Active Recurrence of Lighting Condition for Fine-Grained Change Detection

Qian Zhang[1,2], **Wei Feng**[1,2*], **Liang Wan**[2,3], **Fei-Peng Tian**[1,2], **Ping Tan**[4]

[1] School of Computer Science and Technology, Tianjin University, China

[2] Key Research Center for Surface Monitoring and Analysis of Cultural Relics, SACH, China

[3] School of Computer Software, Tianjin University, China

[4] School of Computing Science, Simon Fraser University, Canada

{qianz, wfeng, lwan, tianfeipeng}@tju.edu.cn, pingtan@sfu.ca

Abstract

This paper addresses active lighting recurrence (ALR), a new problem that actively relocalizes a light source to physically reproduce the lighting condition for a same scene from single reference image. ALR is of great importance for fine-grained visual monitoring and change detection, because some phenomena or minute changes can only be clearly observed under particular lighting conditions. Hence, effective ALR should be able to online navigate a light source toward the target pose, which is challenging due to the complexity and diversity of real-world lighting & imaging processes. We propose to use the simple parallel lighting as an analogy model and based on Lambertian law to compose an instant navigation ball for this purpose. We theoretically prove the feasibility of this ALR strategy for realistic near point light sources and its invariance to the ambiguity of normal & lighting decomposition. Extensive quantitative experiments and challenging real-world tasks on fine-grained change monitoring of cultural heritages verify the effectiveness of our approach. We also validate its generality to non-Lambertian scenes.

1 Introduction

Image-based monitoring and change detection is an important problem in computer vision and machine learning [Feng *et al.*, 2015; Khan *et al.*, 2017]. Since the lighting condition, camera and scene are three key factors jointly determining the appearance of an image, the success of change detection needs to effectively compensate the influence of varied imaging conditions. Unlike classical change detection assuming fixed camera and slowly changed illuminations [Ide *et al.*, 2016], fine-grained change detection aims to find minute changes of high-value scenes, like cultural heritages, in long time intervals under the wild hosting environments [Stent *et al.*, 2016], wherein how to reproduce
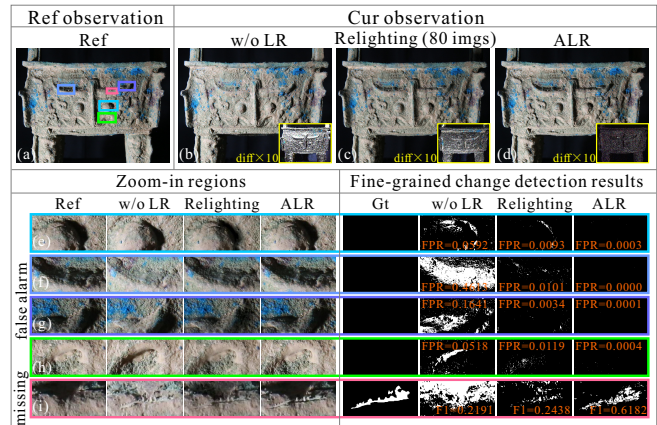
Figure 1: Motivation and importance of active lighting recurrence (ALR) for fine-grained change detection. (a) The reference observation captured under proper near side lighting condition to highlight the rich 3D microstructure of the object. (b)–(d) Current observations of the same object with minute changes, captured under manually aligned lighting (w/o LR), virtually (Relighting using 80 varied illuminations) and ALR reproduced lightings, respectively. The 10-times magnified absolute differences between current and reference observations are shown in the corresponding bottom-right corner. (e)–(i) compare the local microstructures under the reference and different current lightings and their corresponding fine-grained change detection results [Feng *et al.*, 2015]. See text for details.

the reference camera pose and lighting condition is an essential problem and supports many real-world applications, such as accurate surface and material acquisition [Debevec *et al.*, 2000; Ou and Pellacini, 2011], cultural heritage imaging [Tom *et al.*, 2001; Elhabian *et al.*, 2011] and fine-grained change surveillance [Feng *et al.*, 2015; Huang *et al.*, 2017].

Compared to relatively mature camera relocalization (CR) [Williams *et al.*, 2011; Feng *et al.*, 2016; Shi *et al.*, 2018; Miao *et al.*, 2018], lighting recurrence (LR) is more challenging and rarely studied for fine-grained change detection. LR aims to reproduce the lighting condition of a reference image for the same scene with relocalized camera pose. Since slightly varied lightings may cause great false alarmed change detections for the same scene with fixed camera parameters, reproducing the reference lighting condition, either virtually or physically, is a critical issue.
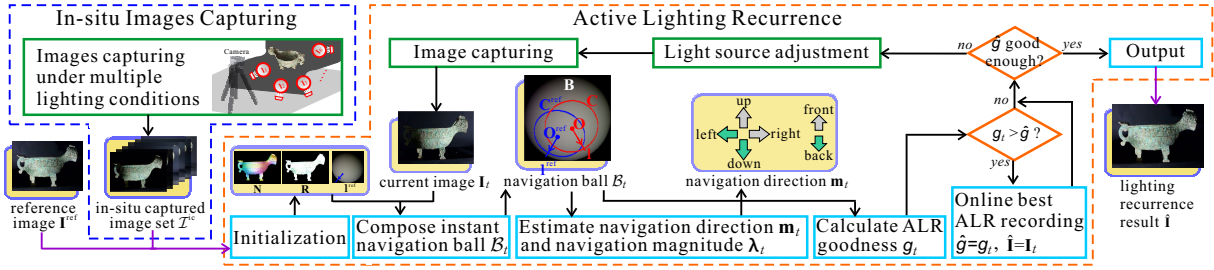
Figure 2: Overall framework of the proposed active lighting recurrence (ALR) approach that works well for realistic near point light sources. The green and light-blue blocks indicate pose-adjustment/image-capturing and online calculation processes, respectively. See text for details.

There are some related works that can deal with the LR problem intuitively, including relighting [Ren *et al.*, 2015; 2013; Liu and Do, 2017], reflectance transformation imaging (RTI) [Tom *et al.*, 2001; Huang *et al.*, 2015; Elhabian *et al.*, 2011], and photometric stereo (PS) [Lu *et al.*, 2015; Shi *et al.*, 2016; Midorikawa *et al.*, 2016]. Nevertheless, most of them just passively take multi-illumination images as input to *virtually* synthesize the target lighting condition, thus belong to *synthetic* or virtual lighting recurrence (SLR). Besides, to achieve higher reconstruction quality, they usually need hundreds of images with different calibrated illuminations [Ren *et al.*, 2015; Tom *et al.*, 2001], complex lighting models and/or sophisticated optimizations [Ren *et al.*, 2015; Lu *et al.*, 2015], which may inevitably lead to reconstruction errors in the relighted image and cannot work in real-time.

