Adaptive Reflectance Transformation Imaging : Acquisition, Automation and Stitching.

Luxman Ramamoorthy

To cite this version:

Luxman Ramamoorthy. Adaptive Reflectance Transformation Imaging : Acquisition, Automation and Stitching.. Signal and Image Processing. Université Bourgogne Franche-Comté, 2023. English. NNT : 2023UBFCK092 . tel-04608993

HAL Id: tel-04608993
https://theses.hal.science/tel-04608993

THESE DE DOCTORAT DE L'ETABLISSEMENT UNIVERSITE BOURGOGNE FRANCHE-COMTE

PREPAREE A l'UNIVERSITE DE FRANCHE-COMTE

Ecole doctorale n° 37

Sciences Physiques pour l'Ingénieur et Microtechniques

Doctorat en informatique
Par
Ramamoorthy Luxman

Adaptive Reflectance Transformation Imaging: Acquisition, Automation and Stitching.

Transformation d'Image de Réflexion Adaptative : Acquisition, Automatisation et Assemblage.

Composition du Jury :

Christian Daul , Professeur, ENSEM, Université de Lorraine, Présidente
Robert Sitnik , Professeur, The Institute of Micromechanics and Photonics, Politechnika Warszawska, Rapporteur
Guillaume Caron, Maître de Conférences HDR, laboratoire MIS, l'Université de Picardie Jules Verne, Rapporteur
Jon Yngve Hardeberg , Professeur, Department of Computer Science, NTNU, Gjøvik, Examinateur
Alamin Mansouri, Professeur, Laboratoire ImViA, Université de Bourgogne, Directeur de thèse
Gaetan Le Goıc, Maitre de Conferences HDR, Laboratoire ImViA, Université de Bourgogne, Co-encadrant
Hermine Chatoux, Maitre de Conferences, Laboratoire ImViA, Université de Bourgogne, Co-encadrant
Franck Marzani, Professeur, Laboratoire ImViA, Université de Bourgogne, Co-encadrant

Declaration of Authorship

I, **Ramamoorthy Luxman**, hereby declare that this thesis titled *Adaptive Reflectance Transformation Imaging: Acquisition, Automation and Stitching* and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.

Signed: ______________________________
Date: 20 October 2023

I dedicate this thesis to Appa and Amma, with love and gratitude.

Acknowledgements

First and foremost, I want to thank my professor, Alamin Mansouri, for always helping and guiding me during my research. He trusted me and let me follow my ideas. I am very thankful for his help and support.

I am also deeply grateful to my co-supervisors, Hermine Chatoux, Gaetan Le Goic, and Franck Marzani, for their invaluable insights, expertise, and constructive feedback that have significantly enhanced the quality of my work.

I would like to extend my sincere thanks to my external advisors from the CHANGE project, Jon Hardeberg and Christian Degrigny, for their valuable contributions and feedback that have enriched my research.

I would also like to express my gratitude to the coordinators of the CHANGE-ITN project from NTNU Norway, Sony and Anneli, for their constant support and guidance.

I would like to thank my family and Kaveri for their patience, support, love, encouragement and the immense sacrifices they have done all along my academic journey. Without their support, nothing would have been possible. I am forever grateful for their patience and understanding.

I would like to express my sincere appreciation to Matthieu and Mathilde from the engineering department for their invaluable assistance in building the LightBot systems. Without their help, it would not have been possible to build the system. Thank you for your invaluable contributions.

I extend my appreciation to my colleagues and friends at the lab, Alpha, Abir, Arsalan, Arul, David, Deepak, Deivid, Diana, Marvin, Reda, Sean, and Yuly for their camaraderie, encouragement, and support.

I would also like to thank my fellow ESRs in the project, Amalia, Agnese, Dipendra, Evi, Athanasia, Sunita, Yoko, Jizhen, Silvia, Jan, Alessandra, and Deepshika, for their invaluable support, inspiration, and encouragement.

Lastly, I thank all those who have played a part in making this research possible, and those who have touched my life in one way or another.

