ADVANCED INTELLIGENT SYSTEMS Open Access

www.advintellsyst.com

GelLight: Illumination Design, Modeling, and Optimization for Camera-Based Tactile Sensor

Jieji Ren, Wenxin Du, Yueshi Dong, Ningbin Zhang, Heng Guo, Boxin Shi, Jiang Zou, and Guoying Gu**

Camera-based tactile sensors attract the attention of the robotics communities by the high-density tactile perception, in which image quality and reconstruction accuracy are significantly determined by the illumination design. However, the influence of illumination has not yet been systematically analyzed, and most existing sensors adopt empirical design and subjective evaluation to determine the light configuration. Herein, a photometric stereo-based modeling, optimization, and evaluation system is proposed to explore the best illumination for typical camera-based tactile sensors. First, this article constructs a tactile benchmark dataset, simulates the contact deformation of elastomer surface, rendering the tactile imaging under various illuminations, and constructs a metrics system to evaluate the performance. Then, the relationship between reconstruct accuracy and illumination direction distribution on the benchmark is depicted, and the best illumination is optimized. The optimized sensor is fabricated and evaluated by standard metrology experiments, which exhibits high reconstruction accuracy and convincingly demonstrates the effectiveness of the proposed design and optimization approach. Furthermore, intensive experiments are conducted on diverse objects, which additionally indicate the generality and adaptability of the designed sensor. Herein, the illumination design can simplify and improve the performance of camera-based tactile sensors.

1. Introduction

Tactile sensing is one of the core perception abilities for robots,[1–4] which have developed rapidly in recent years.[5–8] Camera-based tactile sensors can provide subtle texture of target surface and dense force distribution in interactions,[9] which greatly improve the performance of environment perception and dexterous manipulations for robots. Camera-based tactile sensors utilize the camera to capture the image of silicon gel deformation in the contact process under the illumination (such as light-emitting diodes (LEDs)) from various directions with different colors and then construct the surface details with image intensity and color information, as shown in **Figure 1**a,b.

Since camera-based tactile sensors are proposed[10] and present the high-density perception ability, researchers have begun to explore their potential for robotics.[11] Based on the perception mechanism, diverse designs and improvements are proposed. Researchers minimized the size of GelSight and introduced it into robot gripper[12] grasping and manipulation, which paves the way for robotic applications. After that, researchers began to explore the rich tactile information from camera-based tactile sensors, such as shear force,[13] sliding[14] and hardness,[15] to improve the reconstruction accuracy,[16,17] and propose algorithms for surface texture perception,[18] geometry reconstruction,[19] and force distributions.[20] To facilitate the combination with robots, especially with robot grippers, advanced designs are proposed, such as GelSlim 3.0,[21] GelSight 360,[22] and DigiTac.[23] Besides, camera-based tactile also empowers the high-resolution tactile to the dexterous and integrate with robot fingers[24,25] and hands,[26–29] which presents significant potential applications. With those rapid developments, the design and performance optimization have attracted extensive attention.

J. Ren, Y. Dong, N. Zhang, J. Zou, G. Gu
Robotics Institute
School of Mechanical Engineering
Shanghai Jiao Tong University
Shanghai 200240, China
E-mail: zoujiang@sjtu.edu.cn; guguoying@sjtu.edu.cn

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/aisy.202400596.

DOI: 10.1002/aisy.202400596

W. Du
School of Electronic Information and Electrical Engineering
Shanghai Jiao Tong University
Shanghai 200240, China

H. Guo
School of Artificial Intelligence
Beijing University of Posts and Telecommunications
Beijing 100876, China

B. Shi
National Engineering Research Center of Visual Technology
School of Computer Science
Peking University
Beijing 100871, China

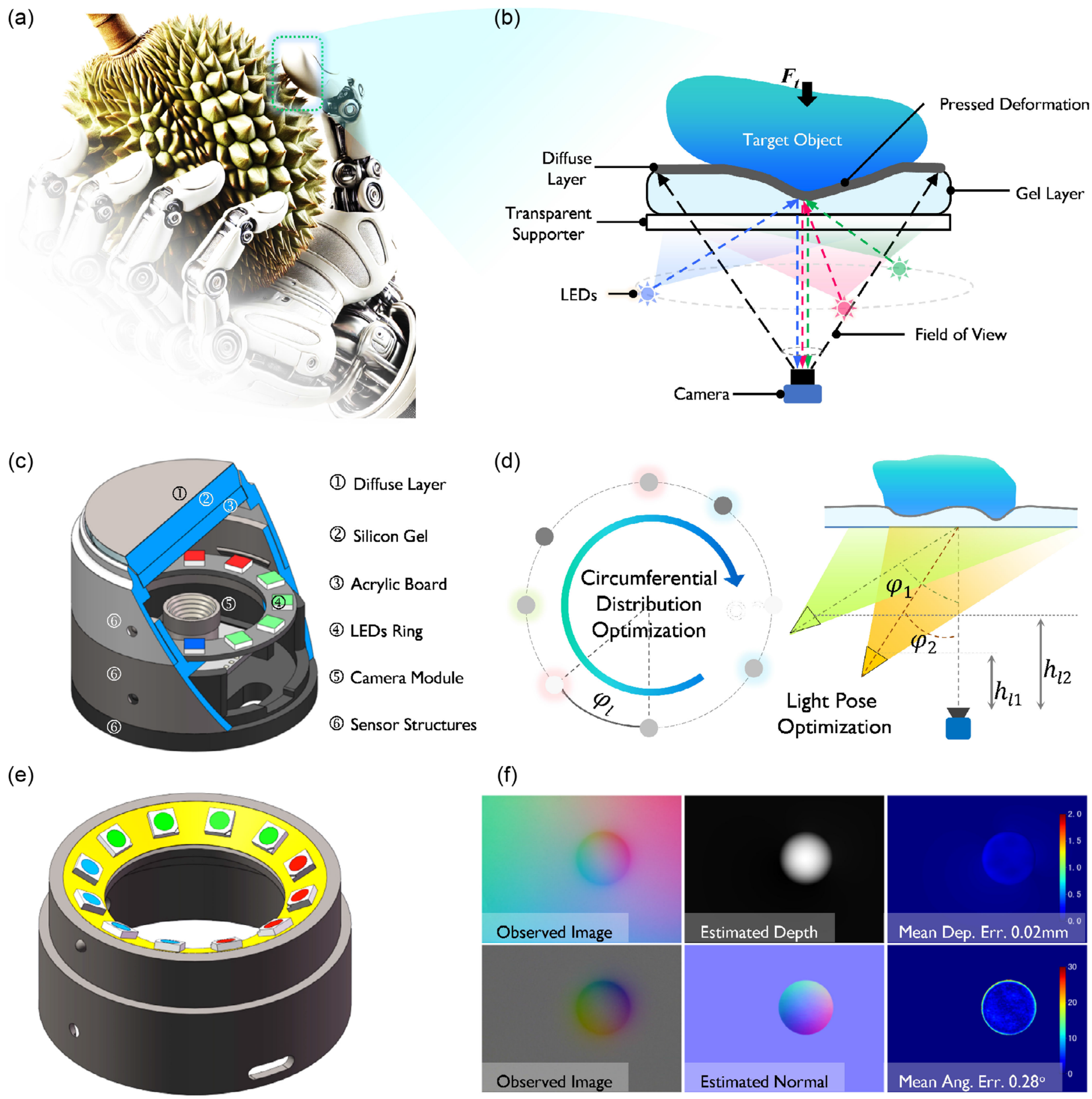

Figure 1. Camera-based tactile sensors can provide high-performance tactile performance for robots. a) High-density tactile sensing is significant for robot hand perception and dexterous manipulations. b) The sensing mechanism of the camera-based tactile sensor involves using the camera to observe the pressed deformation of the gel surface under colorful LED illuminations and reconstructing the deformation. c) Section view of the designed typical camera-based tactile sensor in detail (blue region is the section face). d) Illumination distribution, incident angle, and pose optimization of camera-based tactile sensor. e) The optimized illumination of camera-based tactile sensor. f) Accurate tactile sensing performance (surface geometry reconstruction accuracy less than 50 μm).

However, researchers have gradually discovered that the direction, distribution, and emit (incident) angular of light strongly affect the imaging process of camera-based tactile sensors[30] and lead to the tactile performance dramatically changing. The illumination design significantly determines the performance of camera-based tactile sensors.[12,31] It is important to model the illumination and optimize sensor design, but existing tactile sensors still depend on an empirical approach and subjective vision evaluation for illumination design. From the early stage, researchers have empirically designed the light positions and angles. The GelSight mini[12] takes an optical waveguide to illuminate the gel surface from the side. Then, the improved version[16] takes a large tilt angle to increase the uniformity of illumination and provide color variance for different surface geometries. Besides, the reflection optical path is introduced to reduce the sensor thickness.[17] To increase adaptability, non-planer designs are proposed. For the round head shape,[32] side illumination is designed with the help of total internal reflection. The transparent shells not only support the structure but also prevent the light escape in the complex sensor shape.[33]

Besides, more complex light strips are adopted to illuminate curved surface shapes.[22,24] Researchers also introduce the illumination uniformity matrices and automatic optimized process to design the optical path.[21] However, existing illumination designs strongly depend on the sophisticated experience of researchers. The influence of illumination on reconstruction accuracy is still unclear.

Although some researchers have proposed related design criteria[22,31] and evaluation processes,[16,34,35] most of them have only optimized a single characteristic of sensor design space or just presented limited perspectives on tactile performance, without a systematic and theoretical analysis of illumination. Without systematic quantity metrics, artificial design and subjective evaluation not only slow the sensor optimization but also obstacle the automatic algorithm to accelerate the evaluation process. Besides the lack of an effective design pipeline, researchers cannot conveniently adopt a pipeline to design customized sensors to satisfy diverse applications, which also impedes applications of camera-based tactile sensors in the robotics community. It is expected to build a benchmark and systematic metrics to analyze and evaluate the illumination performance, as well as an effective design and operational optimization method for camera-based tactile sensors.

On the other hand, the physical-based simulation environment of camera-based tactile sensors is significant for illumination design, analysis, and optimization efficiently at low cost. There are famous works such as Taxim,[36] FOTS,[37] and early Sim2Real,[35] but most of them are fit for predefined illumination settings of special sensors. TACTO-[38] and PBR-based methods[30] provide more flexible imaging simulation, but their rendering is based on the depth map (such as those from Pybullet) and fine tune with real sensors. Their main target is to generate a large amount of dataset for training downstream tasks, instead of optimizing the sensor characteristics. To optimize the illumination, we need to build a sensor-independent simulation environment with high-accuracy mechanical deformation simulation of the gel layer under contact force (such as FEM-based method[39]) and a rendering system to image the deformation under flexible and controllable illuminations. Furthermore, quantitative metrics are indispensable for optimization and evaluation. Camera-based tactile sensors adopted simplified photometric stereo[40] to reconstruct the geometry of contact surface,[10] essentially. Consequently, we will refer to the theoretical method of photometric stereo[40] to build the evaluation method and metrics for illumination design and optimization.

