CSIE: Coded Strip-patterns Image Enhancement Embedded in Structured Light-based Methods

Wei Cao[#], Yuping Ye[#], Chu Shi, and Zhan Song[✉]

Guangdong-Hong Kong-Macao Joint Laboratory of Human-Machine
Intelligence-Synergy Systems, Shenzhen Institute of Advanced Technology, Chinese
Academy of Sciences, Shenzhen, 518055, China
{w.cao,yp.ye,chu.shi,zhan.song}@siat.ac.cn

Abstract. When a coded strip-patterns image (CSI) is captured in a structured light system (SLs), it often suffers from low visibility at low exposure settings. Besides degrading the visual perception of the CSI, this poor quality also significantly affects the performance of 3D model reconstruction. Most of the existing image-enhanced methods, however, focus on processing natural images but not CSI. In this paper, we propose a novel and effective CSI enhancement (CSIE) method designed for SLs. More concretely, a bidirectional perceptual consistency (BPC) criterion, including relative grayscale (RG), exposure, and texture level priors, is first introduced to ensure visual consistency before and after enhancement. Then, constrained by BPC, the optimization function estimates solutions of illumination with piecewise smoothness and reflectance with detail preservation. With well-refined solutions, CSIE results can be achieved accordingly and further improve the details performance of 3D model reconstruction. Experiments on multiple sets of challenging CSI sequences show that our CSIE outperforms the existing used for natural image-enhanced methods in terms of 2D enhancement, point clouds extraction (at least 17 % improvement), and 3D model reconstruction.

1 Introduction

Undoubtedly, vision-based 3D reconstruction technology plays an essential role in the development of AR/VR innovations such as the metaverse. While reconstructing an object's 3D surface profiles in active or passive projection coding mode is fundamental and critical to numerous vision-based 3D reconstruction systems, such as binocular stereo vision [18], laser scanning [26], time-of-flight [6], and structured light systems (SLs) [24, 13] etc. Among those systems, SLs in active projection coding mode has gained wide applications in industrial and commercial fields due to their superior advantages, such as highly accurate and dense 3D reconstruction results. Complete SLs typically consists of a projector for projecting coded strip-patterns and a camera for capturing the coded

---

[#]These authors contributed equally to this work. [✉]Corresponding author.

strip-patterns image (CSI), respectively. Ideally, SLs usually use CSI and system calibration parameters to decode the depth information of the target under high exposure settings and then use the triangulation principle [12] to achieve accurate estimation of the 3D point coordinates on the surface of the projected object. However, in the high-speed reconstruction of structured light, the exposure time of a camera needs to be synchronized with that of the projector, i.e., a smaller exposure time setting can support reconstruction at a high frame rate. In other words, the extremely low exposure setting of a camera under high-frame-rate reconstruction makes the coded strips in the CSI bury in the dark and difficult to distinguish, further resulting in poor reconstructed 3D results, such as loss of local details and decrease in accuracy. Obviously, the underexposure problem on the CSI introduced at low exposure settings not only reduces the coded strips' discriminability but also seriously affects the 3D reconstruction performance of object surfaces.

It is worth noting that the existing enhancement methods mainly including three categories, i.e., model-free methods [21, 1, 4], model-based methods [7, 8, 10, 25, 2, 22, 32, 11, 28], and deep learning-based methods [31, 19, 30, 5], have been developed to address underexposure problem in natural images. However, to the best of our knowledge, there is almost no enhancement method specifically designed to address the underexposure problem in CSI under fixed exposure settings within SLs. Specifically, model-free methods are generally simple and effective, but will wash out details and cause oversaturation problems. Model-based methods mostly work according to the simplified Retinex theory assumption that a captured image P can be decomposed into reflectance component R and illumination component L by $P = R \circ L$, where $\circ$ denotes pixel-wise multiplication. Although such methods have been shown to have impressive enhancement effects in recent years, they are typically accompanied by overexposure issues in local regions. Deep learning-based methods rely on complex network structures and synthetic samples, which will increase high hardware resources and time costs in training and will reduce the generalizing ability of multi-scenario applications in terms of the network. More importantly, the methods mentioned above serve natural images, not CSI. Therefore, the core idea of most these methods, although not all, does not consider CSI's inherent properties.

In this paper, we propose a novel method, i.e., CSIE, for underexposed enhancement of CSI in SLs under a fixed low exposure setting. Our work is mainly inspired by an interesting observation: maintaining the visual consistency before and after enhancement can effectively avoid the disruption of regular coded strip patterns within CSI. Inspired by this observation, a valid criterion, bidirectional perceptual consistency (BPC), is first proposed to describe how to guarantee visual perceptual consistency between before and after enhancement. Then, we turn the underexposed enhancement problem into an optimization problem based on the Retinex model for decomposing the illumination and reflectance components under the constraints of the BPC priors. With well-refined illumination and reflectance components, the enhanced CSI results can be obtained accordingly, and further embedded in SLs methods can reconstruct more

3D profiles of objects' surfaces. In particular, the work in this paper has the following **contributions:** 1) We propose a novel CSIE method. To the best of our knowledge, this is the first time that CSIE is embedded in SLs methods to address underexposure problems in CSI under fixed exposure settings. 2) We define BPC (a simple and intuitive criterion including three consistency priors) in CSIE, which explicitly describes how to avoid breaking the visual perceptual consistency that exists in the CSI before and after enhancement. 3) We design BPC-constrained illumination and reflectance estimation via a variational optimization decomposition function, to achieve enhanced CSI results with high visibility and high signal-to-noise ratios. 4) We introduce a block coordinate descent (BCD) [27] technique to solve the convex problem contained in the designing optimization function. Finally, we evaluate the proposed method that includes a comparison with several state-of-the-art image enhancement methods on both multiple CSI sequences captured by SLs and the 3D reconstruction method of [24]. Experiments demonstrate that the proposed method achieves well enhancement effect in enhancing the visual quality for 2D CSI and simultaneously improves the performance of 3D surface reconstruction.