“The more I learn, the more I realize how much I don’t know.” - Albert Einstein

Contents

List of Figures

List of Tables

Introduction

Project context

The EU Marie Curie CHANGE-ITN project (Cultural Heritage Analysis for Next Generation, https://change-itn.eu/) is a multi-disciplinary research initiative aimed at addressing the challenges in studying the changes on cultural heritage artifacts over time. The project brings together leading academic institutions and research organizations from across Europe to develop methodologies for assessing and monitoring changes to cultural heritage artifacts.

The project's main objective is to integrate conventional investigation tools used by cultural heritage experts with new technologies, with the aim of creating a new generation of multi-skilled experts specialized in the study and monitoring of changes to cultural heritage artifacts. This innovative approach involves the selection and training of Early Stage Researchers (ESRs) with different scientific backgrounds relevant to the project, who will collaborate on research projects aimed at reaching the project's main objective.

This thesis is focused on a specific aspect of the EU CHANGE-ITN project - use of an imaging technique called Reflectance Transformation Imaging (RTI) for the analysis of cultural heritage artifacts. The research involves the development and implementation of RTI techniques for the analysis of changes on cultural heritage artifacts and contribute to the project's goal of creating a unified methodology for assessing and monitoring changes to cultural heritage artifacts. The ultimate goal is to provide digital tools for the monitoring of changes in CH artifacts to support conservation efforts. The research have been carried out in collaboration with CH experts and stakeholders from 8 EU countries. As a part of this project, I was trained as an early stage researcher in the assessment of changes in CH objects, with a focus on optimising the capturing and analysis of data to ensure better documentation and long-term preservation of CH.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 813789.

Research overview

Reflectance Transformation Imaging (RTI) is a digital imaging technique that captures the way a surface reflects light coming from different angles. It is commonly used to study cultural heritage artifacts, such as ancient manuscripts, coins and sculptures, because it can reveal detailed surface features that may not be apparent under ambient light. RTI works by capturing a series of images of an object being illuminated from different points on a hemisphere or dome. These images are then combined using specialized software to create a single, high-resolution image that encodes the surface's reflectance properties. The resulting RTI image can then be interactively visualized on a computer, allowing the user to adjust virtually the lighting direction to highlight different features of the surface.

Acquisition, modeling, and visualization of reflectance of complex surfaces is still an active area of research in the RTI field. Complex surfaces can include objects with varying size, shape, material properties, which require specialized and adaptive techniques to accurately capture their reflectance properties. In this thesis, we have focused on addressing the challenges in realizing surface adaptive RTI and automating the acquisition process. We have developed methods for estimating the optimal light configuration for capturing RTI sequence adaptive to the surface being digitized. We also developed methods for stitching together multiple RTI data sets and thus improve the resolution of the RTI data. These methods are developed to improve the accuracy and efficiency of surface adaptive RTI, and to bring advances in the field of digital imaging for cultural heritage and applications.

In the current state of the art, RTI acquisitions are typically carried out by manual placement of a light at different directions (free form) or use of RTI domes with fixed light positions or mechanized dome with movable light source to capture a series of images. Manually positioning the light source is a time-consuming process and lacks accuracy, repeatability. RTI domes are efficient and more reliable, however they are limited to acquisition of smaller sized objects only. To address the limitations pertaining to free form and the dome systems, we investigated the use of robotic arm and automation to streamline the RTI acquisition process. This involves the use of robotic arm to position the light source, use of a XY stage to position the surface as well as automated image capture systems. There are several benefits to automating RTI acquisition. One advantage is the ability to capture RTI images of large surfaces that are generally difficult (or impossible) to acquire using RTI domes.

There are several challenges associated with the automation of RTI acquisition process using robotic arm and XY platform such as building the control systems that can accurately and reliably position the light aligning it to the required angles, collision avoidance in robotic arm planning, integration of these systems into a cohesive and user-friendly workflow, ensuring that the resulting RTI images are of high quality and meet the needs of the user. We studied these challenges in our work, built a fully functional novel robotic arm-based acquisition system and demonstrated the advantage of this system over the other existing systems.