In this work, we establish an effective modeling and optimization framework for camera-based tactile sensor, especially for the most widely used reflective type. Since the ability to reconstruct geometry is one of the most important and distinctive features of camera-based tactile sensors, we prioritize discussing how to enhance the accuracy of static object reconstruction in this work. To facilitate the design process, we first construct a camera-based tactile simulation environment to simulate the gel deformation under object pressure and the tactile imaging with various illuminations. To evaluate the quality of image and performance of tactile sensors under various illuminations, we propose a tactile benchmark with typical surfaces and textures and systematic quantitative metrics. Based on the previously mentioned infrastructures, we utilize a typical camera-based structure as a target sensor to present the design and optimization process of illumination, where the light sources are usually mounted on a ring with the same pose. We first consider the ideal assumption (monochrome directional light and diffuse reflection) and optimize the incident angle and the rotation angle (circumferential distribution) of light by the surface normal reconstruction accuracy with photometric stereo method.[41] To fine-tune the illumination configuration, we then replace the directional light with spotlight (with uniform RGB color illumination) to describe the real situation more accurately and optimize the mounted height of lights to increase the illumination uniformity and linearity of color encoding, as well as the surface reconstruction accuracy. Furthermore, we completely analyze the influence of illumination factors and conclude a general instructive design pipeline for camera-based tactile sensors. Finally, we fabricated the camera-based tactile sensor with the optimized illumination and verified its performance in real world, which can achieve high accuracy better than 50 μm. Intensive experiments exhibit the generality, adaptability, and robustness of the optimized camera-based tactile sensor.

This work not only constructs a theoretical analysis framework of illuminations, which strongly supports the optimization process of light configuration, but also develops a general simulation environment and a systematic benchmark dataset for camera-based tactile sensors. These contributions may provide the community with effective tools and a practical pipeline for sensor design, analysis, and optimization, potentially expanding related research and promoting the further development of camera-based tactile sensors.

2. Design and Modeling

We first describe the mechanism of camera-based tactile sensor and present a typical sensor structure as the following research target. To optimize the illumination, we build a deformation and imaging simulation environment for the sensor and propose a tactile benchmark for evaluation.

2.1. Sensor Design and Benchmark

2.1.1. Camera-Based Tactile Sensor and Task Formulation

To investigate the illumination characteristics of camera-based tactile sensor, we first propose a typical design, as shown in Figure 1c. The sensor mainly consists of the imaging module (camera), the illumination module (LEDs ring), and the gel module (support acrylic board, soft gel layer, with a diameter 50 mm). The camera is accommodated on the bottom base structure and toward the gel. Then the LEDs ring fix on the middle base and connected on the top of the bottom base. The top base holds the acrylic to support the gel layer and provide an optical path for the camera to observe the deformation of the gel surface. A thin layer is painted on the gel surface to provide diffuse reflection. To simplify the analysis and fabrication process, the demo is designed as a cylinder, and the modules are assembled as a stack structure. In this design, we take the LED as the illumination source. Considering the accessibility of manufacturing and fabrication, in this work, the LEDs are co-planar and have the same incident angle.

The tactile sensing performance is strongly related to the characteristics of the camera, illumination, and gel surface and their relative pose. However, most of the existing camera-based tactile sensors utilize commercial cameras, whose performance is fixed upon purchase. Additionally, gel deformation and surface reflections are also determined by the relatively stable formulation and fabrication process. On the other hand, the slight variations in illumination can cause a significant shift in the captured images, as well as in the modeling and the ultimate tactile performance. Thus, the illumination has a more profound effect on the camera-based tactile sensors than other aspects, which plays a critical factor in tactile perception. This work is focused on the design, modeling, and optimization of illumination. The following analysis is based on the predefined camera (field of view), LED (angle of radiation), and the module size (the diameter of the base), and we need to optimize the LED distribution and pose (numbers, incident angle, and the accommodated height).

2.1.2. Benchmark Dataset for Evaluation and Optimization

A comprehensive and representative dataset is important to effectively evaluate the performance of camera-based tactile sensors and optimize the illumination configuration. The benchmark needs to cover diverse surface geometries and be suitable for the characteristics of camera-based tactile sensors. Based on the previous researches[35,36,38,42] and the observation of the surface of the daily objects, we design a benchmark dataset with 10 objects, as shown in **Figure 2**a. We classify the surface into four classes: simple geometry, regular array, periodical surfaces, and complex details, and the shape design parameters are

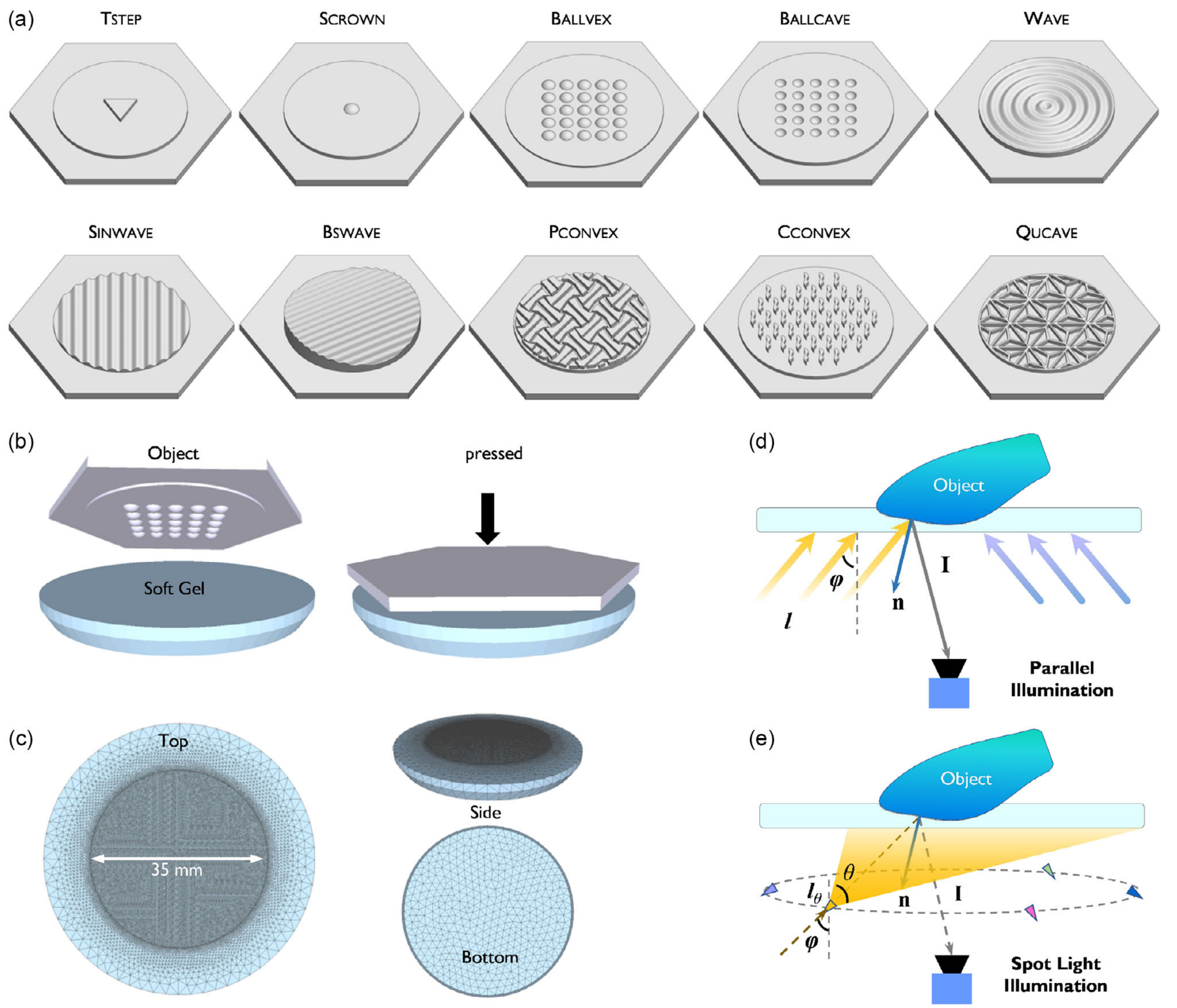

Figure 2. Benchmark dataset and simulation environment. a) The benchmark dataset with 10 typical objects, which includes the simple, regular, period, and complex array groups. b) Simulation process of the tactile contact process. c) Region-wise meshing of the gel for accurate tactile simulation. d) The imaging process with ideal parallel directional illumination; e) the spot LED illumination with a light beam angle θ. The light **l** illuminates the deformable surface with incident angle φ, reflected by the surface (normal vector **n**), and recorded by the camera (intensity **I**).

determined by the gel thickness, mechanical characteristics, and the deformation (details height scale: 1 mm, region diameter: 35 mm).

Simple Group contains a triangular step TSTEP and a spherical crown SCROWN. The step with predefined height can evaluate the depth reconstruction accuracy of tactile geometry, and the top plane can also evaluate the perception variance. The spherical crown has a rich surface normal direction, which can test the spatial correlation of tactile sensing, and can exhibit the reflectance change with normal vector.

Regular Array Group evaluates the tactile performance on dense array geometries. The sphere shape contains nearly all directions surface normal. The BALLVEX is convex ball array, and the BALLCAVE is the concave ball array. The main array introduces strong shadow in the imaging process, which could test the robustness of illumination. Besides, the convex–concave can evaluate the performance of camera-based tactile sensor in the gel compressed and extruding state.

Period Group includes common patterns in daily life, such as the graining and strip on the holder, the details on the non-slip mat, and the texture of the clothes. We design three typical structures to represent them. WAVE provides circle-like pattern, which has the omni-directional fluctuation; SINWAVE provides periodic sine texture, which is widely existing on artificial and natural surfaces and can evaluate the lateral resolution of tactile sensing; BSWAVE simulated the details on a nonplanar by the compound of sine pattern and spherical surface, which may lead to more complex contact and imaging behavior.

Complex Array Group introduces more complex artificial structures. PCONVEX takes polygons convex to densely fill the region, while the CCONVEX takes curved convex to sparsely fill the region. QUCAVE utilizes quadrangle concave to simulate the hollow surface. These complex surface structures can provide more challenging contact and imaging process and lead to comprehensive evaluation for the tactile performance under different illumination configurations.

With these representative objects, we can evaluate the tactile sensing performance in simulation. We first simulate the contact process between the object and the gel layer of sensor and generate the surface deformation of gel. Then, simulate the observation results of the deformed gel surface under different illuminations. Finally, we evaluate the tactile performance by calculating the geometry reconstruction error (such as the mean angular error (MAE) of the surface normal) and the light uniformity (such as the intensity contour map).