As shown in Fig. 1, in real-world fine-grained change monitoring tasks, we usually get the reference observation by casting a particular near side lighting to highlight the rich microstructures of the object, see Fig. 1(a). Hence, in current observation, slight deviations from the reference lighting condition may cause large differences in the relighted image, see Fig. 1(b)–(d). Specifically, without or inaccurate LR may generate two types of errors in fine-grained change detection, that is, rich or particular 3D microstructures caused *false alarms*, and different shading caused *missing*, see Fig. 1(e)–(h) and Fig. 1(i), respectively.[1] Both of them can significantly harm the fine-grained change detection accuracy.

This paper studies ALR that aims to *physically* reproduce the lighting condition from a single reference image. In contrast to SLR, ALR wants to physically relocalize a real light source to exact the same position of the reference one. Clearly, there are three major differences from ALR to the widely-studied SLR. First, besides the reference image, the input of ALR is just a very limited number of currently captured images under varied light source poses. The number of input images for feasible ALR (usually a dozen) is much less than that of SLR (at least hundreds). Hence, mature but sophisticated SLR methods that need large number of images to guarantee reconstruction accuracy are not applicable. Second, ALR is a dynamic process. The purpose of ALR is to produce reliable online navigation guidance from current

lighting pose to the target one, which is not even a problem in any SLR tasks. Third, unlike SLR that prefers accuracy much more than efficiency, instant response and the navigation correctness for each dimension are equally important for the success of ALR. Therefore, ALR is a new problem and no mature SLR methods can be directly used. In fact, to the best of our knowledge, this paper may be the first work on ALR.

In this paper, we focus on ALR of near point light (NPL) that is a commonly-used realistic lighting device in fine-grained change monitoring of cultural heritages. Due to the complexity of NPL model, direct NPL-based ALR is hard to provide online navigation and stable lighting condition estimation. To conquer this problem, as shown in Fig. 2, we propose to use the simple parallel lighting (PL) as an analogy model and compose a navigation ball with two spherical isointensity circles (SICs) indicating the reference and current poses. We theoretically prove that when the two SICs coincide on the navigation ball, the reference shading is strictly recreated on any real object by current light source pose for both PL and realistic NPL models. We also prove the invariance of the proposed ALR strategy to the normal & lighting decomposition ambiguity. As shown in Fig. 1(e)-(i) and our extensive experiments, the proposed ALR works well for both Lambertian and non-Lambertian scenes, and can significantly improve the performance of fine-grained change detection.

2 Active Lighting Recurrence

2.1 Problem Formulation

As shown in Fig. 2, let $\mathbf{I}^{\mathrm{ref}} \in \mathbb{R}^P$ be the reference image, P being the pixel number, $\mathbf{I} = \mathrm{F_L}(\mathcal{L}(\boldsymbol{\rho} \mid \boldsymbol{\Theta}), \mathcal{S}) \in \mathbb{R}^P$ denotes the real lighting/imaging model of the current observation, which is determined by the scene $\mathcal{S}$ (e.g., reflectance, normal, specular regions, if any) and the lighting condition function $\mathcal{L}(\boldsymbol{\rho} \mid \boldsymbol{\Theta})$, where $\boldsymbol{\Theta}$ and $\boldsymbol{\rho}$ indicate the intrinsic (e.g., radiation power, color temperature, intensity distribution) and extrinsic (i.e., position and orientation) parameters of the light source. Since intrinsic lighting parameters $\boldsymbol{\Theta}$ can be easily reproduced using the same light source, we formulate the ALR problem as a *dynamic process* that progressively reproduces the light source pose $\boldsymbol{\rho}$,

$$\hat{\boldsymbol{\rho}} = \arg \min_{\boldsymbol{\rho}} \|\mathrm{F_L}(\mathcal{L}(\boldsymbol{\rho} \mid \boldsymbol{\Theta}), \mathcal{S}) - \mathbf{I}^{\mathrm{ref}}\|_{\mathrm{F}}^2, \qquad (1)$$

$$\begin{aligned} \boldsymbol{\rho}_{t+1} &= \boldsymbol{\rho}_t + \mathrm{diag}(\boldsymbol{\lambda}_t)\mathbf{m}_t, \\ \hat{\boldsymbol{\rho}} &= \lim_{t \to \infty} \boldsymbol{\rho}_t, \end{aligned} \qquad (2)$$

---

[1]The rich 3D microstructures of the object under different near side lightings have different shadows, shadings and specular spots that can cause great false alarmed changes. In contrast, some real changes cannot be clearly observed under the lighting whose direction is close to the normal of the surface where the changes occur.
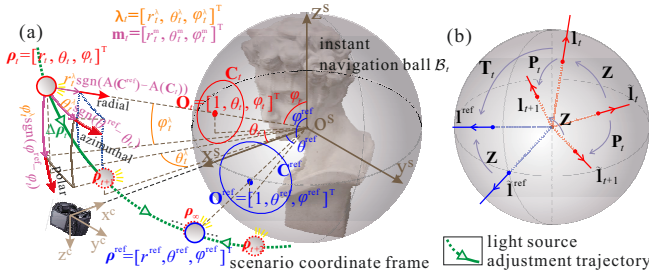
Figure 3: (a) shows instant navigation ball $\mathcal{B}_t$, reference and current SICs $\mathbf{C}^{\mathrm{ref}}$ (blue) & $\mathbf{C}_t$ (red) in the scenario coordinate frame, the parameterization of light source pose $\boldsymbol{\rho}_t = [r_t, \theta_t, \phi_t]^{\mathrm{T}}$ and ALR increment $\Delta\boldsymbol{\rho}_t = \mathrm{diag}(\boldsymbol{\lambda}_t)\mathbf{m}_t$, and light source ALR adjustment trajectory (green line). (b) illustrates the ambiguity matrix $\mathbf{Z}$ of $\mathbf{N}$ & $\mathbf{l}$ decomposition for Lemma 3. See text for details.
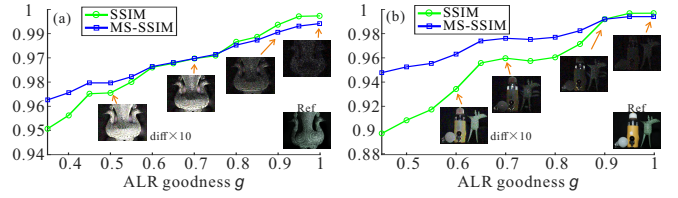