2   Related Work

Model-free Methods. Directly linearly amplifying the intensity of an underexposed image is probably the most straightforward way to recover detail in dark regions. However, this operation comes with a thorny problem, i.e., the pixels in the relatively bright regions may increase the risk of saturation and thus wash out corresponding details. Although the histogram equalization (HE) method can alleviate the above problems by forcing the enhanced intensities of input in a normalized manner, it may lead to noise amplification and distortion appearance. To be more wise, an improved version of HE, content and variational contrast (CVC) [4] enhancement, attempts to improve contrast by introducing different regularization terms on the histogram. It may very likely fall into the trouble of overexposure/underexposure enhancement due to ignoring the contrast change caused by real illumination.

Model-based Methods. Jobson et al. [14,15] did early works based on the Retinex model yet exhibited more unrealistic results in the output. Fu et al. [7] proposed a simultaneous illumination and reflectance estimation (SIRE) method to realize enhancement in logarithmic space, but the reflectance component in it has the problem of detail loss. Thereafter, Fu et al. [8] developed a weighted variational model (WVM) to improve the variation of gradient magnitudes in bright regions. Since the anisotropic smoothness assumption is ignored, the illumination decomposed by the model may fall into erroneous estimation around non-contiguous regions. Guo et al. [10] estimated the initial illumination map from the input and refined it by retaining only the main structural information. This method does well in dealing with underexposed images with globally uniform illumination, however, produces an overexposure issue when processing ones with non-uniform illumination. To this end, Zhang et al. [32]

designed the high-quality exposure correction (HQEC) method, which estimates reasonable illumination to solve the above problem by considering the visual similarity between input and output. The limitation of this method is the blindness of the noise, leading to results with a low SNR in underexposed situations. Considering this limitation, Cai et al. [2] and Hao et al. [11] presented the joint intrinsic-extrinsic prior (JieP) model and the gaussian total variation (GTV) model, respectively, to achieve a balance between visual enhancement and noise compression. However, both optimized models suffer from over-smoothing in the illumination and reflectance, and the latter's piecewise smoothness constraint on the reflectance component makes the model ineffective for enhancing images with extremely low exposure. Different from JieP, Ren et al. [22] proposed the joint denoising and enhancement (JED) method by considering an additional detail enhancement term on the reflectance component, but it also induces local distortion on that. In addition, Xu et al. [28] proposed a structure and texture aware retinex (STAR) model in recent years, which designed an exponential filter through local derivatives and used it to constraint for illumination and reflectance components, respectively. Because of the constraint of local partial derivatives, the coded structures of underexposed CSI face destroyed problems. It is worth noting that Song et al. [25] developed the HDR-based method to improve CSI for alleviating the local underexposure and overexposure problems caused by the high reflectivity of the object surface in the CSI. Different from CSIs (captured with a fixed exposure setting) used in our method, the HDR-based method [25] obtains the solution that requires at least three CSIs (captured with different exposure settings by the same camera under the same scene).

Deep Learning-based Methods. Chen et al. [5] proposed a low-light enhancement method with low photon counts and low SNRs based on the fully convolutional networks in an end-to-end training way. However, the input of network training is raw sensor data, which is difficult to obtain and causes a high training cost. Lore et al. [19] developed a depth encoder to enhance contrast, but its enhanced results only have well performance in noise reduction. Inspired by neurobiological observations, Zhan et al. [31] designed a feature-linking model (FLM) by utilizing the spike times of encoded information, which can preserve information and enhance image detail information simultaneously. Afterward, Zhan et al. [30] also presented a linking synaptic computation network (LSCN) that generates detail-enhanced results, but its noise amplifying problem is obvious. However, almost none of the above methods are suitable for CSI enhancement, mainly due to their enhanced networks being designed for natural images and not considering the regular encoded information in CSI.

3   Proposed Method

In this section, we describe the proposed novel CSIE method for enhancing CSI. We first introduce one Retinex-based model decomposition problem and then restraint it by defining the BPC criterion, including three consistency priors (RG, exposure, texture). Jointing the above priors, we further proposed an opti-

mization function that is designed for decomposing illumination and reflectance components in CSI, respectively. Finally, we use an efficient BCD technique to extract solutions for the convex optimization function iteratively. With well-refined components in solutions, the enhanced CSI can be obtained accordingly.

3.1   Model Decomposition Problem

Let us recall the Retinex model, a commonly used visual model to estimate the illumination and reflectance components in real scenes for image enhancement. Unlike previous enhancement methods based on this model, we apply it here to the processing of CSI instead of natural images and combine the inherent priors in CSI to formulate the image decomposition problem as the minimization of the following optimization function $F(\cdot, \cdot)$:

$$\min_{R,L} F(R, L) = ||R \circ L - P||_2^2 + \lambda_g f_g + \lambda_e f_e + \lambda_r f_r, \tag{1}$$

where f_e, f_g and f_r are three different regularization functions, λ_e, λ_g and λ_r are all non-negative balancing weights. $||R \circ L - P||_2^2$, data fidelity term, constraints the product of the illumination and reflectance components to be close to the original CSI.

In general, decomposing an image into illumination and reflectance components is a highly ill-posed problem due to lack of reasonable priors on the components. Appropriate component priors as constraints can reduce the solution space and converge quickly. To this end, we describe the introduced priors about CSI in Eqn. (1) in detail below.
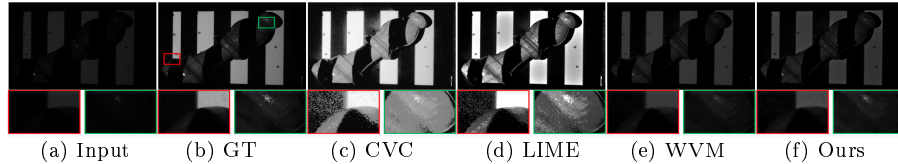

(a) Input   (b) GT   (c) CVC   (d) LIME   (e) WVM   (f) Ours

Fig. 1. Issues existed by previous methods. GT: Ground truth (CSI captured with 2000 ms of exposure value).