2.2. Modeling and Simulation

2.2.1. Gel Deformation Simulation

To accurately simulate the deformation process of elastomer layer, we based on the approach from TacIPC[39] to develop our mechanical simulator, which adopts the FEM-based incremental potential contact (IPC) method.[43] This simulator can calculate the gel surface deformation in the contact and press process, as shown in Figure 2b. Unlike existing FEM-based simulation processes, TacIPC can provide a more accurate contact process and deformation of the gel surface and prevent the gel surface penetration and mesh distortion. In our method, we model the sensor gel elastomer as a soft body represented by a tetrahedral mesh, where its material is characterized by the Neo-Hookean constitutive model. Since the benchmark object is far stiffer and harder than the gel elastomer, we model the object as an affine body, which is almost rigid and is described in ref. [44]. The contact and interaction between the sensor gel elastomer and the object are accurately handled by IPC. IPC introduces a barrier energy term to solve the collision contact, which can improve the numerical stability of the simulator (see mathematical details in S1, Supporting Information).

In our simulation, as the design of Figure 1c, the thickness of the gel is 5 mm, and the diameter of the contact surface is 50 mm. To improve the resolution of the simulation and balance the computation overhead, we take the region-wise mesh generation method, as shown in Figure 2c, and utilize dense mesh (average edge length: 0.1 mm) in the center region (diameter: 35 mm, which can cover the geometry region of the benchmark) to obtain more delicate deformation. To enhance the simulation efficiency, we also implement the algorithm on GPU, incorporating the region-wise meshing strategy.

2.2.2. Imaging Process Simulation

Camera-based tactile sensors capture the deformation image of gel layer under different illuminations and reconstruct the deformation based on photometric stereo.[40] To simulate the imaging process and optimize the illuminations, a computer graphics-based imaging simulation approach (rendering) is adopted. Inspired by previous physical-based[35,38] and learning-based methods,[36,37] the image simulation should provide fast and flexible rendering ability. It is necessary to simulate the light characteristics.[30] We can use the rendering equation to describe the illumination and imaging process. The observed intensity of a point on the reflection layer can be described by:

$$I = f_{\text{BRDF}}(\mathbf{l}, \mathbf{n}, \mathbf{v}) \cdot \mathbf{n} \cdot \mathbf{l} \tag{1}$$

where the f_{BRDF} is the bidirection reflection distribution function, which describes the reflection characteristics of the surface. The $\mathbf{n} = [n_x, n_y, n_z], \mathbf{l} = [l_x, l_y, l_z]^{\mathrm{T}}, \mathbf{v}$ are the surface normal of this position, the illumination light vector, and the observation vector, respectively.

First, to simplify the theoretical analysis, we only consider the directional (parallel), uniform and monochromatic light, which means the light direction and intensity is constant. We also assume the reflection layer of the camera-based tactile sensor is lambertian, which means the reflection albedo can be considered as constant. We also put the surface on the top of camera, which leads the $\mathbf{v} = [0, 0, 1]$. The reflection can be expressed as a constant parameter ρ. So the Equation (1) reduces to:

$$I = \rho \mathbf{n} \cdot \mathbf{l} = \rho [n_x, n_y, n_z] \begin{bmatrix} l_x \\ l_y \\ l_z \end{bmatrix} \tag{2}$$

Based on Equation (2), we can simulate the deformation image of the gel under different directional illuminations. This simulation can support our coarse optimization and find the best distribution, and the incident angle of lights is based on the reconstruction accuracy by photometric stereo method under proposed tactile dataset, as shown in Figure 2d.

However, it is hard to ensure directional and uniform illumination in real world, especially for the common used LEDs in camera-based tactile sensors. Considering the LEDs are diverge and close to the reflection layer in the camera-based tactile sensor, we improve the rendering Equation (1) to describe the real situation with constant ρ and support the fine optimization.

$$I = \mathbf{n} \cdot \mathbf{l}_\theta(z, \varphi) \tag{3}$$

where θ is the divergence angle of LED, $\mathbf{l}(z, \varphi)$ means a light source is placed at the height z with incident angle φ, as shown in Figure 2e. To achieve more uniform illumination, we can adjust the height and the angle of the LEDs (we implement a rendering pipeline with the Unity engine.). Specifically, we set the uniform height and incident angle for all LEDs in the camera-based tactile sensor because it is convenient and robust to place LEDs on a ring in the fabrication process. We can maximize the illumination uniformity on the gel surface and minimize the reconstruction error of touch object to find the best height of LEDs pose and angle.

With the modeling of mechanical press process on the gel and optical imaging under illumination, we can simulate the tactile sensing process, evaluate the performance under different configurations, and optimize the distribution, pose, and angle of illumination hierarchically, to improve the tactile accuracy.

2.3. Optimization Method

Camera-based tactile sensor utilizes the observations of gel surface deformation under illuminations to calculate the target surface geometry, as well as the force distribution.[9] The surface reconstruction accuracy determines the performance of tactile sensing and strongly depend the illumination configurations, such as the incident direction, divergence, and the uniformity of the light source. Based on previous simulations, in this section, we will comprehensively analyze the effect of these factors on reconstruction and optimize the illumination to achieve a high-accuracy tactile sensor.

We adopt the hierarchical light optimization process. First, the coarse optimization is processed under ideal setting, which assumes that the illuminations are parallel and uniform. The coarse optimization analyses the light distribution and the light incident angle. Based on the optimized distribution and incident angle, a more realistic setting (spotlight with divergence LED angle) is introduced. We need to slightly adjust the height of lights to achieve more uniform illuminations.

2.3.1. Coarse Optimization with Ideal Setting

To simplify the analysis process, we first take the parallel, uniform, and monochromatic light in the coarse optimization. As shown in Figure 2d, the number of lights is n and their incident angle is φ. The lights are uniformly distributed on a circle, which leads to equal angular separation $360°/n$ between lights. The incident angle is the angle between light direction and normal of acrylic bottom surface. Considering generality and practicability, we choose the number of lights as $n \in [3, 6, 9, 12, 15, 18, 21]$, which are multiples of three for further analysis (R, G, B LEDs in the chromatic version). To cover a wider angular range, we set the incident angles as $\varphi \in [5°, 85°]$ with 8° step. The 0° and 90° are not easily to set in the real world. Based on the predefined illumination space, we uniformly placed n lights on each incident angle. Each light group is composed of n lights at a fixed incident angle φ. The imaging simulation module renders the observed images under the illumination from each light group $G_{\varphi,n}$ by sequential turning on and off the light.

To effectively evaluate the subtle geometry, we choose the surface normal map and its error as the reconstruction target and metric. With these observed images and the light directions, we can reconstruct the surface normal map of the deformable gel surface by classical photometric stereo method (such as LSPS[40]), and calculate the MAE with the ground truth normal (see details in S2, Supporting Information). For each pixel position i, we can adopt the least-square based photometric stereo[40] to calculate the surface normal $\mathbf{n_i}$ of this point. For all the pixels of the observed image, the surface normal map $\mathbf{N}$ can be calculated pixel-by-pixel. Based on the ground truth (GT) normal map from the benchmark, we can obtain the reconstruction performance by the MAE between the GT normal $\mathbf{n}_{igt}$ and calculated normal $\mathbf{n}_i$:

$$\mathrm{MAE} = \frac{1}{HW}\sum_{i}^{HW} \arccos \frac{\mathbf{n}_{igt} \cdot \mathbf{n}_i}{|\mathbf{n}_{igt}||\mathbf{n}_i|} \tag{4}$$

where H and W are the height and width of the observed image, $\mathbf{n}_i$ and $\mathbf{n}_{igt}$ denote the estimated, and the GT surface normal at pixel position i. Based on the simulation and reconstruction, we can obtain the error distribution under different light positions and incident angles of the camera-based tactile sensor. The reconstructed MAE heat map of SCROWN is shown in **Figure 3**a. (The error changes from small to big with the color changing from blue to red. We set the color map value in $[0°, 30°]$, which makes the block show same color (red) when $\mathrm{MAE} > 30°$.) In the heat map, the number of light increases from 3 to 21 on the vertical axis, and the incident angle increases from 5° to 85° on the horizontal axis. The left image is the reconstructed MAE of SCROWN under different light groups.

However, because the touched surface is close to planar and the imaging process utilizes the ideal setting, the error is too small to find the best light group (close to the modeling error). Considering the in-homogeneous error distribution, for each reconstructed surface normal map, we selected the top %5 and %1 locations to calculate more obvious MAE:

$$\begin{aligned}
\mathrm{Err} &= \arccos \frac{\mathbf{n}_{igt} \cdot \mathbf{n}_i}{|\mathbf{n}_{igt}||\mathbf{n}_i|}, i \in H \times W \\
\mathrm{MAE_{Top5}} &= \mathrm{mean}(\mathrm{Err}[\mathrm{Err} > P_{95}(\mathrm{Err})]) \\
\mathrm{MAE_{Top1}} &= \mathrm{mean}(\mathrm{Err}[\mathrm{Err} > P_{99}(\mathrm{Err})])
\end{aligned} \tag{5}$$

where P_{95} and P_{99} are the percentile, and Err is the angular error matrix. We will statistic the angular error, find the top error points, and calculate the $\mathrm{MAE_{Top5}}$ and the $\mathrm{MAE_{Top1}}$. The middle and right heat maps in Figure 3 show the obvious trend under different light groups, especially on $\mathrm{MAE_{Top1}}$. Besides, the top

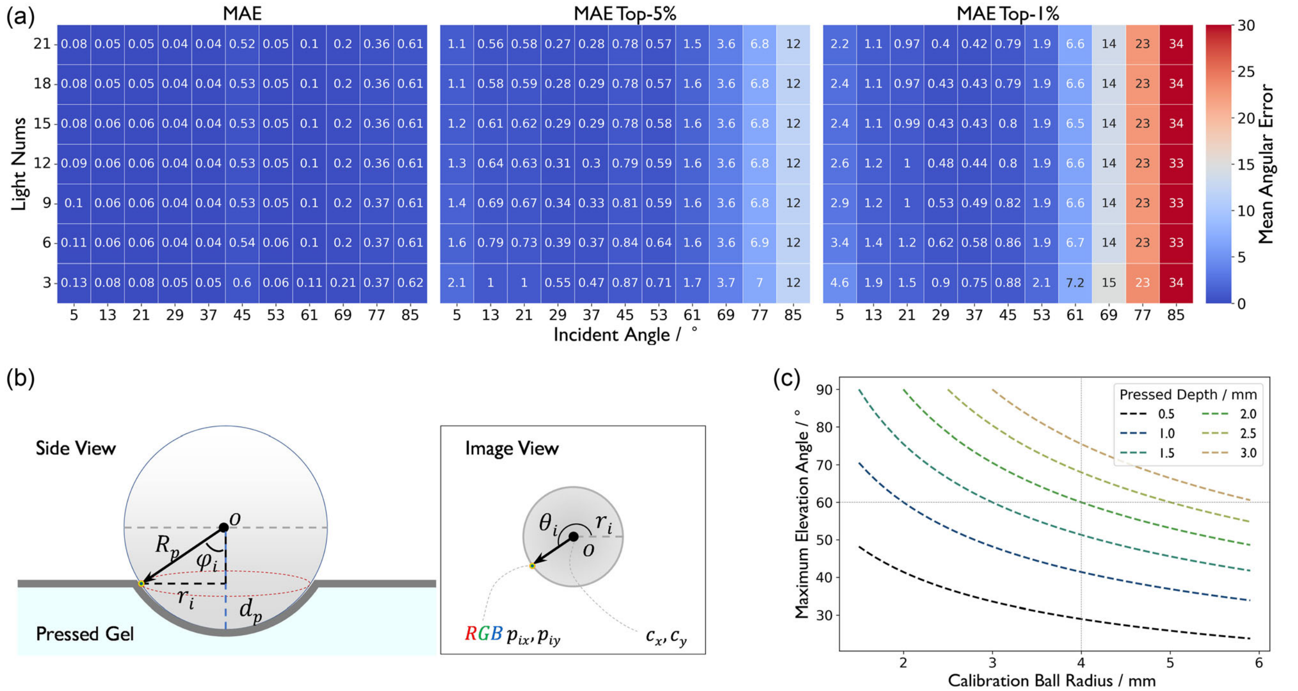

Figure 3. a) Heat map of reconstructed mean angular error (MAE) on SCROWN under different light groups. The left part is the original MAE, the middle and right part is the MAE of the top 5% and 1% error pixels. b) The calibration process could build the mapping look-up table $[\{R_i, G_i, B_i\}, \{G_{ix}, G_{iy}\}]$ for the camera-based tactile sensor. With the help of sphere, we can calculate the gradient for each pixel in the pressed region and record the corresponding pixel value. c) Recorded maximum elevation angle changes with the calibration ball radius at different pressed depth.