Figure 4: ALR goodness g score with (MS-)SSIM change for scenes S1 (a) and S5 (b), respectively. The difference maps of reference image and some intermediate images of ALR are also shown.

where $\mathbf{m}_t$ and $\boldsymbol{\lambda}_t$ indicate the t-th ALR navigation direction and magnitude to adjust the light source pose, $\mathrm{diag}(\cdot)$ is the diagonalization of a vector. Note, in our formulation, Eq. (1) defines the ALR objective, while Eq. (2) is the progressive ALR strategy, i.e., at t-th iteration, we physically adjust the current light source pose by $\Delta\boldsymbol{\rho}_t = \mathrm{diag}(\boldsymbol{\lambda}_t)\mathbf{m}_t$. Hence, the convergence and goodness of an ALR approach relies on the correctness of $\mathrm{diag}(\boldsymbol{\lambda}_t)$ and $\mathbf{m}_t$, and $\lim_{t\to\infty} \mathrm{diag}(\boldsymbol{\lambda}_t) = \mathbf{0}$.

To provide accurate and instant ALR navigation, we first compose an *analogy parallel lighting* (apl) based instant navigation ball via Lambertian law and prove its invariance to normal & lighting decomposition ambiguity in Sec. 2.2. We then prove the feasibility of this simple online ALR-apl strategy to the common realistic near point light sources in Sec. 2.3.

2.2 Analogy Parallel Lighting-Based ALR

We use the simple parallel lighting as an analogy model, abbreviated to apl. Let $\mathbf{l} \in \mathbb{R}^3$ denote the parallel lighting vector, whose magnitude and direction indicate the lighting strength and direction, respectively. Let $\mathbf{N} \in \mathbb{R}^{P\times 3}$ and $\mathbf{R} \in \mathbb{R}^P$ be the normal and grayscale reflectance of the scene, respectively. Under this model, we have $\mathcal{L}(\boldsymbol{\rho} \mid \boldsymbol{\Theta}) = \mathbf{l}$ and $\mathcal{S} = \{\mathbf{R}, \mathbf{N}\}$, with $\mathbf{R}$ and $\mathbf{N}$ being the scene reflectance and normal, respectively. Thus, the ALR formulation Eqs. (1)–(2) can be reduced to a much simpler ALR-apl problem

$$\hat{\mathbf{l}} = \arg\min_{\mathbf{l}} \|\mathbf{R} \circ \mathbf{Nl} - \mathbf{I}^{\mathrm{ref}}\|_{\mathrm{F}}^2, \qquad (3)$$

$$\begin{aligned} \mathbf{l}_{t+1} &= \mathbf{l}_t + \mathrm{diag}(\boldsymbol{\lambda}_t)\mathbf{m}_t, \\ \hat{\mathbf{l}} &= \lim_{t\to\infty} \mathbf{l}_t, \end{aligned} \qquad (4)$$

where the $\circ$ is elementwise multiplication.

Thanks to the simplicity of ALR-apl model, it is possible to online calculate both the navigation direction $\mathbf{m}_t$ and magnitude $\boldsymbol{\lambda}_t$ from the reference and current images, $\mathbf{I}^{\mathrm{ref}}$ and $\mathbf{I}_t$. Specifically, Fig. 2 shows the working flow of the ALR-apl approach. To get a reliable initialization, we first roughly capture K different side lighting images to form the in-situ captured image set $\mathcal{I}^{\mathrm{ic}}$. From $\mathcal{I}^{\mathrm{ic}}$ and $\mathbf{I}^{\mathrm{ref}}$, we obtain scene normal $\mathbf{N}$, reflectance $\mathbf{R}$ and reference lighting vector $\mathbf{l}^{\mathrm{ref}}$. Then, in each ALR-apl iteration, we compose an instant navigation ball $\mathcal{B}$ and online calculate the navigation direction $\mathbf{m}_t$ and

magnitude $\mathrm{diag}(\boldsymbol{\lambda}_t)$ for light source adjustment. We analyze the convergence and efficacy of this process at last.

Initialization

By parallel lighting analogy, we have

$$\mathbf{S} = \mathbf{Nl}, \qquad (5)$$

where $\mathbf{S} \in \mathbb{R}^P$ is shading image that can be obtained by disambiguated intrinsic image decomposition [Zhao *et al.*, 2012], by solving $\mathbf{I} = \mathbf{R} \circ \mathbf{S}$. Clearly, from $\mathcal{I}^{\mathrm{ic}}$ and $\mathbf{I}^{\mathrm{ref}}$, we can obtain $K+1$ shading images (K in-situ captured side lighting images, 1 reference image). They are used to estimate the scene normal $\mathbf{N}$ and reference lighting vector $\mathbf{l}^{\mathrm{ref}}$ via a fast state-of-the-art uncalibrated photometric stereo algorithm, LDR [Favaro and Papadhimitri, 2012]. In practice, we only need $K \approx 12$ side lighting images for initialization.

Instant Navigation Ball $\mathcal{B}$

Given $\mathbf{N}$ and $\mathbf{R}$, current lighting vector $\mathbf{l}_t$ of $\mathbf{I}_t$ can be easily obtained by solving Eq. (5) in closed-form. To eliminate scene dependency in ALR process, we online render both current and reference lightings, $\mathbf{l}_t$ and $\mathbf{l}^{\mathrm{ref}}$, onto a unit sphere $\mathbf{N}^{\mathrm{s}}$, rather than on the real scene normal $\mathbf{N}$. Specifically, let $\mathbf{B}^{\mathrm{ref}}$ and $\mathbf{B}_t$ be the reference and currently rendered images of the unit sphere, i.e., $\mathbf{B}^{\mathrm{ref}} = \mathbf{N}^{\mathrm{s}}\mathbf{l}^{\mathrm{ref}}$ and $\mathbf{B}_t = \mathbf{N}^{\mathrm{s}}\mathbf{l}_t$, where $\mathbf{N}^{\mathrm{s}}$ is the sphere normal. From the Lambertian law, we can easily obtain the following proposition about the spherical isointensity sets $\mathbf{C}^{\mathrm{ref}}$ and $\mathbf{C}_t$, formed by rendered pixels with some particular intensity value, e.g., the median of $\mathbf{B}^{\mathrm{ref}}$.