3.2   Bidirectional Perceptually Consistency

By observing large numbers of CSI sequences pairs between underexposed and well-exposed, we find that inherent properties (RG (defined as the relative grayscale change of neighboring pixels in the CSI), exposure, texture) existed in under-exposed CSI are buried in the dark, which results in SLs method that rely on strip-edge features cannot effectively identify and further significantly affects the 3D reconstructed performance. Compared with the well-exposed CSI (see Fig. 1 (b)), the distortions of relative grayscale, inconsistency of exposure, and loss of details produced by existing methods on the CSI enhancement results are the three main issues. For instance, coded strip features (regular bright-dark order in spatial position) are distorted in the enhanced result by the CVC method due

to the damage of the original relative grayscale, as shown in Fig. 1 (c). Texture degradation and loss of details due to local overexposure appear on the LIME method enhancement results, as shown in Fig. 1 (d). The WVM enhancement results have a local over-dark problem caused by inconsistent exposure, while these regions are contrary to the consistent exposure in the input, as shown in Fig. 1 (e). We intuitively believe that the above three issues can well be removed when we use inherent properties to preserve visual consistency before and after enhancement in CSI, as shown in Fig. 1 (f). Additionally, we analyze their plausibility by defining the following three consistency priors (RG, exposure, and texture) in BPC and then verify the effectiveness of those in BPC in Section 4.

RG Consistency Prior. One fact is that RG change in CSI is caused by the coded projection, which is an inherent feature belonging to CSI. For instance, we show a synthetic patch with the simulated CSI in Fig. 2 (a), which contains weak textures and strong edges/structures. The RG change, such as the relative grayscale change in the red and blue boxes in Fig. 2 (a), enables the strip-edge at the discontinuous region can be identified by SLs-based method [24] to reconstruct the profile of target surface, which can be regarded as an important prior feature. Intuitively, RG inconsistency, such as smaller or no discontinuity differences, will weaken the discrimination of the strip-edge features on the enhanced CSI, thereby reducing the ability to recover depth information for the SLs-based method. To avoid this issue, we take the preprocessed original CSI (see Fig. 2 (b)) as a prior constraint to ensure global consistency (see Figs. 2 (b) and (c)) between before and after enhancement on the relative grayscale. Mathematically, we express the regularization term as:

$$f_g = ||L - G||_2^2, \tag{2}$$

where $G=wgif(P)$ is the output (see Fig. 2 (b)) smoothed by a weighted guided image filtering [17] operation $wgif(\cdot)$. This regularization term takes into account the balance between RG preservation and noise suppression. Therefore, G as a prior can not only be used to maintain the consistency of RG in the illumination smoothing process but also be regarded as the initial value of illumination to guide its smoothing.
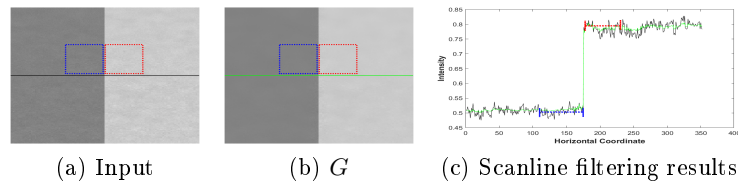

| (a) Input | (b) G | (c) Scanline filtering results |

Fig. 2. Illustration of RG changing.

Exposure Consistency Prior. Exposure consistency emphasizes a reasonable estimation of the illumination distribution. This distribution usually can be understood as piecewise smoothing [7, 8, 10], which follows the illumination distribution on images captured from the real scene and can help to recover clear

details from underexposed regions in CSI. It should be noted that the different existing norms, such as L_1 and L_2, directly employed to constraint illumination will ignore the specificity of CSI, i.e., the structure is buried in the dark or says blind. To alleviate this issue, we use L_γ norm [20] to constrain the illumination gradient to achieve piecewise smoothing and further estimate the illumination with preserved overall structures. Hence, we provide one spatial smoothness regularization term:

$$f_e = ||\nabla L||_\gamma^\gamma, \tag{3}$$

where $|| \cdot ||_\gamma^\gamma$ stands for L_γ norm with $0 \leq \gamma \leq 2$. ∇ is the first-order differential operator contains two directions: ∇_h (horizontal) and ∇_v (vertical). Since the L_γ norm in Eqn. (3) will cause a non-smooth optimization problem, a simple numerical approximation is introduced [2, 11, 20, 3] and Eqn. (3) can be rewritten as:

$$f_e \approx w||\nabla L||_2^2, \tag{4}$$

where $||\nabla L||_\gamma^\gamma \approx w||\nabla L||_2^2$ and

$$w \approx \begin{cases} (|\nabla L| + \varepsilon_L)^{\gamma-2}, & 0 < \gamma \leq 2 \\ \left. \begin{cases} \phi^{-2} & \text{if } |\nabla L| < \phi \\ |\nabla L|^{-2} & \text{otherwise} \end{cases} \right\}, & \gamma = 0 \end{cases}, \tag{5}$$

ϕ is a small constant (typically $\frac{1}{8}$) used for determining gradient sparsity of illumination. The constant ε_L is a small value (typically 0.0001) to avoid the zero denominator. For approximate operations, please refer to [20, 29]. Obviously, this approximation operation, on the one hand, changes the non-smooth L_γ term into L_2 one. On the other hand, it makes the weighted L_2 norm have both abilities that the piecewise smoothness and the L_0 norm pursuit (gradient sparsity), as shown in Fig. 3.
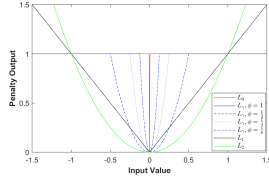

Fig. 3. Distribution of different penalty functions.