errors determine the worst performance of the illumination configuration and contribute the dominate component of the error based on the histogram analysis, which value is representative for error evaluation. So, we will take the $\mathrm{MAE}_{\mathrm{Top1}}$ results for the following analysis.

2.3.2. Fine Optimization with Real Setting

Based on the typical configuration in Section 2.1.1, we uniformly put Red, Green, Blue LEDs on the rings and introduce spot-light model to simulate real illumination. Camera-based tactile sensor[9] needs uniform illumination intensity of channel to ensure accuracy. However, the color-related perception mechanism also needs a sufficiently colorful image. So we need to fine-turn the illumination with the image variance and saturation. Then we can use the standard deviation of illumination intensity δ_I on the gel to evaluate the illumination uniformity of each color.

$$\delta_I = \delta_{\mathrm{r}} + \delta_{\mathrm{g}} + \delta_{\mathrm{b}}$$

$$\begin{aligned} \delta_{\mathrm{r}} &= \sqrt{\tfrac{1}{HW}\textstyle\sum_{i=1}^{HW}\left(I_{\mathrm{ir}} - \hat{I}_{\mathrm{r}}\right)^2}, \\ \delta_{\mathrm{g}} &= \sqrt{\tfrac{1}{HW}\textstyle\sum_{i=1}^{HW}\left(I_{\mathrm{ig}} - \hat{I}_{\mathrm{g}}\right)^2}, \\ \delta_{\mathrm{b}} &= \sqrt{\tfrac{1}{HW}\textstyle\sum_{i=1}^{HW}\left(I_{\mathrm{ib}} - \hat{I}_{\mathrm{b}}\right)^2} \end{aligned} \tag{6}$$

where the I_{ir} and I_{r} are the intensity of each position and the average illumination of the observed region for the red channel. Because we can set same intensity and reflection coefficient, we only need to optimize one channel to simplify the process. Considering the low-intensity value will decrease the image quality, we will take the mean value to normalize the standard deviation (the smaller intensity will lead to a higher normalized standard deviation). The normalized standard deviation (NSD) and saturation (SAT) are written as

$$\begin{aligned} \mathrm{NSD} &= \delta_I/\mathrm{mean}(I) \\ \mathrm{SAT} &= \frac{\max(I_{\mathrm{R}}, I_{\mathrm{G}}, I_{\mathrm{B}}) - \min(I_{\mathrm{R}}, I_{\mathrm{G}}, I_{\mathrm{B}})}{\max(I_{\mathrm{R}}, I_{\mathrm{G}}, I_{\mathrm{B}})} \end{aligned} \tag{7}$$

where mean, max, min is the mean, maximum and minimum function, and I is the image matrix. We integrated perception requirements for uniformity and saturation, using NSD and SAT as metrics for fine-tuning. Combining with the illumination uniformity and color appearance, the height of LEDs will be optimized. For generality, we focus on the illumination-based metrics for the fine optimization by adjusting the mounted height of LEDs.

Except for illuminations, the diameter of the calibration ball and the pressed thickness of the gel layer are also important for the reconstruction accuracy. To sense the target surface, camera-based tactile sensor should transform the observed image into geometry. We can take a calibration ball to build the look-up table.[9] For a known-diameter d_c calibrate ball, we press it on the gel surface, calculate the gradient $[G_x, G_y]$ (see details in S3, Supporting Information) and record the pixel intensity $[R_i, G_i, B_i]$ of the pressed region (the coordinates of pixel can also be taken into consideration.), and build the look-up table from intensity to the surface gradient, as shown in Figure 3b.

However, the pressed depth and the ball radius will affect the recording range of the gradient (especially the elevation angle range), which may influence the reconstruction accuracy. In general, the compressible range d_p of gel surface is around 1–3 mm, so choosing proper calibration ball is necessary. The maximum recording elevation angle can be expressed as the following:

$$\varphi_{\mathrm{emax}} = \arctan\frac{r_{\mathrm{e}}}{R_{\mathrm{p}} - d_{\mathrm{p}}} \tag{8}$$

where R_p is the radius of calibration ball and the r_e is the radius of pressed edge under depth d_p. We can model and plot the maximum elevation angle variation tendency with these two factors.

The pressed depth in the calibration process is related to the hardness and thickness of the gel layer. After designing the gel layer and choosing the elastromer material, we could estimate or measure the typical pressed depth under given load and then can find the best calibration ball from Figure 3c. A good reconstruction result of camera-based tactile sensor needs the calibration process records elevation angle not less than 60°,[16] which needs small calibration ball.

However, the small ball may lead to pressed edge blur and lead to inaccurate edge detection, which may affect the accuracy of the look-up table (see details in S3, Supporting Information). A larger calibration ball may help to alleviate this problem because the edge may easily distinguish and the relative error is smaller. Besides, large pressed depth leads to serve shadow effect, which may decrease the accuracy of the high-latitude region on the calibrated ball (for example, with 29° incident angle of previous setting, shadows will appear in areas over 71° elevation angle). Considering the range of elevation angle, the accuracy of look-up table, and the compressible range of gel layer (5 mm thickness of this configuration), we choose the 4 mm radius ball and press 2 mm to calibrate the look-up table.

3. Results

3.1. Optimization Results

We implement the aforementioned design concept into a simulation environment, thereby generating a model. Through a hierarchical optimization, we determine the most favorable illumination configuration. In the initial stage, utilizing the principles of photometric stereo, we conduct a coarse optimization of the light distribution and the initial orientation. Subsequently, considering the perceptual requirement of the camera-based tactile sensor, we fine-tune the placement and orientation of the light source, ultimately yielding the optimal illumination setup.

3.1.1. Coarse Optimization Results

For the coarse optimization process, the light distribution and the incident angle should be determined. We follow the previously presented method to evaluate the reconstruction performance on the benchmark dataset and depict the heat map of $\mathrm{MAE}_{\mathrm{Top1\%}}$ for each shape. Details are shown in **Figure 4**a. It is clear that the reconstruction error decreases with the increase of light number, but tends to be flat after exceeding 12 lights. On the other hand, the error first decreases then increases with the incident angle increasing, which means there is a best illumination angle for camera-based tactile sensor on this object. The minimum error region is the mid-low latitude regions (top-left region on the heat map), which could guide the design of illumination distribution and incident direction in camera-based tactile sensor.

The heat map can depict the surface geometries distributions. For sample shapes, such as SCROWN and the SINWAVE, show large blue region with small error. On the other hand, the red region expands with the complexity of surface, such as the CAVE and CCONVEX. The sample geometry with high frequency may lead to large error, such as the simple TSETP with a rapid changing step. Furthermore, comparing the heat map of SINWAVE and BSWAVE, we can find that the low-frequency shape may not obviously affect the accuracy.

We further analyze the performance trend with the incident angle under different lights of each object, as shown in Figure 4b. It is obvious that the small and big incident angles are not friendly for camera-based tactile sensing. This is mainly caused by the shadows under larger angles (side incident) and the less contrast feature under small incident angles. The complex surface, such as CCONVES and QUCAVE, shows a more severe degradation curve with the increase of incident angle. Furthermore, for each object, the best performance angular is almost the same under arbitrary numbers of illuminations and their performance variation larger than the difference under light numbers. The design of camera-based tactile sensor should pay more attention to the illumination design, especially the incident angle. We also average the performance on all benchmarks and calculate the best incident angle as 32.2°.

On the coarse optimization stage, we balance the performance degradation trend with the light number and incident angle and determine the light number as 12 (ensure high accuracy and simplify the fabrication process with small number of lights) and the incident angle φ as 29° (most close to the average minimum error). We will refine the illumination light field in the following optimization process.

3.1.2. Fine Optimization Results

For the fine optimization process, we first analyze the intensity variation of different colors under colorful spotlight simulation. In the spotlight simulation, the gel layer is put at a height of 40 mm (h_g) from the camera, as shown in Figure 1c. We first mount the led ring at a height of 20 mm (h_l) from the camera, set the spotlight angle as 160°, and reference the radiation pattern as.[45] With the measured irradiance distribution, we can simulate the LED with high fidelity (Figure S9, Supporting