Proposition 1 (SICs & shading equivalence). *With analogy parallel lighting and Lambertian law, given an arbitrary lighting condition $\mathbf{l}$, the spherical isointensity set $\mathbf{C}$ always forms a circle under the view of lighting direction, which can be named as spherical isointensity circle (SIC). Iff the reference and current SICs $\mathbf{C}^{\mathrm{ref}}$ and $\mathbf{C}_t$ coincide completely, the reference and current images, $\mathbf{I}^{\mathrm{ref}}$ & $\mathbf{I}_t$, are the same.*

Due to the regular shape of a SIC, we can always make the reference and current SICs, $\mathbf{C}^{\mathrm{ref}}$ and $\mathbf{C}_t$, coincide on the rendered sphere by adjusting light source pose. Proposition 1 guarantees the shading equivalence between the real reference and current images, once the corresponding SICs coincide. Hence, as shown in Fig. 3(a), we can dynamically compose an *instant navigation ball* $\mathcal{B}_t$ to provide effective instant ALR navigation,

$$\mathcal{B}_t = \{\mathbf{B}_t, \mathbf{C}_t, \mathbf{O}_t, \mathbf{C}^{\mathrm{ref}}, \mathbf{O}^{\mathrm{ref}}\}, \qquad (6)$$

where $\mathbf{B}_t$ is the current rendered image on the unit sphere, $\mathbf{O}^{\mathrm{ref}}$ and $\mathbf{O}_t$ are the center coordinates of $\mathbf{C}^{\mathrm{ref}}$ and $\mathbf{C}_t$, re-

spectively. As shown in Fig. 3(a), In fact, the center coordinate of a SIC indeed indicates the corresponding lighting direction, i.e., $\mathbf{O}^{\text{ref}} = \frac{\mathbf{l}^{\text{ref}}}{\|\mathbf{l}^{\text{ref}}\|}$ and $\mathbf{O}_t = \frac{\mathbf{l}_t}{\|\mathbf{l}_t\|}$.

Online Estimation of Navigation Direction $\mathbf{m}_t$

Note, see Fig. 3(a), an analogy parallel light source pose $\boldsymbol{\rho}_t$ can be expressed by vector $[r_t, \theta_t, \phi_t]^{\text{T}}$ in the light source spherical coordinate system, with $r_t = \|\mathbf{l}^{\text{ref}}\|$ being the lighting strength, θ_t & ϕ_t indicating the lighting direction in the azimuthal and polar axes, respectively. Therefore, according to Proposition 1, we can effectively estimate the ALR navigation direction from current lighting $\mathbf{l}_t$ toward the target one $\mathbf{l}^{\text{ref}}$ in radial, azimuthal and polar axes as

$$\mathbf{m}_t = \text{sgn}([A^{\text{ref}}, \theta^{\text{ref}}, \phi^{\text{ref}}]^{\text{T}} - [A_t, \theta_t, \phi_t]^{\text{T}}), \qquad (7)$$

where $A^{\text{ref}} = \text{A}(\mathbf{C}^{\text{ref}})$ and $A_t = \text{A}(\mathbf{C}_t)$ are the area of SICs $\mathbf{C}^{\text{ref}}$ and $\mathbf{C}_t$, respectively, $\text{A}(\cdot)$ is region area, $\text{sgn}(\cdot)$ is the sign function. As illustrated by Fig. 3(a), $\mathbf{m}_t$ reflects the positive $+1$ or negative -1 ALR directions along the three axes of light source spherical coordinate system, for the apl model.

Online Estimation of Navigation Magnitude $\boldsymbol{\lambda}_t$

With $\mathbf{m}_t$ and fast estimation of current lighting vector $\mathbf{l}_t$, we can establish a manual control loop with $[A^{\text{ref}}, \theta^{\text{ref}}, \phi^{\text{ref}}]^{\text{T}}$ as desired set-point (SP) and $[A_t, \theta_t, \phi_t]^{\text{T}}$ being the measured process variable (PV). In fact, in our real-world tasks of fine-grained change detection of cultural heritages, we mainly relied on this *manual ALR* to do the job, due to its great portability.

Nevertheless, as shown in Fig. 3(a), we can also effectively estimate the ALR navigation magnitude $\boldsymbol{\lambda}_t$ via *bisection approaching*, which together with $\mathbf{m}_t$ can enable an *automatic ALR* process, with the help of a robotic arm. Specifically, given an initial $\boldsymbol{\lambda}_0 = [r_0^{\lambda}, \theta_0^{\lambda}, \phi_0^{\lambda}]^{\text{T}}$, where $r_0^{\lambda}, \theta_0^{\lambda}, \phi_0^{\lambda}$ indicate the initial light source adjustment magnitude in radial, azimuthal and polar axes, respectively. In our experiments, we empirically set $r_0^{\lambda} = \theta_0^{\lambda} = \phi_0^{\lambda} = 3$mm in our robotic platform. In t-th ($t > 1$) ALR iteration, we have

$$a_t^{\lambda} = \begin{cases} \frac{1}{2}a_{t-1}^{\lambda} & a_t^{\text{m}}a_{t-1}^{\text{m}} < 0, \\ \mu a_{t-1}^{\lambda} & \text{otherwise,} \end{cases} \qquad (8)$$

where $a \in \{r, \theta, \phi\}$ denotes the three independent spherical axes, μ is the speed-up rate of navigation magnitude and is empirically set as 1.2 in our experiments. With Eq. (8), we can efficiently update $\boldsymbol{\lambda}_t$ and obtain the ALR increment $\Delta\boldsymbol{\rho}_t = \text{diag}(\boldsymbol{\lambda}_t)\mathbf{m}_t$, which is directly applied on the robotic platform to finish the t-th ALR adjustment.