Texture Consistency Prior. Existing researches [9, 23] regard the decomposed R as the part that can remain unchanged under illumination variations because it contains inherent properties of the object surface, namely edges. Consistent properties are also included in edges of the captured CSI. So, texture consistency needs to consider the consistency of edge degradation between reflectance component and input before and after the CSI enhancement, which can also be called reflectance gradient degradation prior. In addition, edges, as dependency features that need to be identified in the 3D model reconstruction

method [24], should be preserved in the output after CSI enhancement to the greatest extent possible. To this end, we project edges of the noise-reduced original CSI P onto ones of R and express the gradient fidelity term as follows:

$$f_r = ||\nabla R - M||_2^2, \tag{6}$$

where

$$M(\nabla P; \varepsilon_m) := \begin{cases} 0, & |\nabla P| < \varepsilon_m \\ \nabla P, & \text{otherwise}, \end{cases} \tag{7}$$

is a truncated function that only penalizes gradients whose magnitudes are smaller than the small threshold ε_m (typically 1e-5), as shown in Fig. 4. ε_m is used for determining whether there is an edge at a pixel in the CSI. In other words, small ε_m can better preserve textures, and vice versa. So, this regularization term separates the reflectance component (with noise-suppressed and detail-preserved) from CSI by removing small non-zero gradients, i.e., flattening tiny textures/noises.
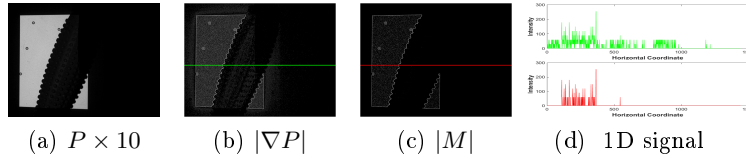

(a) $P \times 10$ (b) $|\nabla P|$ (c) $|M|$ (d) 1D signal

Fig. 4. Illustration of truncated ability in M. For observation easily, the intensity of original CSI P is linearly scaled by 10.

3.3   Exact Solver to Problem (1)

Although the illumination smoothness term constrained by L_γ norm is non-smooth, the results of numerically approximated operation act as a tight surrogate to change the non-smooth L_γ term into a piecewise smooth L_2 one. By doing so, the optimization function in Eqn. (1) can be rewritten as the following equivalent form:

$$\min_{R,L} F(R, L) = ||R \circ L - P||_2^2 + \lambda_e w ||\nabla L||_2^2 + \lambda_g ||L - G||_2^2 + \lambda_r ||\nabla R - M||_2^2. \tag{8}$$

One can see that the optimization function in Eqn. (8) can be regarded as a problem that only contains two variables (R and L) to solve. Block coordinate descent (BCD) [27] technique is a common choice to solve this problem. To facilitate the solution and analysis, we further divide the objective function in Eqn. (8) into two sub-problems corresponding to R and L as follows:
L **sub-problem:** Collecting the L involved terms from Eqn. (8) gives the following problem:

$$L_{k+1} = \arg\min_{L} ||R \circ L - P||_2^2 + \lambda_e w ||\nabla L||_2^2 + \lambda_g ||L - G||_2^2. \tag{9}$$

R **sub-problem:** With L_{k+1} acquired from the above solution, the minimization corresponding to R in Eqn. (8) can be formulated as the following optimization problem:

$$R_{k+1} = \arg\min_R ||R \circ L_{k+1} - P||_2^2 + \lambda_r ||\nabla R - M||_2^2. \tag{10}$$

As can be seen from the two sub-problems above, they are both minimization problems containing only classical least-squares terms and therefore have closed-form global optimal solutions. The solutions are detailed as follows:

a) Solution for (T1): The sub-problem in Eqn. (9) can be reformulated in matrix notation as:

$$(\mathbf{R} \circ \mathbf{L} - \mathbf{P})^T (\mathbf{R} \circ \mathbf{L} - \mathbf{P}) + \lambda_g (\mathbf{L} - \mathbf{G})^T (\mathbf{L} - \mathbf{G}) + \lambda_e (\mathbf{L}^T \mathbf{D}_h^T \mathbf{W}_h \mathbf{D}_h \mathbf{L} + \mathbf{L}^T \mathbf{D}_v^T \mathbf{W}_v \mathbf{D}_v \mathbf{L}), \tag{11}$$

where $\mathbf{D}_h$ and $\mathbf{D}_v$ are Toeplitz matrices from discrete gradient operators with forward difference, respectively. $\mathbf{W}_h$ and $\mathbf{W}_v$ denote diagonal matrices containing weights w_h and w_v, respectively. The unique solution of the variable $\mathbf{L}$ in Eqn. (11) can be easily obtained by performing linear system operation like:

$$\mathbf{L}_{k+1} = (\mathbf{R}_k^T \mathbf{R}_k + \lambda_g \mathbf{1} + \lambda_e \bar{\mathbf{W}}_k)^{-1} (\mathbf{R}_k^T \mathbf{P} + \lambda_g \mathbf{G}), \tag{12}$$

where $\mathbf{1}$ is an identity matrix and $\mathbf{D}_h^T \mathbf{W}_h \mathbf{D}_h + \mathbf{D}_v^T \mathbf{W}_v \mathbf{D}_v$ represents a symmetric positive definite Laplacian matrix [16].

b) Solution for (T2): Similar to the solution of *(T1)*, the update of the closed-form solution R_{k+1} can be directly obtained by the following operations:

$$\mathbf{R}_{k+1} = (\mathbf{L}_{k+1}^T \mathbf{L}_{k+1} + \lambda_r \bar{\mathbf{V}}_k)^{-1} (\mathbf{L}_{k+1}^T \mathbf{P} + \lambda_r \bar{\mathbf{M}}), \tag{13}$$

where $\bar{\mathbf{V}} = \mathbf{D}_h^T \mathbf{D}_h + \mathbf{D}_v^T \mathbf{D}_v$. Note that $\bar{\mathbf{M}} = \text{vec}^{-1}(\mathbf{D}_h^T \dot{\mathbf{M}}_h + \mathbf{D}_v^T \dot{\mathbf{M}}_v)$, where $\dot{\mathbf{M}}_h$ and $\dot{\mathbf{M}}_v$ are the vector versions of the corresponding $\mathbf{M}_h$ and $\mathbf{M}_v$, respectively, while $\text{vec}^{-1}(\cdot)$ represents the inverse vectorization operator for reshaping vectors back to their matrix format.