Thesis overview

In this manuscript, we have organized the content by introducing in chapter 1 the concepts associated to reflectance. Then, we discuss the RTI technology on which our work is focused on. This sets the stage for the rest of the thesis. Then our contributions are presented in four chapters, each covering a separate research problem related to RTI addressed in these works. Each of these chapters follows a similar structure, starting with an introduction/ review of the problem and then presenting the experiments, results, and conclusions for the proposed solutions.

In chapter 2 we present a robotic acquisition system named *LightBot* for automating the process of capturing RTI images, that we built and demonstrated in this Ph.D. In chapter 3 we present a framework for stitching RTI data, which allows to combine multiple RTI data into a single, comprehensive data. In Chapter 4, we present the *NBLP-RTI benchmark dataset*, which provides a standard set of data for evaluating the good light configuration for RTI surfaces. In Chapter 5, we present the NBLP methods developed and tested in my Ph.D. Finally, we conclude the thesis with a summary of the main contributions and findings presented in the previous chapters. We also discuss the implications of these results for future research in the field, and outline some potential directions for future work. We also reflect on the challenges and lessons learned during the course of our research, and provide some concluding thoughts on the broader implications of our work.

In the appendix , we provide the supplementary material associated with this thesis including the detailed information on the mechanical design, API for the LightBot, and the blender plugin called *surfaceAdaptive-RTI* that we built for development of methods in virtual space.

CHAPTER 1

Reflectance Imaging: State of the art

Chapter overview

Reflectance is a fundamental concept in the field of imaging and relates to the way surfaces reflect light. It is a key factor in determining the appearance of an object, and is intrinsically associated with the surface itself. Understanding surface reflectance is important for a wide range of applications like quality inspection, surface characterization, graphics and 3D shape recovery. In this chapter, we provide an overview of the fundamental concepts of reflectance. We then introduce the RTI technique which is central to our work. We discuss the factors that can affect surface reflectance and the methods that are used to measure it, as well as the challenges and opportunities presented by this field of study. Finally, we introduce the RTI acquisition process and explain the need for automation in the acquisition process.

Interaction of light with matter is described by different phenomenon such as reflectance, scattering, absorption and transmittance. Transmittance is the passage of light through a medium and it describes how much light is transmitted from a surface. Absorptance of a medium is defined by the ratio of absorbed radiant power to incident radiant power. Reflectance is a property of surfaces that describes how the light rays are reflected. It is determined by the composition, structure, and surface properties of the material, the angle and wavelength of the incoming light. Scattering is the process of waves getting deviated due to certain anomalies in the medium. Scattering takes place due to the wave nature of matter whereas reflection is a consequence of the particle nature of matter. Reflectance is an essential property of surface appearance. There are many applications that rely on surface reflectance that span a wide range of fields. Some examples of these applications as illustrated in Figure.1.1 include computer graphics David F [1991], Amanatides [1987], 3D imaging Zhang et al. [1999], remote sensing Nansen and Elliott [2016], Govender et al. [2007], material science Philips-Invernizzi et al. [2001], cultural heritage conservation Tsai [2007], Huang et al. [2016], medicine Elangovan and Jeyaseelan [2016], photometry Bessell et al. [2005], color science Byrne and Hilbert [2003], or manufacturing Zendagui et al. [2019]. These applications rely on the ability to accurately characterize and analyze surface reflectance in order to gain insights into the properties of materials and their interactions with light. In this thesis, our focus is on techniques that utilize surface reflectance for applications in the field of cultural heritage. Imaging techniques are often favored in the cultural heritage for authenticity, provenance, and conservation purposes due to their non-invasive nature. These techniques allow for the examination and analysis of artifacts without causing any damage or alteration to the original object. This makes them particularly useful for authenticity assessments, as well as for studying the manufacturing techniques and materials used in the creation of the object. In addition, imaging techniques can be used to monitor and track the condition of cultural heritage items over time, helping to identify any changes or degradation that may occur and allowing for the development of conservation strategies to preserve the item. The use of imaging techniques in cultural heritage offers many benefits and became an important tool in the field of cultural heritage conservation.