See Fig. 3(a), Eq. (8) indeed defines a bisection approaching strategy. To understand the convergence of this strategy, we conceptually separate the ALR process into two stages. First, regardless of the spherical position (θ_t, ϕ_t) of the center $\mathbf{O}_t$ of current SIC $\mathbf{C}_t$, we first make the area of $\mathbf{C}_t$ equal to that of $\mathbf{C}^{\text{ref}}$. That is, we first reproduce the target lighting strength of the analogy parallel light source. In this condition, current lighting pose $\boldsymbol{\rho}_t$ and the reference pose $\boldsymbol{\rho}^{\text{ref}}$ lie on the same great circle of a bigger sphere whose center coincides with that of the scenario coordinate frame. $\mathbf{m}_t$ indicates the 1D relative position from $\boldsymbol{\rho}_t$ toward $\boldsymbol{\rho}^{\text{ref}}$ on this great circle.
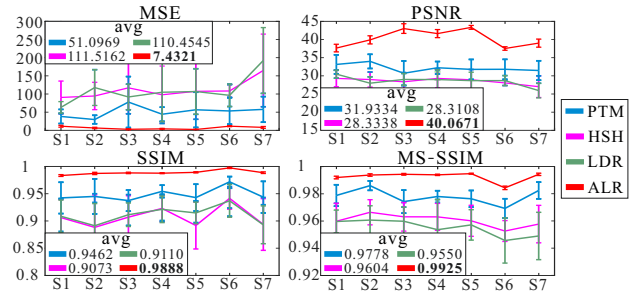

Figure 5: Quantitative comparisons of our ALR approach and 3 baselines for 7 scenes with near point light source.

Hence, the second stage is to use the bisection strategy to approach $\boldsymbol{\rho}^{\text{ref}}$, by iteratively checking the relative direction of $\mathbf{m}_t$ and $\mathbf{m}_{t-1}$. That is, on the great circle, if $\boldsymbol{\rho}_t$ do not reach $\boldsymbol{\rho}^{\text{ref}}$, we enlarge the step size; if $\boldsymbol{\rho}_t$ crosses $\boldsymbol{\rho}^{\text{ref}}$, we bisect the step size. Since in each cases $\mathbf{m}_t$ guarantees the right direction, it is clear that such strategy can effectively lead $\boldsymbol{\rho}_t$ toward $\boldsymbol{\rho}^{\text{ref}}$, i.e., $\lim_{t \to \infty} \text{diag}(\boldsymbol{\lambda}_t) = \mathbf{0}$. See the green curve of Fig. 3(a) for an example of light source adjustment trajectory.

Algorithm & Implementation Details

Fig. 2 shows the detailed working flow of our ALR-apl approach. This process is terminated by an ALR goodness g that measures the recurrence accuracy by the overlap ratio of $\mathbf{C}^{\text{ref}}$ and $\mathbf{C}_t$,

$$g = \frac{\text{A}(\text{reg}(\mathbf{C}_t) \cap \text{reg}(\mathbf{C}^{\text{ref}}))}{\text{A}(\text{reg}(\mathbf{C}_t) \cup \text{reg}(\mathbf{C}^{\text{ref}}))}. \qquad (9)$$

If g is good enough, e.g., $g > 0.99$, the ALR process stops.

Invariance to $\mathbf{N}$ & $\mathbf{l}$ Decomposition Ambiguity

According to Eq. (5), the $\mathbf{N}$ & $\mathbf{l}$ decomposition generally subject to an ambiguity matrix $\mathbf{Z}$,

$$\mathbf{S} = \mathbf{N}\mathbf{l} = \tilde{\mathbf{N}}\mathbf{Z}^{-1}\mathbf{Z}\tilde{\mathbf{l}}, \qquad (10)$$

where $\tilde{\mathbf{N}} \in \mathbb{R}^{P \times 3}$ and $\tilde{\mathbf{l}} \in \mathbb{R}^3$ are the ground truth normal and lighting condition. Therefore, our initialization also exits an ambiguity matrix $\mathbf{Z}$ between the estimated and real scene normal $\mathbf{N}$ and lighting direction $\mathbf{l}$, which may influence the correctness of ALR. Fortunately, we prove that if the estimation error of scene normal $\mathbf{N}$ is smaller than $\frac{\pi}{3}$, the ambiguity matrix $\mathbf{Z}$ does not affect the effectiveness and convergence of our ALR approach.

Lemma 1. *The ambiguity matrix $\mathbf{Z}$ generated by the decomposition of Eq. (5) is a rotation matrix.*

Lemma 2. *The radial-axis navigation direction, i.e., r_0^{m}, is independent to $\mathbf{Z}$.*

Lemma 3. *Let $\langle \beta, \mathbf{e} \rangle$ be axis-angle representation of $\mathbf{Z}$. For the azimuthal and polar ALR direction guidance in Eq. (7), if $\beta < \frac{\pi}{3}$, $\mathbf{Z}$ does not affect our ALR process. We can faithfully relocalize the light source to reference apl pose.*

Brief proof. *Lemma 2 ensures the lighting strength can be reproduced, regardless of $\mathbf{Z}$. We need only relocalize $\mathbf{O}_t$ to $\mathbf{O}^{\text{ref}}$ on the unit sphere. In this condition, as shown in Fig. 3(b), light source increment $\Delta\boldsymbol{\rho}_t$ is equivalent to*
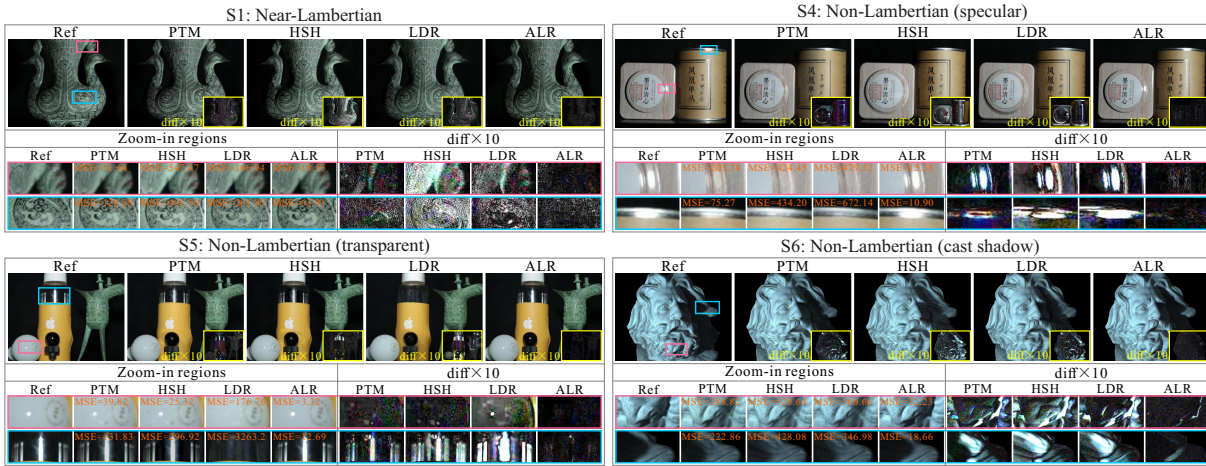
Figure 6: Visual comparisons of our ALR method and 3 baselines for one near-Lambertian scene (S1) and 3 non-Lambertian scenes (S4–6). For each scene, the first row shows the reference image and the lighting recurrence images of 4 methods. The second row shows the zoom-in regions and corresponding difference maps (magnified by 10). Besides, the MSE values of zoom-in regions are also provided.