The above iterative estimation is repeated until the convergence conditions $||\mathbf{L}_{k+1} - \mathbf{L}_k||/\mathbf{L}_k \leq \varpi$ and $||\mathbf{R}_{k+1} - \mathbf{R}_k||/\mathbf{R}_k \leq \varpi$ are satisfied or the maximum number of iterations exceeds a preset constant K. The whole iterative optimization process is summarized as Algorithm 1. To further improve the visibility of images, the final enhanced CSI $\mathbf{P}_f$ is achieved by projecting the adjusted illumination back to the reflectance, i.e., $\mathbf{P}_f = \mathbf{R} \circ \mathbf{L}_f$, where $\mathbf{L}_f$ is gamma-corrected (empirically set as 2.2) illumination estimation.

4   Experiments

In this section, we run all enhancement experiments using Matlab R2021a on a laptop with Windows 11 OS, 16G RAM and Intel Core i7-2.3 GHz CPU. In our experiments, the parameters λ_e, λ_g, λ_r, ϖ, and γ (in L_γ norm) are empirically

---

Algorithm 1 Exact Solver to Optimization Problem (8)

Require: The weight coefficients λ_e, λ_g, λ_r; the original CSI **P**, initial illumination map **G**, and truncated result **M**; the positive parameter ϖ and the maximum number of iterations K.

Ensure: Optimal solutions $\mathbf{R}_{k+1}$ and $\mathbf{L}_{k+1}$.

1: initial $\mathbf{L}_0 \leftarrow \mathbf{P}$, $\varpi \leftarrow 0.001$, $k \leftarrow 0$, and $K \leftarrow 20$;
2: **repeat**
3:     Update $\mathbf{L}_{k+1}$ using Eqn. (12);
4:     Update $\mathbf{R}_{k+1}$ with $\mathbf{L}_{k+1}$ using Eqn. (13);
5:     $k \leftarrow k + 1$;
6: **until** ($||\mathbf{L}_{k+1} - \mathbf{L}_k||/\mathbf{L}_k \leq \varpi$ and $||\mathbf{R}_{k+1} - \mathbf{R}_k||/\mathbf{R}_k \leq \varpi$) or $k > K$.

---

set to 0.01, 0.15, 5, 0.001, and 0.6 respectively. In addition, the experimental setup in the SLs' hardware configured with an off-the-shelf projector (TI3010, with resolution 1280×720 pixels, ≤2800fps) and a camera (BFS-U3-16S2M, with resolution 1440×1080 pixels, USB interface) with 50ms of exposure time, as shown in Fig. 5 (a). Meanwhile, we use the 3D model reconstruction method [24] (for showing the reconstructed 3D model performance of enhanced CSI) and its coding strategy (see Fig. 5 (b)), and implement it by writing C++ in the VS2019 environment. To fully evaluate the proposed method, we test 16 sets CSI sequences of object surfaces with different reflectance under a fixed low exposure setting, such as samples in Fig. 5 (c). Then, we compare proposed CSIE subjectively and objectively with existing state-of-the-art methods, including CVC [4], SIRE [7], WVM [8], LIME [10], JieP [2], JED [22], HQEC [32], GTV [11], STAR[28], and LSCN [30].
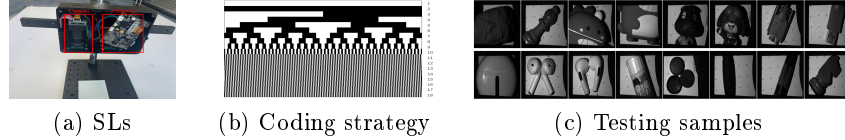

(a) SLs   (b) Coding strategy   (c) Testing samples

Fig. 5. Experimental configuration. (a) The SLs consists of an off-the-shelf projector and a camera, and the two devices are synchronized. (b) Coding strategy (with 18 pattern images to be projected) of Gray code combined with binary shifting strip with widths of 4 pixels [24]. (c) Samples from 16 sets of the captured CSI sequences.

Retinex decomposition. One can observe that visual artifacts appear in the illumination and reflectance components on the SIRE simultaneously, such as the edges of the frog's eyes in Fig. 6 (b). These artifacts are generated due to an unreasonable assumption that illumination is with unconstrained isotropic smoothness. Furthermore, JieP, JED, and STAR all produce strong artifact and over-smoothing problems on the reflectance component, such as the frog's hair and mouth in Figs. 6 (c), (d), and (e). For these visual distortions in the decomposed components, in essence, it is because those methods reduce the ability to discriminate structures-textures at very low exposures. In contrast, the illumination component constrained by the L_γ norm in our method to be as far as

possible both removed most of the tiny textures and preserved the prominent structures. Meanwhile, the reflectance component constrained by the L_2 norm in our method to maintain both more textures and details of high-frequency, such as the frog's face in Fig. 6 (f). More importantly, unlike SIRE, thanks to the suppression of noise factors, our model can effectively avoid noise amplification on the reflectance component, which is particularly suitable for recovering the details of underexposed CSI.
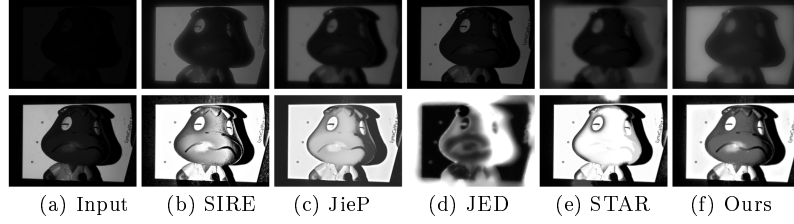