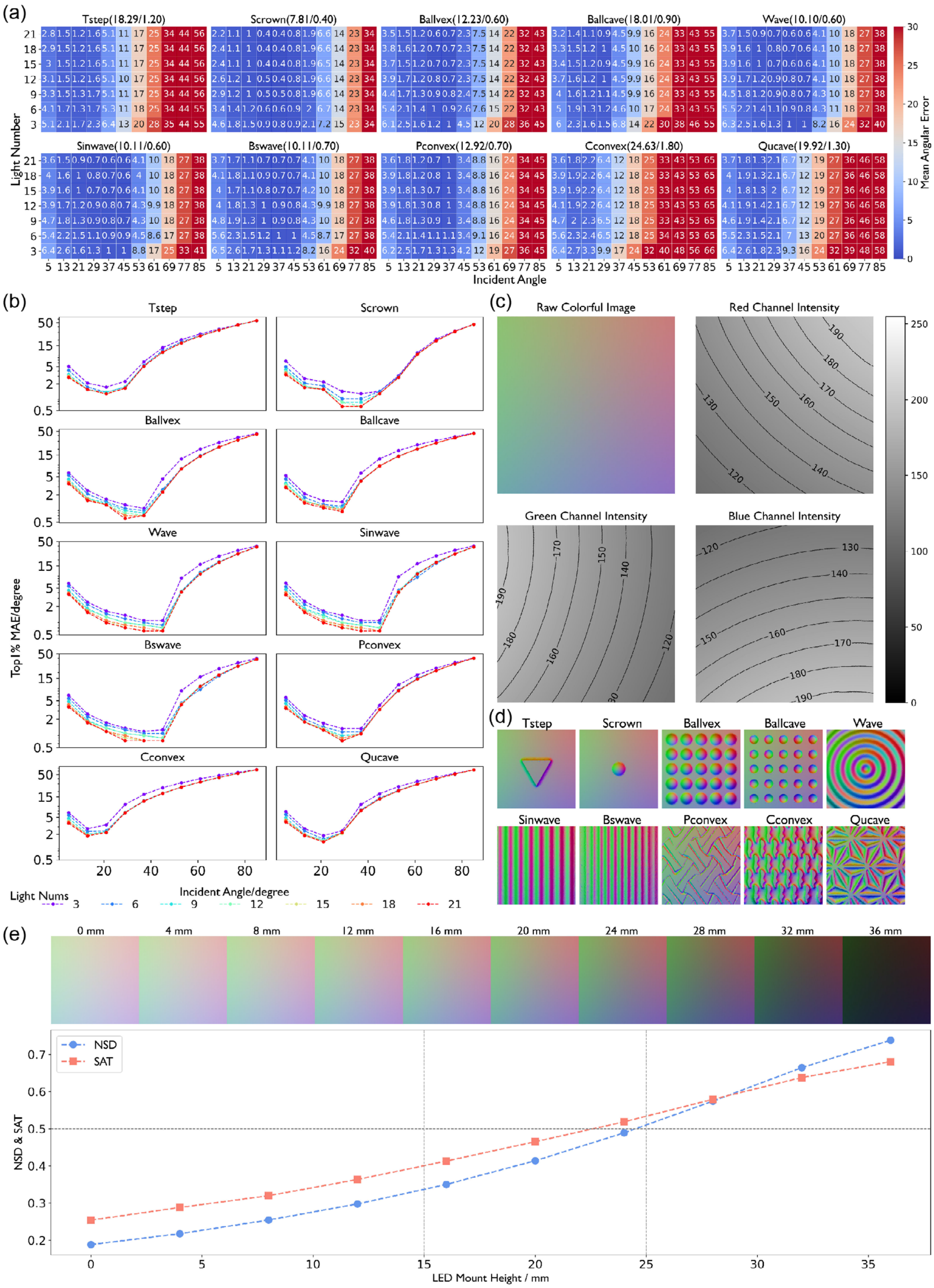

Figure 4. Simulation and optimization results. a) Heat map of Top1% reconstructed mean angular error (MAE) on each object in the benchmark under parallel and uniform illumination of light groups with different direction and distribution. The vertical axis represents the number of lights, and the horizontal axis means the increase of incident angle. b) The performance changes with the incident angle on each object, with various light numbers. To clearly show the trend, we set the y-axis as logarithmic coordinates. c) Simulation image of a standard plane (upper-left) and the intensity of each channel with intensity contour line. d) Simulation images of objects from benchmark on 20 mm height illumination. e) Simulation images at different height (upper). The normalized standard deviation (NSD) and image saturation (STA) change trend are plotted on the bottom as blue and red lines.

Information). We take the previous best configuration: incident angle 29°, 12 uniformly placed lights with [RRRRGGGGBBBB] color from 0° to 360° anticlockwise. We first illuminate a standard plane to measure the intensity variation of each channel. The distribution is shown in Figure 4c, which is consistent with LED distribution.

We also render the images of objects from the benchmark as shown in Figure 4d. We can clearly see color change with the surface normal under the colorful spotlight illuminations. However, the uneven color distribution and the intensity influence simulation results, such as the light color region on top and bottom of the BALLVEX and BALLCAVE, and the strong red or blue color region of wave-like shapes. We will further optimize the illumination configurations.

The mounted height h_l of LEDs will strongly impact the illumination uniformity in the realistic spotlight, we then optimize the mounting height $h_l \in [0, 36]$mm. In this process, we also take the standard plane to analyze the intensity distribution. Based on the sensing mechanism of the camera-based tactile sensor,[9] we prefer a more uniform intensity of each channel (related to the stability and consistency in perception range) and a more colorful of the image (related to the direction-sensitive color value).

As in the previous setting, we change the LEDs height h_l and calculate the normalized standard deviation (NSD), and the results are shown in Figure 4e. We can observe that the intensity decreases with the distance because the LED's maximum radiation direction deviates from the target center, and the total amount of light irradiated on the gel surface is reduced. So we should put the illumination source as far away from the gel surface as possible (in the sensor structure). On the other hand, the saturation (SAT) of the image should have a proper value. In Figure 4e, the images are fading on the upper-left with a small light height and dim on the bottom-right corner with a large light height. We also plot the saturation change of each image as the red line. For qualitative observation, higher saturation is preferred. However, higher saturation will present uneven color distribution of the image, which leads adverse effects for color intensity-based reconstruction (50% is preferable and the default setting of common applications). Considering the derivation of each channel and the saturation of the whole image, we properly choose saturation around 50% (light height h_l around 20 mm, with smaller NSD), as shown in **Figure 5**a.

3.2. Real-World Evaluation

Based on the previous optimized results, we fabricate a camera-based tactile sensor and evaluate the reconstructed accuracy in real-world experiments. The gel deformation under press, imaging quality, illumination uniformity, and reconstruction are comparable with simulation.

3.2.1. Fabrication

We first manufacture the mechanical parts and design electrical components for camera-based tactile sensor, as shown in Figure 5b. We cast and spray reflection coating for the gel layer, assemble each module, and integrate them as the complete sensor as shown in Figure 5c.

To facilitate the fabrication process, we take the 3D printer to build the mechanical modules and gel casting mold. As shown in Figure 5b(i), to cast the gel layer, we put a laser-cut acrylic board (thickness 3 mm, diameter 50 mm) in the bottom of the 3D-printed mold and poured the mixed PDMS silicon gel (Wesitru PDMS 0030) into the mold. After curing, we will spray a thin layer gray of silicon ink on the gel surface as a diffuse reflection layer. The processed gel will be mounted into the printed gel module for the following integration.

For the camera module, we designed a self-lock buckle to host the camera board and LED-driven board. We only need to insert these PCB boards into the printed compartment and connect the signal lines for assembly. Here we chose the camera with quite similar parameters as the simulation (it is hard to find the commercial camera module with exactly same parameters as simulation, not to mentioned the non-ideal imaging process) and put the camera in a precisely consistent pose and position as the simulation configuration.

For the illumination module, because the optimized design prefers the inclined incident angle (12 distributed lights, 29° with normal), we designed a customized flexible printed circuit (FPC) board with 12 programmable LEDs (WS2812B SMD5050) to achieve this configuration. Because the inclined LEDs ring should be placed on a side surface of truncated cone, so we should expand it as the contour of the FPC. Details are shown in Figure 5b(ii). After FPC manufacturing, we can fit the flexible board into the mounting surface of the printed illumination module, insert the LED signal pins into the tangential curved slot to connect with the driven board, and achieve compact inclined incident angle for the gel surface.

Based on the assembled modules, we can integrate them by mortise and tenon structure. The camera module acts as the sensor basis and assembles the illumination module on the top. Finally, the gel module will cover the top of the illumination module to complete the fabrication of a camera-based tactile sensor, as shown in Figure 5b(iii). To improve the stability and durability, we glue around the connection between components and modules (real-world sensor is shown in Figure S1, Supporting Information). The service performance is also verified by a cyclic loading experiment (Figure S8, Supporting Information).

3.2.2. Illumination Analysis

The observed image under optimized illumination is shown in Figure 5d, which is close to the simulated results (The discrepancy between real-world sensor and simulation is due to the simulation cannot perfectly model the camera response function, LED wavelength, and the color bias of the reflection layer.). The appearance of the observed image shows the uniformity and consistency of illumination, which ensures the reconstruction accuracy. We quantitatively analyze the intensity distribution of the observed image by the contour.

The average intensity distribution is stable, as shown in Figure 5d, middle of the first row. We show the average intensity with color-filled contour and plot the contour line with a value label. To relieve the effect of noise and outliers, we blur the original image before showing the contour. In the range of [0, 255],