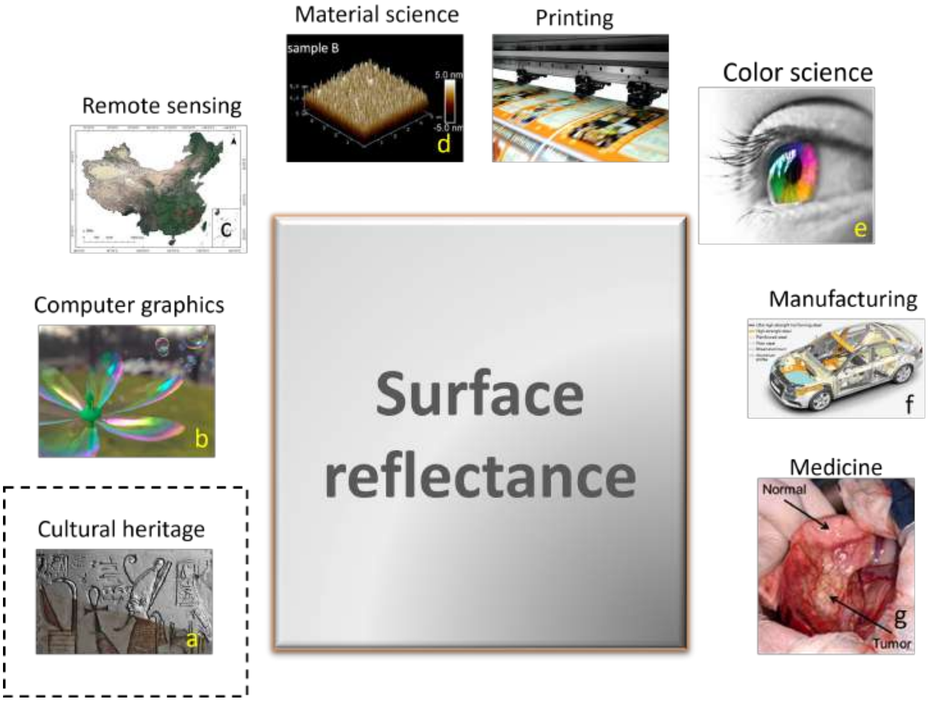


Figure 1.1: Applications that relies on surface reflectance.
Copa-Hypotheses [2019][a], Yoshiki Kaminaka and Kaneda [2022][b], Zhou et al. [2022][c], Chung et al. [2017][d], RTP-Company[e], Berladir et al. [2017][f], Holt et al. [2014][g],

1.1 Reflectance: Basic concepts

When light hits a surface, a part of the energy is absorbed and a part of it is transmitted, a part is reflected. The amount of light that is reflected depends on the surface properties of the material, as well as the angle and wavelength of the incoming light. The reflection of light from a surface influence how materials look and interact with the surroundings. There are several literature that provide a comprehensive introduction to surface reflectance concepts, such as Nayar, Fleming et al. [2003]. In the following section, we draw upon information from these literature to provide a comprehensive overview of the concepts of reflectance.

When a light ray hits an object, generally it can be reflected in different ways such as specular, diffuse and gloss reflection. The mechanism of specular and diffuse reflection is illustrated in Figure.1.2. Specular reflection, also known as surface reflection, occurs when light is totally reflected at a specific angle off a surface. This type of reflection is characterized by a highly directional, mirror-like reflection. It occurs when light bounces off a shiny surface, such as a mirror or a polished metal surface. Diffuse reflection, also known as body reflection, occurs when light is reflected equally in all directions off a surface. This type of reflection is characterized by a more even and diffused reflection, rather than a highly directional one. It occurs when light bounces off a matte surface, such as a white sheet of paper or a rough concrete wall.