(a) Input   (b) SIRE   (c) JieP   (d) JED   (e) STAR   (f) Ours

Fig. 6. Comparison of several Retinex decomposition results. (a) Input and corresponding ten times intensity levels (used for observation). **Top Row in (b) to (f):** Illumination component. **Bottom Row in (b) to (f):** Reflectance component.
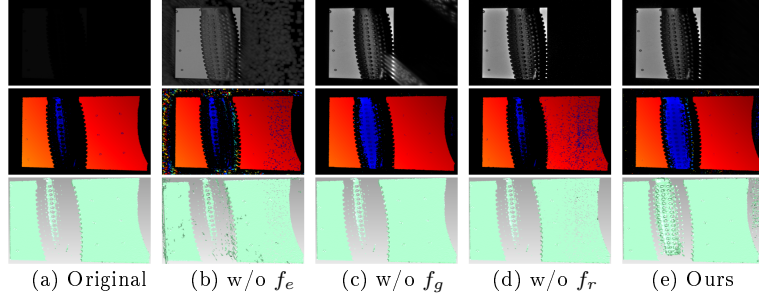

(a) Original   (b) w/o f_e (c) w/o f_g (d) w/o f_r (e) Ours

Fig. 7. Ablation study of different prior on 3D reconstruction performance of the proposed CSIE. (a) The images from the first to the last rows are original CSI, depth maps, and 3D point clouds, respectively. From (b) to (e): The enhanced versions corresponding to (a).

3D Impact of model with different prior. Since RG, exposure, and texture consistency priors have not been adopted in previous Retinex algorithms, we analyze their impacts in the following experiments. Here, we set $\lambda_e = 0$, or $\lambda_g = 0$, or $\lambda_r = 0$ in (8), respectively, and update them according to Eqns. 9 to 13 describe, and thus have three baselines: CSIE w/o f_e, CSIE w/o f_g and CSIE w/o f_r. From Fig. 7, one can see that CSIE w/o f_e tends to generate visual artifacts on the coded strip at the right black that should not be appeared on the enhanced CSI results. This problem will lead to lots of holes appearing on 2D depth maps and 3D point clouds simultaneously. The bright streaks of CSIE w/o f_g on the enhanced CSI results destroy the coded information at the depth discontinuities, causing the details loss of the screwdriver's local profile on the 2D depth map and 3D point clouds, respectively. However, the increase in point clouds numbers of both demonstrates that the two regularization terms

have a positive impact on improving the 3D reconstruction performance of the object surface. For CSIE w/o f_r, the local textures lost in the enhanced CSI results directly lead to the emergence of large numbers of black holes in the 2D depth image, which seriously affects the numbers of 3D point clouds extraction. Obviously, the prior f_r is far more important than the former two for CSI enhancement and 3D reconstruction. By considering three, the proposed CSIE obtains output with pleasing visibility and undistorting coded strip, which not only satisfies the visual expectations (consistency between RG, exposure, and texture) on the enhanced CSI results but also reasonably recovers more useful depth details buried in the dark.

Convergence Behavior. Analyzing the convergence rate of CSIE with different prior from Fig. 8, one can observe that different components with joint priors are faster in convergence rate than ones without λ_e, or λ_g, or λ_r. This is mainly due to the BPC criterion applying the different priors' constraints to reduce the solution space of the objective function. Besides, one can see that the iterative error ϖ for different components with arbitrarily different priors reduces to less than 0.001 in 20 iterations. This favorable convergence performance benefits from the proposed CSIE exact solver converging to global optimal solutions for the problem (8).
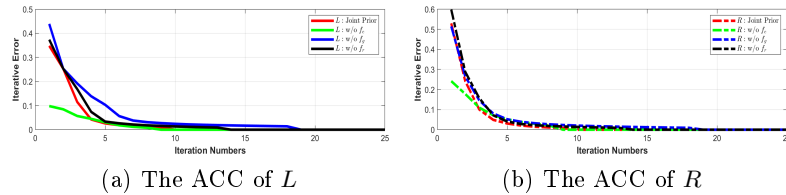

(a) The ACC of L (b) The ACC of R

Fig. 8. Convergence rate of our model with different prior. ACC: Average convergence curves.

Subjective Comparison. For comparison fairness, we use the CSI sequence with 2000 ms of exposure time as the reference (ground truth) instead of the one with 50 ms of that. Fig. 9 shows several visual comparisons. From which, we can find that WVM, HQEC, STAR, and GTV cannot effectively recall the information in the dark region on the original CSI, which significantly affects the depth information extraction and the 3D model reconstruction, such as the bells in Fig. 9 (a). This problem is almost always generated, especially on GTV results. Although the enhanced CSI results of CVC and LSCN exhibit higher visibility in the given cases, the coded features with relatively regular bright-black strips present on them are destroyed. While, unreasonable coded strips indirectly cause both to produce invalid depth and point cloud maps, as shown in the second and three rows in Figs. 9 (a), (b) and (c). As for LIME, it recovers more obvious and high-contrast details, however, its depth and point cloud information will be lost due to amplified noises and overexposed textures on the enhanced CSI results' local regions. In contrast, the enhanced CSI results by CSIE are closer to the ground-truth CSI than other methods, as shown in Figs. 9 (a), (b) and (c). Besides, our method obtains a balanced enhanced performance between depth