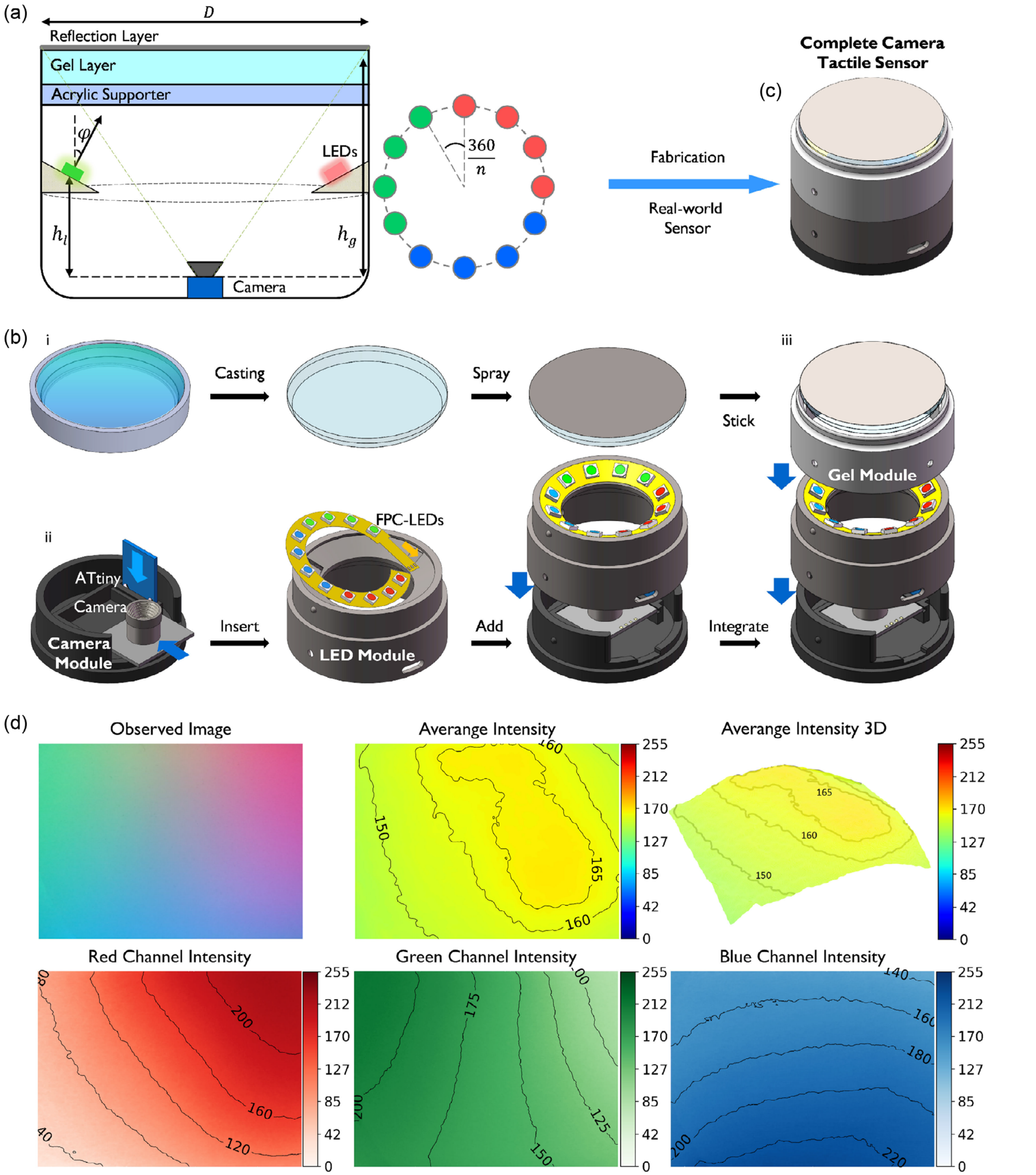

Figure 5. Optimized illuminations, fabrication process, and appearance of observed image. a) The optimized incident angle φ and mounted height of proposed method. We adopt the 29° incident angle for illumination and 20 mm to mount the illumination LEDs ring in implementation (the sensor diameter D is 50 mm and the distance from camera to gel surface h_g is 40 mm). b) Fabrication process of designed sensor. We use the 3-D printer to manufacture the mechanical base of the camera module, illumination module, gel module, and gel casting mold. We prepare the gel layer in (i), assemble each module with key components in (ii), and integrate modules as the optimized camera-based tactile sensor in (iii). c) Camera-based tactile sensor demo with optimized illumination. d) Uniformity analysis of designed illumination in real camera-based tactile sensor. The first row presents the original observed image, the average intensity, and the 3D version (we enlarge the amplitude to show the slight intensity change, the intensity distribution is uniform and only changes 15 (≈5%) from center to peripheral region (96.17% values locate in $\pm 2\sigma$)). The second row shows the intensity of each channel. The average intensity is uniform and stable with small variance. All the color intensity ranges are [0, 255].

the mean intensity of the observed image is 159.02, and the standard deviation is 6.67. To quantitatively analyze the uniformity, we count the number of pixels whose value is located between mean value and standard deviation range $\pm n\sigma$ and calculate their percentage of the whole image that 64.05 % for $\pm\sigma$, 96.17 % for $\pm 2\sigma$, and 99.90 % for $\pm 3\sigma$. The illumination intensity is stable and dominates large observed area, which strongly verify the effectiveness and progressiveness of the optimized illumination. For each color channel, we also present their intensity distribution and plot the contour line with value label in the second row of Figure 5d. The directional and continuously changed light can effectively encode the surface deformation geometries on the reflection layer.

3.2.3. Sensor Calibration

To reconstruct the target surface geometry from the tactile image, we should calibrate the corresponding real-world size of each pixel on the sensor surface and build the mapping relationship between the RGB image and the surface geometry. As shown in **Figure 6**a, we first calibrate the pixel size with a cylinder probe with an 8 mm diameter. We press the probe on the sensor surface in different locations, respectively, then manually depict the edge and use circle-fitting algorithm to determine the diameter in pixel P_d as 340.17 pixel ($\overline{P}_d = 340.17$, $\sigma = 0.47$). Then, we can calculate physical resolution of tactile sensor as 0.02352 mm pixel^{-1}.

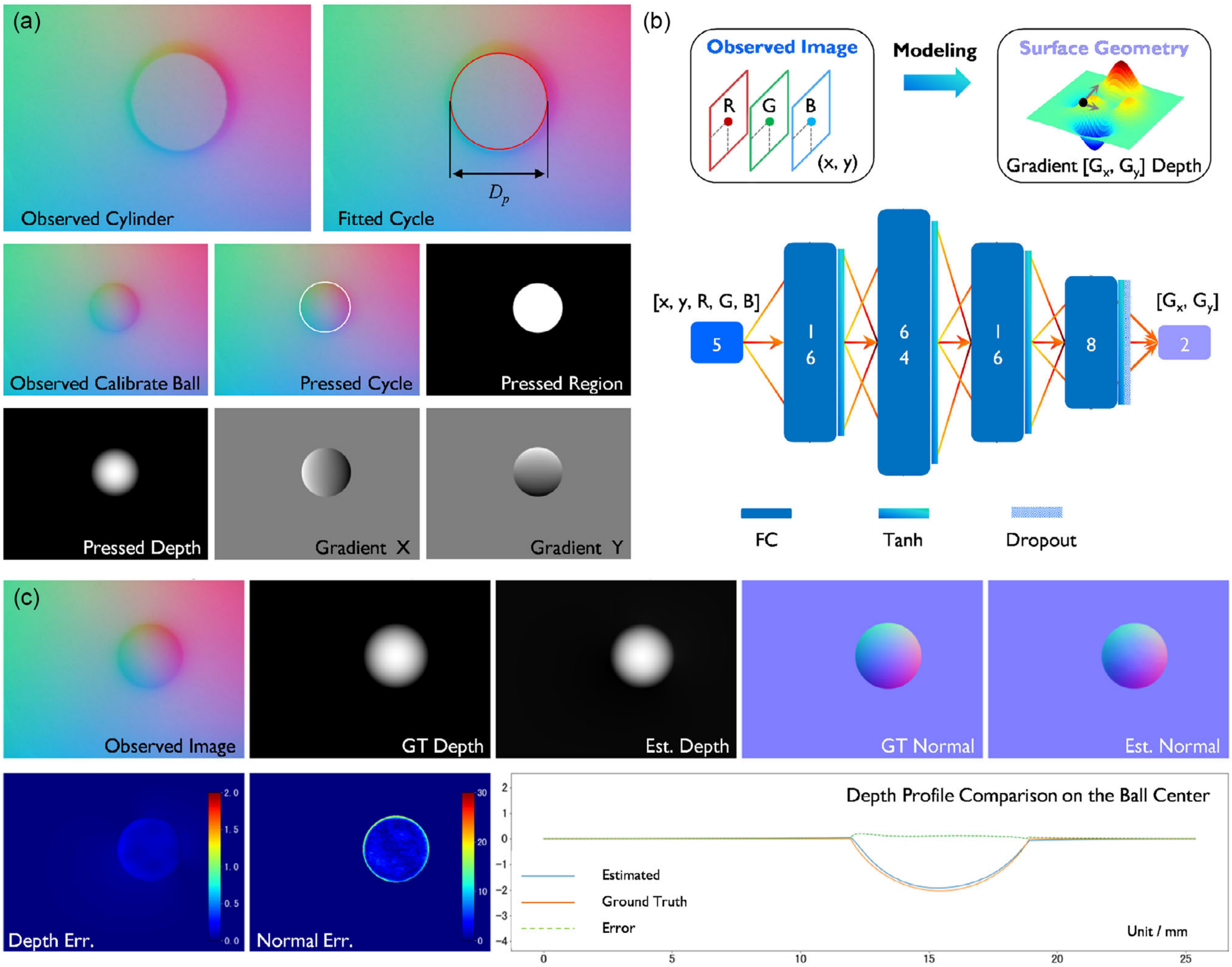

Figure 6. Calibration, modeling, and reconstruction results. a) Calibration process of the camera-based tactile sensor. We first take the cylinder press the surface to calibrate the real-world size of each pixel. Then, we utilized a ball to calibrate the mapping relationship between observed image and target geometry, by fitting the pressed cycle region, matting the cycle, calculating the height map, and generating the gradient map in x direction and y directions. b) The modeling process and the structure of proposed GelNet, which consists of five fully connection layers and a drop-out layer, with tanh activation. c) Reconstruction analysis of proposed sensor. First row shows the observed tactile image of pressed ball, the ground truth (GT), and the reconstruction (Est.) of pressed depth (the more white region means pressed more deeply), and the GT and Est. of surface normal map (N_x , N_y , N_z are colorized with R, G, B). The second row shows the error map of depth with 0.02 mm mean error, the error map of surface normal with 0.28° mean angular error, and the section profile of depth reconstruction on the horizontal center line of the ball, which shows the optimized sensor can achieve an reconstruction accuracy in micrometer scale.

We then need to construct the mapping relationship between the RGB value in the tactile image and the surface geometry (see details in S3, Supporting Information). Although we have optimized the illumination, we cannot totally eliminate the nonuniformity of light. As mentioned in refs. [17,22], we also record the coordinates of each calibrated pixel and take the algorithm to fitting the mapping between $\{R_i, G_i, B_i, P_{ix}, P_{iy}\}$ and $\{G_{ix}, G_{iy}\}$. We first take an 8 mm diameter ball to press the sensor surface as completely as possible and captured more than 100 pressed images with different locations. The captured images are substrate by the original image (without any object press) to obtain a difference image. Then, we can take the sphere function to calculate the gradient of each point in the pressed region, as described in Figure 3b. We use a semiautomatic method (manually initial labeling and machine vision-based detection) to accurately fit the pressed circle in the image as much as possible (see Figure S2, Supporting Information). Finally, we collected a N-by-7 table with $N = 5,702,702$ points from captured calibration images, each row contains pixel locations, values, as well as the corresponding gradient: $\{R_i, G_i, B_i, P_{ix}, P_{iy}\}, \{G_{ix}, G_{iy}\}$.

3.2.4. Nonlinear Mapping Modeling

The illumination intensity has unavoidable spatial nonuniformity, which introduces the nonlinearity into mapping and decreases the overhead of table look-up process, we build a neural network to model the mapping process between the $(R_i, G_i, B_i, P_{ix} P_{iy})$ and (G_{ix}, G_{iy}), as shown in Figure 6b.

Taking into account the dimensions and complexity of the nonlinear mapping, we build a 5-layer fully connected neural network named GelNet to model the nonlinear mapping between the observation and the gradient. For each location in the tactile image, the coordinates and observed value are concatenated as a vector and sent into the neural network. To satisfy the range to gradient, we take the hyperbolic tangent function (Tanh) as activation for each layer. The dropout layer is also introduced before the last layer to alleviate the overfitting of the network (implemented by pyTorch). The detailed structure is shown in Figure 6b. We determined the optimal network parameters through extensive experimentation.

Considering the compact size of GelNet, we take the CPU to train the network with a 128 batch size over 150 epochs, set the loss function as L1, and use the Adam optimizer with 10^{-6} learning rate. The captured dataset is spited with 9:1 for training and testing phrase. It is worth noting that we have normalized the original dataset to increase the training velocity, convergence, and generality. With the help of the GelNet, we can quickly process the observed values accurately, instead of the verbose table look-up (see details in Figure S2, Supporting Information).

3.2.5. Evaluation of Reconstruction Accuracy

After modeling the nonlinear mapping, we take the calibrated ball press the gel surface (obtain ground truth value from the spherical equation), utilize the trained model to estimate the corresponding gradient, reconstruct the geometry of the pressed region (2.5D depth map and 3D surface normal map), and evaluate the reconstruction accuracy with the optimized illumination. Details are shown in Figure 6c.