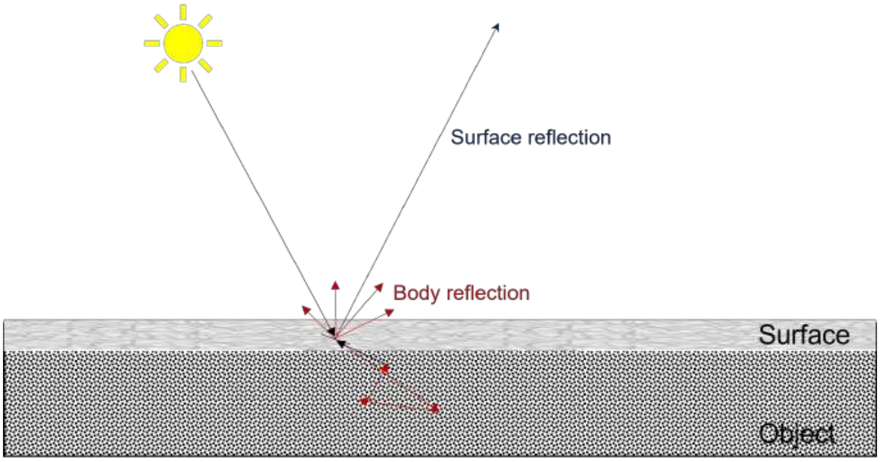


Figure 1.2: Mechanism of reflection

The appearance of a surface can change depending on the specular and diffuse components of its reflection. The contributions of these two components majorly influences the overall appearance of the surface, as shown in Figure.1.3. A surface with a higher proportion of the specular component will appear glossy, while a surface with a low specular component and a high diffuse component will appear matte. A surface with a balance of both components will appear semi-glossy. Light reflected by mirror like polished surface contains fully specular component. Reflectance lobes are visible features in the shape of a reflectance graph. There are different types of reflectance lobes, including specular, diffuse, and semi-diffuse lobes. Specular lobes represent a tightly concentrated, bright highlight, while diffuse lobes represent a more spread-out, soft/rough highlight1 as illustrated in Figure.1.4.

(a) Gloss (b) Semigloss (c) Matte

Figure 1.3: Example of gloss, semi gloss and matte surface reflections

Radiometry is a branch of science that deals with the measurement and analysis of electromagnetic radiation, such as visible light, ultraviolet, and infrared Kirkpatrick [2005]. It uses specialized instruments like radiometers and spectroradiometers to measure the intensity, wavelength, and other characteristics of electromagnetic radiation McCluney [2014]. Radiometric techniques in optics characterize the distribution of the radiation's power in space in entire optical radiation spectrum while *photometric* techniques, is limited to visible light's interaction with the human eye Hébert et al. [2015]. The reflectance properties of a surface depend on its material properties.

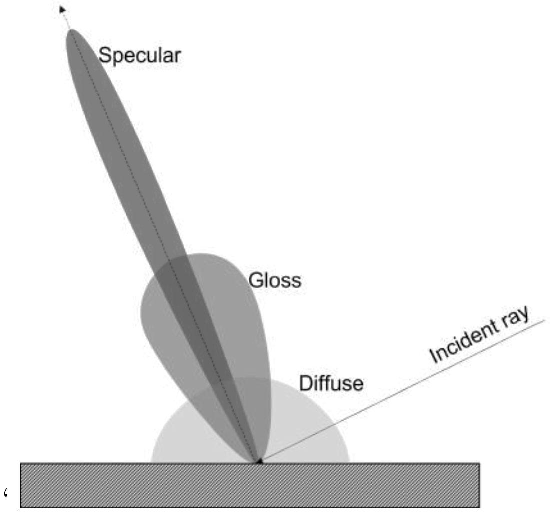


Figure 1.4: Components of reflectance

In order to quantitatively understand the reflectance property, various physical measurements are used that are briefly explained below. These quantities, in turn, are used to calculate the Bidirectional Reflectance Distribution Function (BRDF), which serves as a representative of the surface's reflectance property.