a rotation matrix $\mathbf{T}_t$. Considering $\mathbf{Z}$, we have $\mathbf{T}_{t+1} = \mathbf{T}_t \mathbf{Z} \mathbf{P}_t^{-1} \mathbf{Z}^{-1}$, where $\mathbf{P}_t$ is the real light source adjustment realized by robotic arm movement or manual adjustment. Normally, $\mathbf{P}_t = \mathbf{T}_t$. So, $\mathbf{T}_{t+1} = \mathbf{T}_t \mathbf{Z} \mathbf{T}_t^{-1} \mathbf{Z}^{-1}$. Then, the conclusion follows by Theorem 1 of [Feng et al., 2016]. □

The convergence necessary condition $\beta < \frac{\pi}{3}$ of Lemma 3 can be easily satisfied by current photometric stereo methods [Favaro and Papadhimitri, 2012], which is empirically verified and discussed in detail in Sec. 3.1.

2.3 ALR under Realistic Near Point Lighting

In fine-grained change detection, near point lighting (NPL) is a commonly-used realistic model that assigns different scene points with distance-related lighting directions,

$$\mathbf{S}_p = e\mathbf{N}_p(\boldsymbol{\rho} - \mathbf{X}_p)/\|\boldsymbol{\rho} - \mathbf{X}_p\|^3, \quad (11)$$

where $\mathbf{S}_p$ is the shading intensity of pixel p, $\boldsymbol{\rho} \in \mathbb{R}^3$ is the near point light source position, $\mathbf{X}_p \in \mathbb{R}^3$ indicates the spatial coordinate of point p, e is the lighting power. Eq. (11) describes the quadratic attenuation effect of lighting strength w.r.t. the distance between the light source and scene. In the following, we prove that the proposed ALR-apl approach still works for such realistic NPL light source, thus avoiding sophisticated estimation of the NPL parameters. In this case, $\mathbf{l}^{\text{ref}}$ and $\mathbf{l}$ in Fig. 2 now denote the equivalent parallel light vector of the near point lighting.

Proposition 2 (SICs & shading equivalence). *Under NPL model, if we locate the point light source to an arbitrary position, then the spherical isointensity set $\mathbf{C}$ acquired from the rendered ball image $\mathbf{B}$ always forms a circle, under the view that light source points to the sphere center. Iff the reference and current SICs $\mathbf{C}^{\text{ref}}$ and $\mathbf{C}_t$ coincide completely, the reference and current images, $\mathbf{I}^{\text{ref}}$ & $\mathbf{I}_t$, are the same.*

Proposition 3 (ALR navigation equivalence). *The ALR approach for the apl model is also applicable to NPL model.*

Brief proof. We use superscripts apl and NPL to distinguish the apl model and NPL model. Given any two points p and

q, it is easy to verify $\mathbf{B}_p^{\text{apl}} \star \mathbf{B}_q^{\text{apl}} \Leftrightarrow \mathbf{B}_p^{\text{NPL}} \star \mathbf{B}_q^{\text{NPL}}$, where $\star$ denotes an operator in $\{>, <, =\}$. That is, $\mathbf{B}_p^{\text{apl}}$ and $\mathbf{B}_p^{\text{NPL}}$ have the same order relation. So, the proposed ALR guidance also applies to near point lighting condition. □

3 Experimental Results

We build 7 scenes (S1–7) to evaluate our ALR method and baselines. S1–3 are near-Lambertian scenes. S4–7 are non-Lambertian scenes, including many specular (S4, S7), transparent (S5) and cast shadow (S6) regions respectively. For each scene, we collect 20 multi-illumination images by a Canon 5D Mark III camera. We use a small LED bulb as the near point light source to carry out the experiments.

3.1 Convergence and Effectiveness Validation

In our ALR, we use ALR goodness g to determine the best lighting recurrence result. To verify the effectiveness, we record all images and the corresponding g during ALR for S4 and S5, then we compute SSIM and MS-SSIM [Wang et al., 2003] for each image. Fig. 4(a) and (b) show the relation of ALR goodness g and the average (MS-)SSIM for S4 and S5, respectively. Besides, the difference maps (magnified by 10) of reference image and some recorded images during ALR are also shown. Since the minute change of lighting condition may cause large differences in image, it is reasonable that the (MS-)SSIM may decrease slightly sometimes, but the (MS-)SSIM increases with the increase of g in global. It means that g is an effective goodness criterion for our ALR.

