p1,2}$ is smaller than a threshold.

Fifth, we repeat the second through fourth steps, and find the candidate of the diagonal matrix with the maximal number of inliers. Finally, we re-compute the diagonal matrix from those inlier pixels by solving eq.(10) via alternating least squares.

References

[1] J. Barron and J. Malik. Shape, illumination, and reflectance from shading. *IEEE TPAMI*, 38(7):1670–1688, 2015. 5

[2] H. Barrow and J. Tenenbaum. Recovering intrinsic scene characteristics from images. *Computer Vision Systems*, pages 3–26, 1978. 1, 2

[3] P. Belhumeur, D. Kriegman, and A. Yuille. The bas-relief ambiguity. *IJCV*, 35:33–44, 1999. 1

[4] M. Berry, M. Browne, A. Langville, V. Pauca, and R. Plemmons. Algorithms and applications for approximate nonnegative matrix factorization. *Computational Statistics & Data Analysis*, 52(1):155–173, 2007. 2, 3

[5] S. Bi, N. Kalantari, and R. Ramamoorthi. Deep hybrid real and synthetic training for intrinsic decomposition. In *Proc. EGSR2018*, 2018. 3

[6] M. Brill. Image segmentation by object color: a unifying framework and connection to color constancy. *JOSA A*, 7(10):2041–2047, 1990. 1

[7] R. Byrd, M. Hribar, and J. Nocedal. An interior point algorithm for large-scale nonlinear programming. *SIAM Journal on Optimization*, 9(4):877–900, 1999. 5

[8] O. Drbohlav and M. Chaniler. Can two specular pixels calibrate photometric stereo? In *Proc. IEEE ICCV2005*, pages II–1850–1857, 2005. 1

[9] Q. Fan, J. Yang, G. Hua, B. Chen, and D. Wipf. Revisiting deep intrinsic image decompositions. In *Proc. IEEE CVPR2018*, pages 8944–8952, 2018. 3

[10] M. Fischler and R. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM*, 24(6):381–395, 1981. 5

[11] B. Funt, M. Drew, and M. Brockington. Recovering shading from color images. In *Proc. ECCV92*, pages 124–132, 1992. 2

[12] E. Garces, C. Rodriguez-Pardo, D. Casas, and J. Lopez-Moreno. A survey on intrinsic images: Delving deep into lambert and beyond. *IJCV*, 130:836–868, 2022. 2

[13] R. Grosse, M. Johnson, E. Adelson, and W. Freeman. Ground-truth dataset and baseline evaluations for intrinsic image algorithms. In *Proc. IEEE ICCV2009*, pages 2335–2342, 2009. 2

[14] B. Horn. *Obtaining Shape from Shading Information*, pages 123–171. MIT Press, 1989. 1

[15] E. Hsu, T. Mertens, S. Paris, S. Avidan, and F. Durand. Light mixture estimation for spatially varying white balance. *ACM TOG*, 27(3):1–7, 2008. 1

[16] A. Hyvärinen and E. Oja. Independent component analysis: algorithms and applications. *Neural Networks*, 13(4–5):411–430, 2000. 2

[17] N. Kong, P. Gehler, and M. Black. Intrinsic video. In *Proc. ECCV2014*, pages 360–375, 2014. 3

[18] P.-Y. Laffont, A. Bousseau, and G. Drettakis. Rich intrinsic image decomposition of outdoor scenes from multiple views. *IEEE TVCG*, 19(2):210–224, 2013. 3

[19] E. Land and J. McCann. Lightness and retinex theory. *JOSA*, 61(1):1–11, 1971. 1, 2

[20] L. Lettry, K. Vanhoey, and L. Van Gool. DARN: a deep adversarial residual network for intrinsic image decomposition. In *Proc. IEEE WACV2018*, pages 1359–1367, 2018. 3

[21] Z. Li and N. Snavely. Learning intrinsic image decomposition from watching the world. In *Proc. IEEE CVPR2018*, pages 9039–9048, 2018. 3

[22] W. Ma, H. Chu, B. Zhou, R. Urtasun, and A. Torralba. Single image intrinsic decomposition without a single intrinsic image. In *Proc. ECCV2018*, pages 201–217, 2018. 3

[23] R. Matsuoka, T. Baba, and M. Okuda. Reflectance estimation and white balancing using multiple images. In *Proc. IEEE ICIP2015*, pages 407–411, 2015. 2

[24] T. Narihira, M. Maire, and S. Yu. Direct intrinsics: Learning albedo-shading decomposition by convolutional regression. In *Proc. IEEE ICCV2015*, pages 2992–2992, 2015. 3

[25] S. Nayar and R. Bolle. Reflectance based object recognition. *IJCV*, 17:219–240, 1996. 1

[26] R. Oya, R. Matsuoka, and T. Okabe. NMF vs. ICA for light source separation under AC illumination. In *Proc. VISAPP2020*, pages 460–465, 2020. 1, 2, 3, 5

[27] S. Shafer. Using color to separate reflection components. *COLOR Research and Application*, 10(4):210–218, 1985. 1, 2, 4

[28] M. Sheinin, Y. Schechner, and K. Kutulakos. Computational imaging on the electric grid. In *Proc. IEEE CVPR2017*, pages 6437–6446, 2017. 1, 2, 5

[29] M. Sheinin, Y. Schechner, and K. Kutulakos. Rolling shutter imaging on the electric grid. In *Proc. IEEE ICCP2018*, pages 1–12, 2018. 2

[30] J. Shi, Y. Dong, H. Su, and S. Yu. Learning non-Lambertian object intrinsics across shapenet categories. In *Proc. IEEE CVPR2017*, pages 1685–1694, 2017. 3

[31] M. Tappen, W. Freeman, and E. Adelson. Recovering intrinsic images from a single image. *IEEE TPAMI*, 27(9):1459–1472, 2005. 2

[32] M. Vollmer and K.-P. Möllmann. Flickering lamps. *European Journal of Physics*, 36(035027), 2015. 1

[33] Y. Weiss. Deriving intrinsic images from image sequences. In *Proc. IEEE ICCV2001*, pages 68–75, 2001. 2

[34] G. Ye, E. Garces, Y. Liu, Q. Dai, and D. Gutierrez. Intrinsic video and applications. *ACM ToG*, 33(4), 2014. 3

[35] H. Zhou, X. Yu, and D. Jacobs. GLOSH: Global-local spherical harmonics for intrinsic image decomposition. In *Proc. IEEE ICCV2019*, pages 7820–7829, 2019. 3