We can observe the reconstructed depth map is very close to the ground truth (the whiter region means pressed more deep) in the first row of Figure 6c. To meticulously analyze the details, we can also calculate the surface normal map of reconstructed geometry, which can obviously present the high-frequency change of the surface.[41] The normal distribution is consistent with the ground truth and even shows the surface tiny terrain in a real-world tactile process (enlarge the estimated normal map for details). Although the edge has a relative high error in the second row of Figure 6c, which is caused by the unavoidable shadows, the proposed illuminated design can significantly alleviate these effects and ensure the error under a satisfying range.[41]

The optimized sensor exhibits high-performance reconstruction accuracy under various metrics, such as the mean absolute error of depth map (0.02 mm), the MAE of surface normal map (0.28°), as well as the section profile on the horizontal ball center line.

3.3. Performance on Diverse Real-World Objects

3.3.1. Benchmark Surface

To further evaluate and exhibit the performance of the illumination-optimized tactile sensor, we conducted rich real-world experiments. We first fabricate the objects (see in Figure S3a, Supporting Information) of the previously proposed dataset by precise 3D printing (Photo-polymer resin, Kexcelled Inc.), press them by the sensor, and reconstruct the geometry by the proposed approach. The results are shown in **Figure 7**. In each row, there are the original observed tactile image, difference image with reference image, estimated gradient, reconstructed surface normal, and depth map, from left to right. The regular shapes (TSTEP, SCROWN, BALLCAVE, BALLVEX) and the period geometries (WAVE, SINWAVE, BWAVE) are clearly reconstructed, which exhibit the sensor perception ability from shape to details. These structures are representative of daily life and provide a solid foundation for practical application.

Besides, the complex surfaces (PCONVEX, CCONVEX, QUCAVE) are also be used to test the tactile ability on diverse geometry features. The fluctuated and multiscale textures are challenging for tactile. However, the optimized camera-based tactile sensor can clearly reconstruct the detail (subtle elevations and depressions), which evaluates the effectiveness of optimized illumination.

3.3.2. General Surface

We also test the sensor's capability on diverse ordinary-life objects and surfaces (see Figure S3b–d, Supporting Information). We selected typical elements, tools, and objects to test the performance, such as screws, clothes, bobbins, knife hilt, as well as human fingertips. We have classified these objects into four groups (industrial elements surface, regular and period surfaces, planar surface with subtle texture, and complex surface that involves bio-metrics and complex texture.) and utilize the sensor

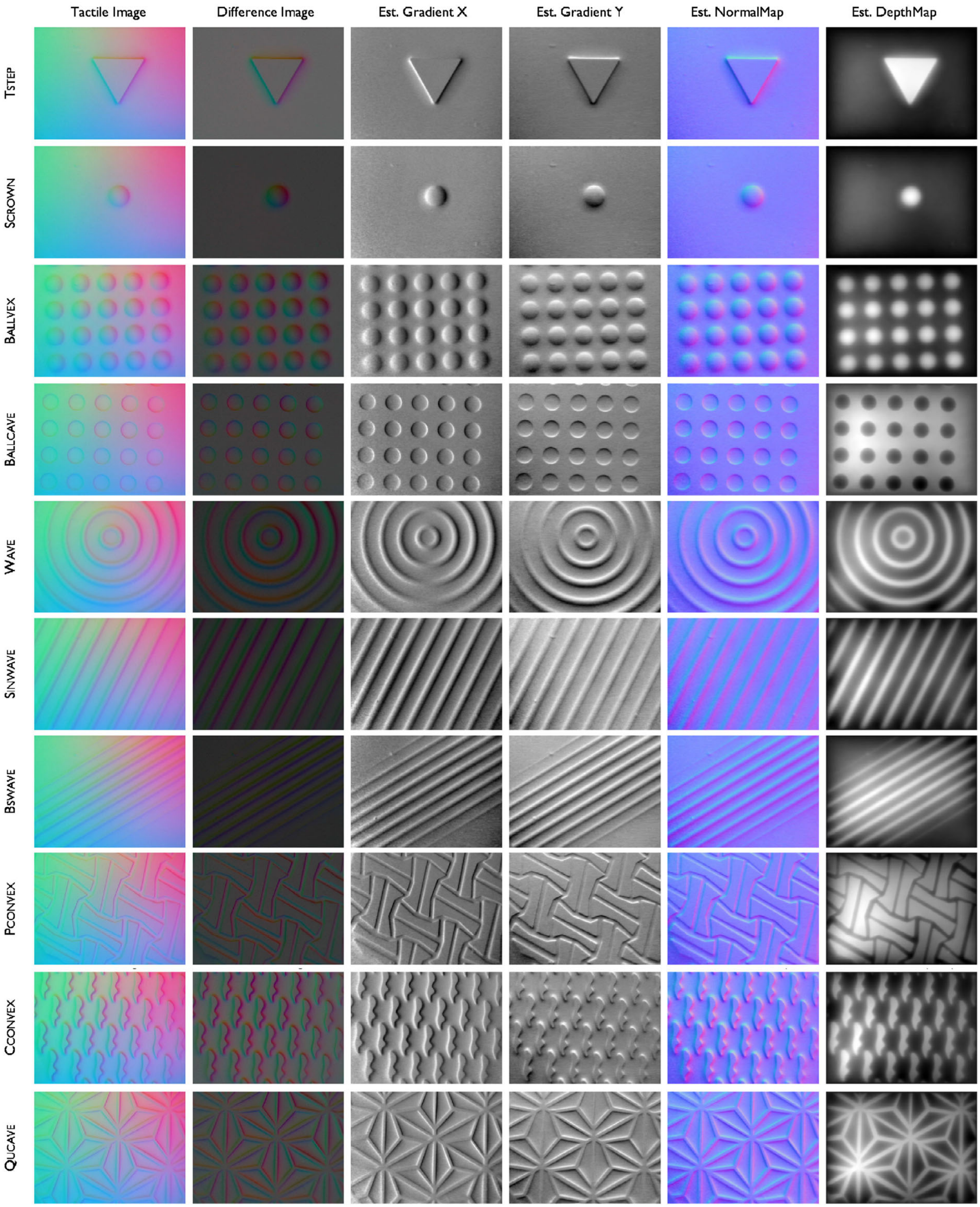

Figure 7. Experiments on real-world objects from proposed tactile dataset. The large-scale shape and small-scale details of each group (simple, regular array, period and complex shapes with concave–convex regions) are clearly reconstructed (best viewed in color and zoom-in).

press on the surface of the target, record the tactile image, and reconstruct their surface geometry, as shown in Figure S4–S7, Supporting Information.

For the industrial elements in Figure S4, Supporting Information, such as screws, connectors, and caps, are widely used in various industries. We select typical elements for tactile experiments to show the perception ability and reconstruction accuracy of the fabricated sensor with the proposed optimized method. The thread of screws and the inscription of the connector can be clear reconstructed, which can provide robots with high-density and precise tactile in industrial grasping and manipulation.

For the regular and periodic surfaces in Figure S5, Supporting Information, such as shape array, non-skid bars, and handle grip, are common daily. They are attached to the surface of bottles, handles, tools, and equipment for robust holding and grasping. Typical patterns are selected to exhibit the perception ability of ordinary life. Detailed tactile perception can help robot to recognize the target and adjust the strategies and methods to interact with them.

For the subtle planar surfaces in Figure S6, Supporting Information, such as reliefs, coins, are popular in a wide range of objects for decoration and information. We conduct experiments on a common target to show the recovery capacity of subtle texture, which shows the potential application on micro-scale tactile and reconstruction, such as archaeology and cultural relief conservation.

For the complex surfaces in Figure S7, Supporting Information, such as human body surface and clothes, are unavoidable in the human–robot interaction. We choose the fingertip print, hair, and beard to exhibit the high-density tactile potential applications on service robots and the tactile involved personal care process. The optimized tactile sensor can support extensive daily service tasks, such as clothes arrangement, sewing, and shaving. This group experiments also show the possibilities of bio-metrics and micro-scale measurement.

4. Discussion

4.1. Effectiveness of Illumination Optimization

Based on the proposed design and optimization method, the distribution, the incident angle, and pose of light source are adjusted to optimum. The observed tactile images are improved with a more uniform appearance, which supports adaptive and robust modeling and helps to achieve high-precision surface geometry reconstruction. Based on the photometric stereo, we establish a solid theoretical framework for systematic design and optimization of camera-based tactile sensors. The proposed design methodology and optimized approach may provide a standard process for camera-based tactile sensors.

4.2. Dataset and Simulation Environment

To achieve highly efficient design and optimization, we propose a tactile surface shape dataset and design metrics for evaluation benchmark, implement a FEM-based mechanical simulation and rendering-based imaging simulation pipeline, and demonstrate the fabrication. The dataset is representative and covers typical surface textures in daily and industrial environments. Furthermore, we develop and combine the simulation of precise deformation and illuminated imaging as a system for camera-based tactile sensor. The simulation system is fast and accurate, which can alleviate the dataset shortage in tactile-related algorithms and decrease time-consuming experiments in fabrication iteration. These evaluations benchmark and assistant components can provide highly efficient tools to the camera-based tactile community for further research. Although there is still an unavoidable sim to real gap, which mainly arises from inaccurate mechanical parameters, insufficient mesh density, the color difference of LED, the variation in surface reflection, and the simple camera response in the simulation system, this gap can be regarded as a systematic bias in simulation. The simulated illuminations and observations can reliably reflect the primary features of the real-world sensor, and this gap does not influence illumination optimization.

4.3. Generalization of Proposed Framework

This work provides generally qualitative insights and quantitative findings for camera-based tactile sensor design, especially for the illumination design and optimization. Our research focuses on reflective camera-based tactile sensors with reflective coatings, which are the most widely applied type. Various variants can apply the proposed optimized approach to their illumination optimization. Based on the photometric stereo theory and the simulation results of this work, a uniformly distributed light setup with inclined incident angle offers superior performance for camera-based tactile sensor. This implies that a (axial or rotational) symmetric design is preferable. It is better to design a slope to host the light source. These insights offer foundational guidance for researchers, especially new entries, embarking on an initial design. With the proposed framework, researchers can take their initial design parameters into this framework to enhance their system's performance. This work not only provide an optimization process of a specific camera-based tactile sensor but also present a distinctive perspective and a systematic approach for the optimization of camera-based tactile sensor in general.

4.4. Wide Range of Potential Applications

Intensive experiments on diverse targets and scenarios show the high-density and accurate tactile performance can open big imaging space for related fields, such as industry, daily life, personal care, and bio-metrics. The proposed illumination optimization method can provide a simple way to design high-performance camera-based tactile sensors for various industries, lower the barrier to entry tactile, and increase the widely application of camera-based tactile sensors, which can provide robots with high-performance tactile with rich physical information of the target, and significantly improve their perception ability in complex grasping, dexterous manipulation and safe human–machine interactions, as well as emotion communication with touch interfaces.