Light flux, also known as radiant flux, is a measure of the power emitted per unit solid angle. The light flux can be used to describe the total amount of electromagnetic radiation emitted by a light source, or it can be used to describe the amount of electromagnetic radiation that is received by a surface. Figure.1.5 illustrates the light flux and the solid angle subtended by the light source over the surface, where $d\omega$ is the unit solid angle, dA' is the foreshortened area (a foreshortened area refers to the perceived reduction in the area of a surface patch as seen from a particular point), r is the distance of the light source from the object, dA is the unit surface area, θ is the incident angle. Reflectance can be characterized by two quantities called surface irradiance and surface radiance.

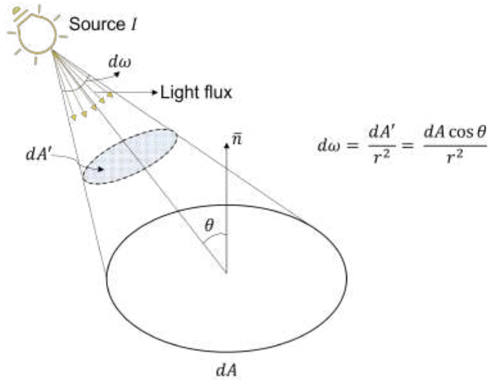


Figure 1.5: Illustration of light flux- power emitted per unit solid angle, $watts/steradians$

Surface irradiance is the measure of illumination of a surface. It is measured as the light flux incident per unit surface area. It is obtained by the Eq.1.1.

$$E = \frac{d\Phi}{dA} = I\frac{d\omega}{dA} = \frac{\frac{IdAcos\theta}{r^2}}{dA} = \frac{Icos\theta}{r^2} = \frac{1}{r^2}(\overline{n}\cdot\overline{s}), \tag{1.1}$$

$$E = \frac{Icos\theta}{r^2}, \tag{1.2}$$

where, E is the surface irradiance, Φ is the light flux, dA is the unit surface area, I is the radiant intensity of the light source, $d\omega$ is the unit solid angle, θ is the incident angle and $\bar{n}$ is the surface normal.

Surface radiance is the measure of light flux emitted per unit foreshortened area (the area shaded in the figure) per unit solid angle. Figure.1.6 illustrates the surface radiance quantity. It is defined by the Eq. 1.3.

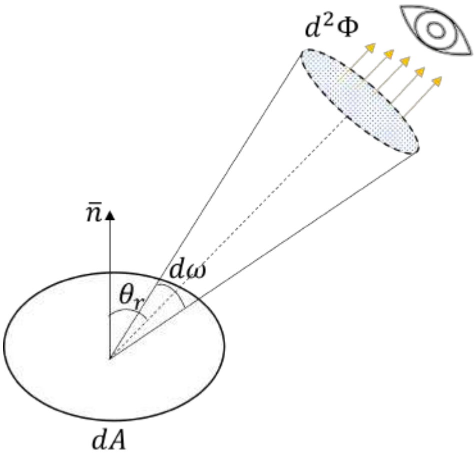


Figure 1.6: Illustration of surface radiance - the light flux emitted per unit area per unit solid angle.

$$L = \frac{d^2\Phi}{(dAcos\theta_r)d\omega} \tag{1.3}$$

where, L is the surface radiance, Φ is the light flux, θ_r is the reflectance angle.

Bidirectional Reflection Distribution Function (BRDF) is a measure of how much light is reflected by a surface as a function of the incident and reflected light directions Montes and Ureña [2012], Bartell et al. [1981], Voss et al. [2000], Marschner et al. [2000]. It is widely used in fields such as computer graphics Kurt and Edwards [2009], Ngan et al. [2005]. The BRDF of a surface is typically defined as the ratio of the reflected radiance to the incident irradiance. For a surface illuminated by a light source and observed by a camera as shown in Figure. 1.7, the BRDF is defined by Eq.1.4:

$$\text{BRDF},\ f(\theta_i, \phi_i, \theta_r, \phi_r, \lambda) = \frac{L(\theta_r, \phi_r, \lambda)}{E(\theta_i, \phi_i, \lambda)} \tag{1.4}$$

where, L and E are the surface irradiance and radiance respectively as defined previously λ is the wavelength of the incident ray, (θ_i, ϕ_i) and (θ_r, ϕ_r) are the azimuth and elevation angles of the incident and reflected rays respectively. Note that the Bidirectional Reflectance Distribution Function (BRDF) may not exhibit consistent behavior across different wavelengths of the incident light. This means that the reflectance characteristics can change depending on whether the incident light is blue, infrared, or another color.