We prove that the ambiguity matrix $\mathbf{Z}$ does not influence our ALR method if $\beta < \frac{\pi}{3}$, where β is the angle of axis-angle representation of matrix $\mathbf{Z}$. In fact, the condition is easy to be met. To verify this, we introduce the dataset [Xiong et al., 2015] which includes 7 statue scenes (e.g., Cat, Frog, Hippo) and corresponding ground truth normals. Each scene has 20 multi-illumination images. We estimate the scene normal by LDR [Favaro and Papadhimitri, 2012] and calculate the corresponding mean angle error (MAE) for each scene. We find that the average, maximum and minimum MAEs are $\frac{\pi}{27}$, $\frac{\pi}{15}$
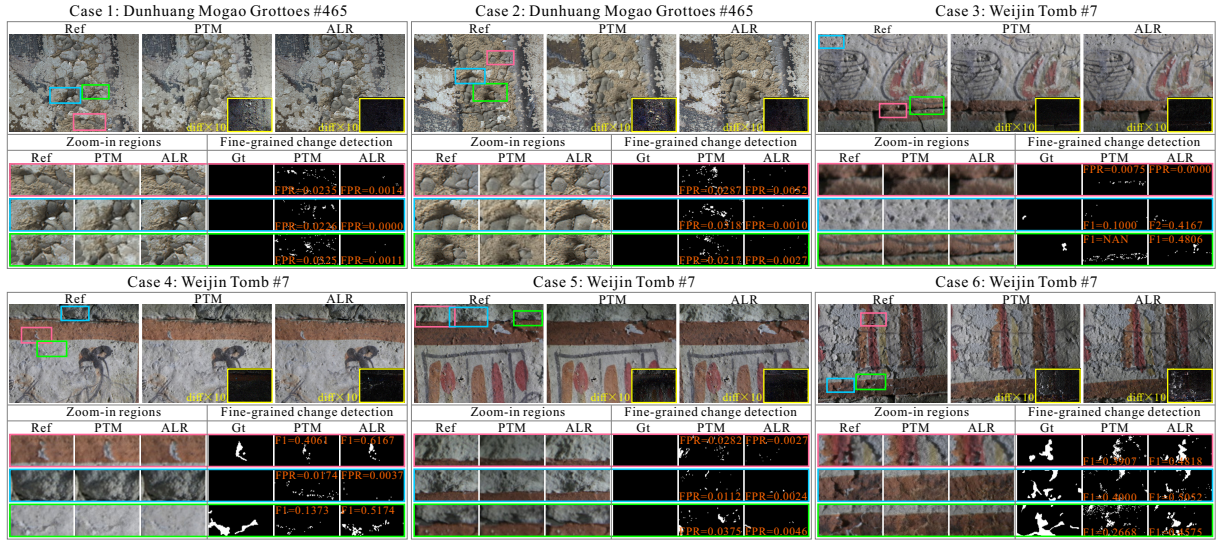
Figure 7: Comparisons of ALR and PTM methods on fine-grained change detection of ancient murals in Dunhuang Mogao Grottoes #465 (Case 1–2) and Weijin Tomb #7 (Case 3–6). We use 20 images under different lighting conditions to conduct PTM method. For each case, the zoom-in regions and corresponding change detection results are shown in the bottom, including the FPR/F1-Measure values.

and $\frac{\pi}{45}$ respectively, which are much less than $\frac{\pi}{3}$. Therefore, according to Lemma 3, we can confidently say that the ambiguity matrix $\mathbf{Z}$ does not influence our ALR method.

Our ALR method needs not large number of images. To verify this, we carry out ALR for 7 scenes (mentioned later) with variable number of multi-illumination images. The 2-th to 5-th rows of Table 1 show the averages of 4 commonly-used criteria MSE, PSNR, SSIM and MS-SSIM under different image numbers. The 6-th row of Table 1 shows the MAE between each estimated scene normal and the one using 100 images. We can see that the MAE reduces with the increase of image number, and the recurrence accuracy of ALR is stable and always well. It also proves that our ALR is effective as long as the estimated scene normal satisfies Lemma 3. In practice, we commonly collect about 20 images for our ALR.

3.2 Quantitative Comparison

To compare with our ALR method, we use PTM [Tom *et al.*, 2001], HSH [Elhabian *et al.*, 2011], LDR [Favaro and Papadhimitri, 2012] as our baselines. PTM [Tom *et al.*, 2001] and HSH [Elhabian *et al.*, 2011] are two image-based relighting methods, and we use a light probe to calibrate the lighting direction for each captured image. LDR [Favaro and Papadhimitri, 2012] is a state-of-the-art uncalibrated photometric stereo method based on parallel lighting model. We carry out the 3 methods using the captured 20 images and generate 20 lighting recurrence images. We use MSE, PSNR, SSIM, MS-SSIM [Wang *et al.*, 2003] as accuracy metrics.

# Imgs	5	10	20	40	60	80	100
MSE	3.4289	3.2433	2.5803	2.8792	2.6086	2.7137	2.8708
PSNR	42.7792	43.0208	44.0140	43.5379	43.9665	43.7950	43.5506
MSSSIM	0.9953	0.9953	0.9961	0.9957	0.9961	0.9960	0.9957
MS_MSSIM	0.9964	0.9964	0.9967	0.9965	0.9967	0.9966	0.9966
MAE	$\frac{\pi}{20}$	$\frac{\pi}{23}$	$\frac{\pi}{49}$	$\frac{\pi}{87}$	$\frac{\pi}{136}$	$\frac{\pi}{305}$	0

Table 1: ALR accuracy vs. #images used in initialization.

Near-Lambertian Scenes

Fig. 5 shows the quantitative comparisons of our ALR method and 3 baselines for near-Lambertian scenes (S1–3). Each node indicates the average evaluation of 20 lighting recurrence results and the up and down bar of each node denotes the variance. Besides, the average of the 4 criteria for all scenes are also shown. Fig. 6 shows the visual comparisons of 4 methods for S1, including the zoom-in regions and corresponding difference maps. We see that our ALR method can generate more accurate lighting recurrence results than baselines for both quantitative and visual comparisons. PTM has the next best recurrence accuracy, and our ALR improves PSNR score by about 29 percent than PTM. Besides, we can see that even if the difference map has been magnified by 10, the difference in image is not obvious.

Non-Lambertian Scenes

As shown in Fig. 5 (S4–7) and Fig. 6 (S4–6), we can also see that our ALR method is more effective than baselines for non-Lambertian scenes. From the zoom-in regions and difference maps in Fig. 5, we find that the baselines can not faithfully reproduce the surface details for non-Lambertian scenes generally, e.g., the missing of specular region for S4 in Fig. 5. Besides, compared with the near-Lambertian scenes, the baselines generally generate mush worse results for the non-Lambertian scenes.

3.3 Real-World Applications

We apply our ALR to actively capture and measure the fine-grained changes of ancient murals in two World Cultural Heritage Sites, Dunhuang Mogao Grottoes (Case 1–2) and Weijin Graves (Case 3–6) in Fig. 7. Specifically, given the reference image, we first relocalize current camera pose via ACR [Feng *et al.*, 2016], then we do ALR to physically reproduce the lighting condition of reference image and take current images for evaluation. Besides, we also capture 20 images under

different lighting conditions in current observation and carry on PTM [Tom *et al.*, 2001] to generate relighting images for comparison. Fine-grained changes are detected by FGCD algorithm [Feng *et al.*, 2015] for both ALR and PTM results.

Fig. 7 shows 6 real cases of fine-grained changes of ancient murals, which are caused by either natural deteriorations (Case 1–2) or artificial intervention, e.g. mural restoration (Case 3–6), respectively. Refer to the zoom-in regions and corresponding FGCD results, we can clearly see that our ALR can generate much higher F1-Measure and lower FPR errors. This is because that some surface details cannot be faithfully reproduced by PTM, e.g., the cast shadow of Case 2 in Fig. 7. In contrast, our ALR supports much more accurate fine-grained change detection.