5. Conclusion

Camera-based tactile sensors can provide high-density tactile information. To improve the performance of them, in this work, we propose a systematic method for illumination design and optimization, which significantly determinate the performance of camera-based tactile sensors. We build a benchmark with tactile shape and metrics to evaluate their performance, implement a simulation environment with press deformation and lambertian imaging process, and finally obtain optimized illumination configuration for demonstrate sensor. We fabricate the sensor and model the tactile perception. Intensive experiments on the standard meteorology and diverse target surfaces present the accuracy and adeptness of sensors and verified the effectiveness of proposed illumination optimization.

The proposed method systematically analyses the influence of illumination, provides an effective design and optimization approach for camera-based tactile sensor, and reduces the design complexity of the most important part of the sensor. The standard pipeline can lower the entrance difficulty of the camera-based tactile sensor design. Furthermore, the assistant components of this work will facilitate the community for further research of camera-based tactile sensors.

Although the reconstruction efficiency and geometry distortion still exist in the camera-based tactile sensor, intensive experiments have convincingly demonstrated the application prospects of high-precise and high-density tactile perception. In further research, we will develop algorithms to eliminate the distortion and increase the accuracy of geometry reconstruction, improve the perception frequency, and explore the applications in various industries.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

This article was partly supported by National Natural Science Foundation of China under grant 52025057, Natural Science Foundation of Shanghai under grant 23ZR1435500, Chenguang Program of Shanghai Education Development Foundation and Shanghai Municipal Education Commission under grant 22CGA11, China Postdoctoral Innovation Talents Support Program under grant BX20220204, and China Postdoctoral Science Foundation under grant 2022M722085. The authors also want to thank Yanbing Zhou, Wenqiang Xu, Qianyu Guo, Jin Liu, Rui Li, Shan Luo and Wenzhen Yuan for constructive discussion and technical support, the GelSight and GelSlim team for their open-source technical details, and the openbayes.com for their computing platform.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Jieji Ren: conceptualization (equal); data curation (lead); formal analysis (lead); investigation (lead); methodology (lead); project administration (lead); software (lead); validation (equal); visualization (lead); writing—original draft (lead); writing—review and editing (equal). **Wenxin Du**: software (equal); writing—original draft (supporting). **Yueshi Dong**: validation (supporting); writing—original draft (supporting). **Ningbin Zhang**: funding acquisition (supporting); resources (supporting); visualization (supporting); writing—review and editing (supporting). **Heng Guo**: methodology (supporting); validation (supporting); writing—original draft (supporting). **Boxin Shi**: methodology (supporting); validation (supporting); writing—review and editing (supporting). **Jiang Zou**: conceptualization (supporting); funding acquisition (supporting); validation (supporting); writing—review and editing (supporting). **Guoying Gu**: conceptualization (equal); funding acquisition (lead); resources (lead); supervision (lead); validation (equal).

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Keywords

camera-based tactile sensors, photometric stereo, robot perception, tactile perception

Received: July 15, 2024
Revised: December 10, 2024
Published online:

[1] N. F. Lepora, *Sci. Rob.* **2024**, *9*, eadq1501.
[2] B. Shih, D. Shah, J. Li, T. G. Thuruthel, Y.-L. Park, F. Iida, Z. Bao, R. Kramer-Bottiglio, M. T. Tolley, *Sci. Rob.* **2020**, *5*, eaaz9239.
[3] F. Liu, S. Deswal, A. Christou, Y. Sandamirskaya, M. Kaboli, R. Dahiya, *Sci. Rob.* **2022**, *7*, eabl7344.
[4] C. Bartolozzi, L. Natale, F. Nori, G. Metta, *Nat. Mater.* **2016**, *15*, 921.
[5] C. Hegde, J. Su, J. M. R. Tan, K. He, X. Chen, S. Magdassi, *ACS Nano* **2023**, *17*, 15277.
[6] J. Qu, B. Mao, Z. Li, Y. Xu, K. Zhou, X. Cao, Q. Fan, M. Xu, B. Liang, H. Liu, X. Wang, *Adv. Funct. Mater.* **2023**, *33*, 2306249.
[7] C. Wang, L. Dong, D. Peng, C. Pan, *Adv. Intell. Syst.* **2019**, *1*, 1900090.
[8] S. Gao, Y. Dai, A. Nathan, *Adv. Intell. Syst.* **2022**, *4*, 2100074.
[9] W. Yuan, S. Dong, E. H. Adelson, *Sensors* **2017**, *17*, 2762.
[10] M. K. Johnson, E. H. Adelson, in *2009 IEEE Conf. Computer Vision and Pattern Recognition*, IEEE, Piscataway, NJ **2009**, pp. 1070–1077.
[11] S. Zhang, Z. Chen, Y. Gao, W. Wan, J. Shan, H. Xue, F. Sun, Y. Yang, B. Fang, *IEEE Sens. J.* **2022**, *22*, 21410.
[12] R. Li, R. Platt, W. Yuan, A. Ten Pas, N. Roscup, M. A. Srinivasan, E. Adelson, in *2014 IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, IEEE, Piscataway, NJ **2014**, pp. 3988–3993.
[13] B. Fang, J. Zhao, N. Liu, Y. Sun, S. Zhang, F. Sun, J. Shan, Y. Yang, *Adv. Intell. Syst.* **2024**, *7*, 2400290.
[14] W. Yuan, R. Li, M. A. Srinivasan, E. H. Adelson, in *2015 IEEE Int. Conf. Robotics and Automation (ICRA)*, IEEE, Piscataway, NJ **2015**, pp. 304–311.
[15] W. Yuan, C. Zhu, A. Owens, M. A. Srinivasan, E. H. Adelson, in *2017 IEEE Int. Conf. Robotics and Automation (ICRA)*, IEEE, Piscataway, NJ **2017**, pp. 951–958.
[16] S. Dong, W. Yuan, E. H. Adelson, in *2017 IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS)*, IEEE, Piscataway, NJ **2017**, pp. 137–144.
[17] S. Wang, Y. She, B. Romero, E. Adelson, in *2021 IEEE Int. Conf. Robotics and Automation (ICRA)*, IEEE, Piscataway, NJ **2021**, pp. 6468–6475.
[18] W. Yuan, Y. Mo, S. Wang, E. H. Adelson, in *2018 IEEE Int. Conf. Robotics and Automation (ICRA)*, IEEE, Piscataway, NJ **2018**, pp. 4842–4849.

[19] S. Wang, J. Wu, X. Sun, W. Yuan, W. T. Freeman, J. B. Tenenbaum, E. H. Adelson, in *RSJ Int. Conf. Intelligent Robots and Systems (IROS)*, Madrid, Spain **2018**, pp. 1606–1613.
[20] D. Ma, E. Donlon, S. Dong, A. Rodriguez, in *2019 Int. Conf. Robotics and Automation (ICRA)*, IEEE, Piscataway, NJ **2019**, pp. 5418–5424.
[21] I. H. Taylor, S. Dong, A. Rodriguez, in *2022 Int. Conf. Robotics and Automation (ICRA)*, IEEE, Piscataway, NJ **2022**, pp. 10781–10787.
[22] M. H. Tippur, E. H. Adelson, in *2023 IEEE Int. Conf. Soft Robotics (RoboSoft)*, IEEE, Piscataway, NJ **2023**, pp. 1–8.
[23] N. F. Lepora, Y. Lin, B. Money-Coomes, J. Lloyd, *IEEE Rob. Autom. Lett.* **2022**, *7*, 9382.
[24] J. Zhao, E. H. Adelson, in *2023 IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS)*, IEEE, Piscataway, NJ **2023**, pp. 8979–8984.
[25] I. Andrussow, H. Sun, K. J. Kuchenbecker, G. Martius, *Adv. Intell. Syst.* **2023**, *5*, 2300042.
[26] S. Q. Liu, L. Z. Yañez, E. H. Adelson, in *2023 IEEE Int. Conf. Soft Robotics (RoboSoft)*, IEEE, Piscataway, NJ **2023**, pp. 1–6.
[27] J. Zhao, E. H. Adelson, (Preprint) arXiv:2309.10886 **2023**.
[28] S. Q. Liu, E. H. Adelson, in *2024 IEEE International Conference on Robotics and Automation (ICRA)*, Yokohama, Japan **2024**, pp. 691–697.
[29] B. Romero, H.-S. Fang, P. Agrawal, E. Adelson, in *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Abu Dhabi, United Arab Emirates **2024**, pp. 1853–1860.
[30] A. Agarwal, T. Man, W. Yuan, in *2021 IEEE Int. Conf. Robotics and Automation (ICRA)*, IEEE, Piscataway, NJ **2021**, pp. 1–7.
[31] Y. Ma, A. Agarwal, S. Q. Liu, W. Yuan, E. H. Adelson, in *2024 IEEE 7th Int. Conf. Soft Robotics (RoboSoft)*, IEEE, Piscataway, NJ **2024**, pp. 1068–1074.
[32] D. F. Gomes, Z. Lin, S. Luo, in *2020 IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS)*, IEEE, Piscataway, NJ **2020**, pp. 9903–9909.
[33] B. Romero, F. Veiga, E. Adelson, in *2020 IEEE Int. Conf. Robotics and Automation (ICRA)*, IEEE, Piscataway, NJ **2020**, pp. 4796–4802.
[34] D. F. Gomes, P. Paoletti, S. Luo, in *Robotics: Science and Systems (RSS)* **2023**.
[35] D. F. Gomes, P. Paoletti, S. Luo, *IEEE Rob. Autom. Lett.* **2021**, *6*, 4177.
[36] Z. Si, W. Yuan, *IEEE Rob. Autom. Lett.* **2022**, *7*, 2361.
[37] Y. Zhao, K. Qian, B. Duan, S. Luo, *IEEE Rob. Autom. Lett.* **2024**, *9*, 5647.
[38] S. Wang, M. Lambeta, P.-W. Chou, R. Calandra, *IEEE Rob. Autom. Lett.* **2022**, *7*, 3930.
[39] W. Du, W. Xu, J. Ren, Z. Yu, C. Lu, *IEEE Rob. Autom. Lett.* **2024**, *9*, 2559.
[40] R. J. Woodham, *Opt. Eng.* **1980**, *19*, 139.
[41] F. Wang, J. Ren, H. Guo, M. Ren, B. Shi, in *Proc. IEEE/CVF Int. Conf. Computer Vision,* Paris, France **2023**, pp. 9477–9487.
[42] Z. Chen, S. Zhang, S. Luo, F. Sun, B. Fang, *IEEE Rob. Autom. Lett.* **2023**, *8*, 1239.
[43] M. Li, Z. Ferguson, T. Schneider, T. R. Langlois, D. Zorin, D. Panozzo, C. Jiang, D. M. Kaufman, *ACM Trans. Graphics* **2020**, *39*, 49.
[44] L. Lan, D. M. Kaufman, M. Li, C. Jiang, Y. Yang, *ACM Trans. Graphics* **2022**, *41*, 1.
[45] D. GmbH, Dialux Luminaire Finder - Product Data Sheet: Flex Strip ip65 Digital 30 **2024**, https://luminaires.dialux.com/en/article/ezhz33QUQc-zxRas3_VPaQ (accessed: April 2024).