estimation, point clouds extraction, and **3D** model reconstruction benefit from the reasonable consideration of BPC introduced for CSIE.
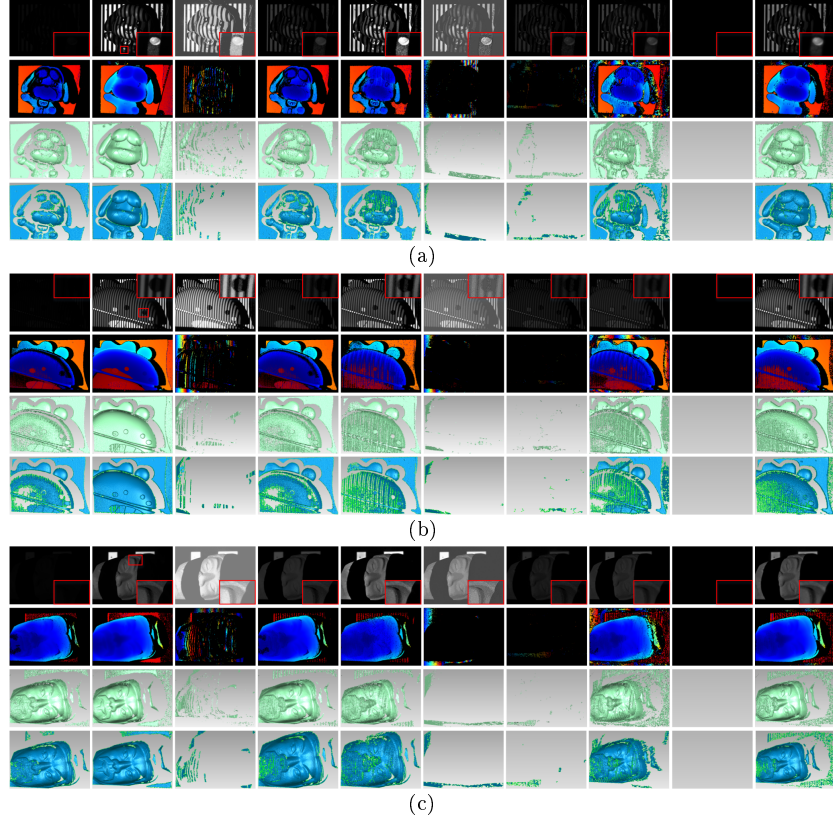

Fig. 9. Result comparison of different methods on CSI. **From Left to Right in Three Cases:** Original CSI, Ground truth, CVC, WVM, LIME, LSCN, HQEC, STAR, GTV, and Ours, respectively. **From Top Row to Bottom Row in Three Cases:** Enhanced results by different methods, corresponding 2D depth map, 3D point clouds, and 3D mesh map, respectively.

Objectively Comparison. Here, we apply four indicators (peak signal-to-noise ratio (PSNR), structural similarity index measure (SSIM), root mean square error (RMSE), and number of points (NPs)) to evaluate the performance of our CSIE objectively. As we can see, our CSIE method performs favorably against existing image enhancement methods adapted to natural images, which shows that the proposed CSIE method has superior performance in terms of structure preservation and noise compression, as shown in Table 1. Additionally, we note that STAR has higher NPs (see Fig. 10), but its other measurement results (with smaller PSNR and SSIM values and larger RMSE values) are less than the performance of our proposed method. It is worth noting that our method ranks second on NPs, but the NPs extracted by our method have better

3D reconstruction performance than the ones extracted by STAR, as shown in Figs. 9 (a), (b) and (c). The main reason is that using local partial derivatives in STAR generates numerous invalid NPs, which in turn recovers more erroneous 3D surface details that do not belong to the object. Obviously, the CSIE method achieves not only higher PSNR and SSIM values but also lower RMSE values in contrast to mentioned above methods, which is highly correlated with its better decomposition of the illumination and reflectance components on the CSI. Overall, the CSIE method not only has a better comprehensive performance compared to other methods, but its results are at least 17% improvements better than the original CSI sequences in terms of NPs extraction.

Table 1. Quantitative measurement results on 16 sets of CSI sequences with a fixed exposure value (50 ms) in terms of average PSNR, SSIM, and RMSE.

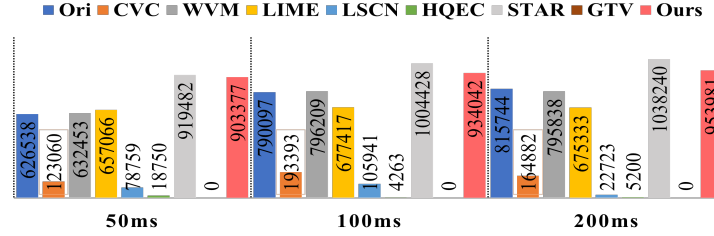| Metrics	Original	CVC	WVM	LIME	LSCN	HQEC	STAR	GTV	Ours
PSNR	19.38	15.89	23.55	22.07	19.76	21.82	23.44	19.23	**25.23**
SSIM	0.3969	0.3152	0.7167	0.5602	0.0931	0.6261	0.7858	0.3807	**0.8098**
RMSE	0.1839	0.3979	0.1089	0.1133	0.3271	0.1317	0.1056	0.1817	**0.1009**

Fig. 10. Quantitative comparison results on 16 sets of CSI sequences with different exposure values (50 ms, 100 ms, and 200 ms) in terms of average NPs.

5 Conclusion

In this paper, we propose a novel CSIE method implemented in a variational optimization decomposition, which can be used for CSI enhancement and further improve the performance of the existing 3D model reconstruction method based on SLs. The proposed CSIE is based on a visually bidirectional perceptual consistency criterion, including RG, exposure, and texture consistency priors. Combined with all priors, the proposed CSIE can simultaneously obtain the illumination component with piecewise smoothness and the reflectance component with detail preservation on the enhanced CSI under a low exposure setting. Experimental results on 2D enhancement and 3D model reconstruction demonstrate that our CSIE outperforms other existing enhancement methods adapted to the natural image in terms of visual enhancement and coded strip feature preservation.