4 Conclusion

We have studied a new problem, active lighting recurrence (ALR), that actively reproduces the actual lighting condition of a single reference image. To achieve instant and accurate ALR guidance, we propose a simple yet effective analogy parallel lighting (apl) based ALR approach. We show that the proposed approach works well for the commonly-used realistic near point lighting model, with strict theoretical equivalence and convergence guarantees. Extensive quantitative experiments and challenging real-world tasks on fine-grained change monitoring of cultural heritages verify the effectiveness and superiority of our approach. In the future, we plan to explore full 6D ALR for more general near surface lighting model. We hope this work could encourage further studies of active recurrence of imaging conditions, for both camera and lighting, and their real-world applications. Besides, we are also interested in exploring spherical panorama [Zhao *et al.*, 2015] based active environment lighting recurrence and using region-level strategy [Feng *et al.*, 2010] to further speedup and stabilize the ALR process.

References

[Debevec *et al.*, 2000] P. Debevec, T. Hawkins, C. Tchou, H.P. Duiker, W. Sarokin, and M. Sagar. Acquiring the reflectance field of a human face. In *ACM SIGGRAPH*, 2000.

[Elhabian *et al.*, 2011] S.Y. Elhabian, R. Ham, and A.A. Farag. Towards accurate and efficient representation of image irradiance of convex-lambertian objects under unknown near lighting. In *ICCV*, 2011.

[Favaro and Papadhimitri, 2012] P. Favaro and T. Papadhimitri. A closed-form solution to uncalibrated photometric stereo via diffuse maxima. In *CVPR*, 2012.

[Feng *et al.*, 2010] W. Feng, J. Jia, and Z.Q. Liu. Self-validated labeling of Markov random fields for image segmentation. *IEEE TPAMI*, 32(10):1871–1887, 2010.

[Feng *et al.*, 2015] W. Feng, F.P. Tian, Q. Zhang, N. Zhang, L. Wan, and J. Sun. Fine-grained change detection of misaligned scenes with varied illuminations. In *ICCV*, 2015.

[Feng *et al.*, 2016] W. Feng, F.P. Tian, Q. Zhang, and J. Sun. 6D dynamic camera relocalization from single reference image. In *CVPR*, 2016.

[Huang *et al.*, 2015] X. Huang, M. Walton, G. Bearman, and O. Cossairt. Near light correction for image relighting and 3d shape recovery. In *Digital Heritage*, 2015.

[Huang *et al.*, 2017] R. Huang, W. Feng, Z. Wang, M.Y. Fan, L. Wan, and J. Sun. Learning to detect fine-grained change under variant imaging conditions. In *ICCVW*, 2017.

[Ide *et al.*, 2016] T. Ide, D.T. Phan, and J. Kalagnanam. Change detection using directional statistics. In *IJCAI*, 2016.

[Khan *et al.*, 2017] S. Khan, X. He, F. Porikli, M. Bennamoun, F. Sohel, and R. Togneri. Learning deep structured network for weakly supervised change detection. In *IJCAI*, 2017.

[Liu and Do, 2017] S. Liu and M.N. Do. Inverse rendering and relighting from multiple color plus depth images. *IEEE TIP*, 26(10):4951–4961, 2017.

[Lu *et al.*, 2015] F. Lu, I. Sato, and Y. Sato. Uncalibrated photometric stereo based on elevation angle recovery from BRDF symmetry of isotropic materials. In *CVPR*, 2015.

[Miao *et al.*, 2018] D. Miao, F.P. Tian, and W. Feng. Active camera relocalization with RGBD camera from a single 2D image. In *ICASSP*, 2018.

[Midorikawa *et al.*, 2016] K. Midorikawa, T. Yamasaki, and K. Aizawa. Uncalibrated photometric stereo by stepwise optimization using principal components of isotropic brdfs. In *CVPR*, 2016.

[Ou and Pellacini, 2011] J. Ou and F. Pellacini. LightSlice: Matrix slice sampling for the many-lights problem. *ACM TOG*, 30(6):179:1–179:8, 2011.

[Ren *et al.*, 2013] P. Ren, J. Wang, M. Gong, S. Lin, X. Tong, and B. Guo. Global illumination with radiance regression functions. *ACM TOG*, 32(4):130, 2013.

[Ren *et al.*, 2015] P. Ren, Y. Dong, S. Lin, X. Tong, and B. Guo. Image based relighting using neural networks. In *ACM SIGGRAPH*, 2015.

[Shi *et al.*, 2016] B. Shi, Z. Wu, Z. Mo, D. Duan, S.-K. Yeung, and P. Tan. A benchmark dataset and evaluation for non-Lambertian and uncalibrated photometric stereo. In *CVPR*, 2016.

[Shi *et al.*, 2018] Y.B. Shi, F.P. Tian, D. Miao, and W. Feng. Fast and reliable computational rephotography on mobile device. In *ICME*, 2018.

[Stent *et al.*, 2016] S. Stent, R. Gherardi, B. Stenger, and R. Cipolla. Precise deterministic change detection for smooth surfaces. In *WACV*, 2016.

[Tom *et al.*, 2001] M. Tom, G. Dan, and W. Hans. Polynomial texture maps. In *ICME*, 2001.

[Wang *et al.*, 2003] Z. Wang, E.P. Simoncelli, and A.C. Bovik. Multiscale structural similarity for image quality assessment. In *IEEE Asilomar Conference on Signals, Systems and Computers*, 2003.

[Williams *et al.*, 2011] B. Williams, G. Klein, and I. Reid. Automatic relocalization and loop closing for real-time monocular slam. *IEEE TPAMI*, 33(9):1699–712, 2011.

[Xiong *et al.*, 2015] Y. Xiong, A. Chakrabarti, R. Basri, S.J. Gortler, D.W. Jacobs, and T. Zickler. From shading to local shape. *IEEE TPAMI*, 1(1):1, 2015.

[Zhao *et al.*, 2012] Q. Zhao, P. Tan, Q. Dai, L. Shen, E. Wu, and S. Lin. A closed-form solution to retinex with non-local texture constraints. *IEEE TPAMI*, 34(7):1437–1444, 2012.

[Zhao *et al.*, 2015] Q. Zhao, W. Feng, L. Wan, and J. Zhang. SPHORB: A fast and robust binary feature on the sphere. *IJCV*, 113(2):143–159, 2015.