References

1. Abdullah-Al-Wadud, M., Kabir, M.H., Dewan, M.A.A., Chae, O.: A dynamic histogram equalization for image contrast enhancement. IEEE Transactions on Consumer Electronics **53**(2), 593–600 (2007)
2. Cai, B., Xu, X., Guo, K., Jia, K., Hu, B., Tao, D.: A joint intrinsic-extrinsic prior model for retinex. In: IEEE International Conference on Computer Vision. pp. 4000–4009 (2017)
3. Cao, W., Wu, S., Wu, J., Liu, Z., Li, Y.: Edge/structure-preserving texture filter via relative bilateral filtering with a conditional constraint. IEEE Signal Processing Letters **28**, 1535–1539 (2021)
4. Celik, T., Tjahjadi, T.: Contextual and variational contrast enhancement. IEEE Transactions on Image Processing **20**(12), 3431–3441 (2011)
5. Chen, C., Chen, Q., Xu, J., Koltun, V.: Learning to see in the dark. In: IEEE Conference on Computer Vision and Pattern Recognition. pp. 3291–3300 (2018)
6. Cho, H., Kim, S.W.: Mobile robot localization using biased chirp-spread-spectrum ranging. IEEE Transactions on Industrial Electronics **57**(8), 2826–2835 (2009)
7. Fu, X., Liao, Y., Zeng, D., Huang, Y., Zhang, X.P., Ding, X.: A probabilistic method for image enhancement with simultaneous illumination and reflectance estimation. IEEE Transactions on Image Processing **24**(12), 4965–4977 (2015)
8. Fu, X., Zeng, D., Huang, Y., Zhang, X.P., Ding, X.: A weighted variational model for simultaneous reflectance and illumination estimation. In: IEEE Conference on Computer Vision and Pattern Recognition. pp. 2782–2790 (2016)
9. Grosse, R., Johnson, M.K., Adelson, E.H., Freeman, W.T.: Ground truth dataset and baseline evaluations for intrinsic image algorithms. In: IEEE International Conference on Computer Vision. pp. 2335–2342 (2009)
10. Guo, X., Li, Y., Ling, H.: LIME: Low-light image enhancement via illumination map estimation. IEEE Transactions on Image Processing **26**(2), 982–993 (2016)
11. Hao, S., Han, X., Guo, Y., Xu, X., Wang, M.: Low-light image enhancement with semi-decoupled decomposition. IEEE Transactions on Multimedia **1**(1), 1–14 (2020)
12. Hartley, R., Zisserman, A.: Multiple view geometry in computer vision. Cambridge University Press (2003)
13. Van der Jeught, S., Dirckx, J.J.: Real-time structured light profilometry: a review. Optics and Lasers in Engineering **87**, 18–31 (2016)
14. Jobson, D.J., Rahman, Z.u., Woodell, G.A.: A multiscale retinex for bridging the gap between color images and the human observation of scenes. IEEE Transactions on Image Processing **6**(7), 965–976 (1997)
15. Jobson, D.J., Rahman, Z.u., Woodell, G.A.: Properties and performance of a center/surround retinex. IEEE transactions on image processing **6**(3), 451–462 (1997)
16. Krishnan, D., Fattal, R., Szeliski, R.: Efficient preconditioning of laplacian matrices for computer graphics. ACM Transactions on Graphics **32**(4), 1–15 (2013)
17. Li, Z., Zheng, J., Zhu, Z., Yao, W., Wu, S.: Weighted guided image filtering. IEEE Transactions on Image processing **24**(1), 120–129 (2014)
18. Liu, Z., Yin, Y., Wu, Q., Li, X., Zhang, G.: On-site calibration method for outdoor binocular stereo vision sensors. Optics and Lasers in Engineering **86**, 75–82 (2016)
19. Lore, K.G., Akintayo, A., Sarkar, S.: Llnet: A deep autoencoder approach to natural low-light image enhancement. Pattern Recognition **61**(0), 650–662 (2017)
20. Min, D., Choi, S., Lu, J., Ham, B., Sohn, K., Do, M.N.: Fast global image smoothing based on weighted least squares. IEEE Transactions on Image Processing **23**(12), 5638–5653 (2014)

21. Pisano, E.D., Zong, S., Hemminger, B.M., DeLuca, M., Johnston, R.E., Muller, K., Braeuning, M.P., Pizer, S.M.: Contrast limited adaptive histogram equalization image processing to improve the detection of simulated spiculations in dense mammograms. Journal of Digital imaging **11**(4), 193–200 (1998)
22. Ren, X., Li, M., Cheng, W.H., Liu, J.: Joint enhancement and denoising method via sequential decomposition. In: IEEE International Symposium on Circuits and Systems. pp. 1–5 (2018)
23. Rother, C., Kiefel, M., Zhang, L., Schölkopf, B., Gehler, P.: Recovering intrinsic images with a global sparsity prior on reflectance. Advances in Neural Information Processing Systems **24** (2011)
24. Song, Z., Chung, R., Zhang, X.T.: An accurate and robust strip-edge-based structured light means for shiny surface micromeasurement in 3-d. IEEE Transactions on Industrial Electronics **60**(3), 1023–1032 (2012)
25. Song, Z., Jiang, H., Lin, H., Tang, S.: A high dynamic range structured light means for the 3d measurement of specular surface. Optics and Lasers in Engineering **95**, 8–16 (2017)
26. Sun, Q., Chen, J., Li, C.: A robust method to extract a laser stripe centre based on grey level moment. Optics and Lasers in Engineering **67**, 122–127 (2015)
27. Tseng, P.: Convergence of a block coordinate descent method for nondifferentiable minimization. Journal of Optimization Theory and Applications **109**(3), 475–494 (2001)
28. Xu, J., Hou, Y., Ren, D., Liu, L., Zhu, F., Yu, M., Wang, H., Shao, L.: Star: A structure and texture aware retinex model. IEEE Transactions on Image Processing **29**(1), 5022–5037 (2020)
29. Xu, L., Zheng, S., Jia, J.: Unnatural l_0 sparse representation for natural image deblurring. In: IEEE Conference on Computer Vision and Pattern Recognition. pp. 1107–1114 (2013)
30. Zhan, K., Shi, J., Teng, J., Li, Q., Wang, M., Lu, F.: Linking synaptic computation for image enhancement. Neurocomputing **238**(0), 1–12 (2017)
31. Zhan, K., Teng, J., Shi, J., Li, Q., Wang, M.: Feature-linking model for image enhancement. Neural Computation **28**(6), 1072–1100 (2016)
32. Zhang, Q., Yuan, G., Xiao, C., Zhu, L., Zheng, W.S.: High-quality exposure correction of underexposed photos. In: ACM International Conference on Multimedia. pp. 582–590 (2018)