It is important to note that the BRDF exhibits the Helmholtz reciprocity property Clarke and Parry [1985], Potton [2004], which states that if the positions of the camera and light source are reversed, the BRDF remains unchanged. Also, the total reflected power for a given direction of incident radiation is less than or equal to the energy of the incident light. This means that a surface cannot reflect more light than it receives. These properties of the BRDF are useful for understanding and predicting the reflection of light from a surface under different lighting conditions and viewing directions.

There are various analytical models that have been developed to describe the BRDF of materials. These models can be divided into two categories: empirical models, which are based on observations and measurements, and physically-based theoretical models, which are based on the underlying

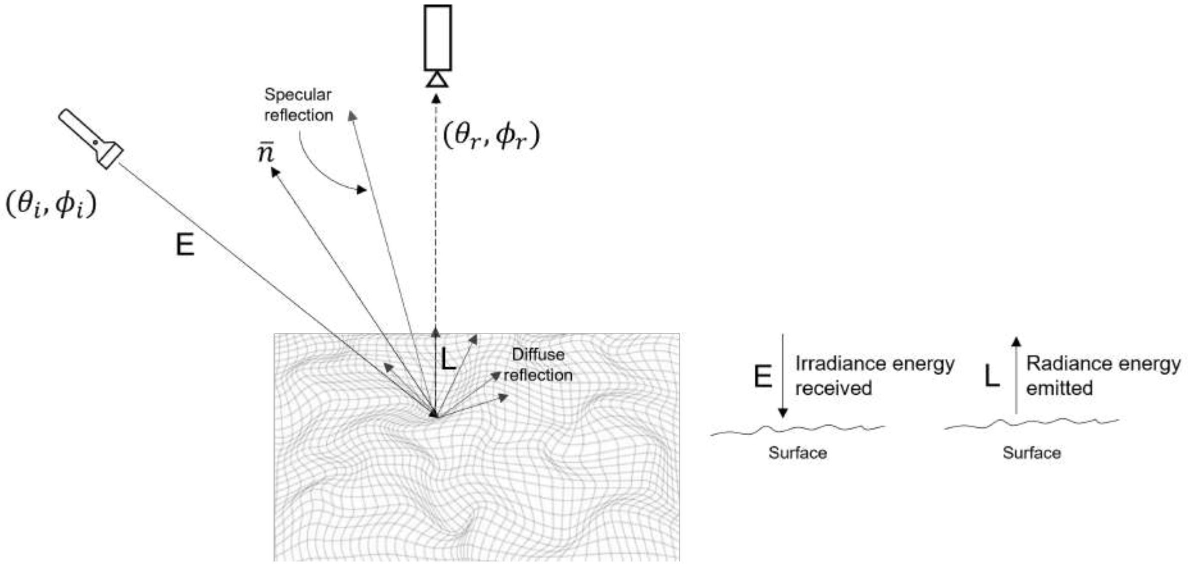


Figure 1.7: Reflectance of a material.

physical properties of materials. Both models are approximations of the reflectance properties of materials. Empirical models are often derived from statistical analysis of measured data, and are useful for approximating the BRDF of surfaces when no detailed knowledge of their material properties is available. These models are typically simpler and easier to use than physically-based models, but are less accurate and may not be suitable for all materials. Physically-based models are derived from fundamental principles of optics and material physics, and aim to accurately model the BRDF of any surface material and geometry. These models are often more complex and computationally expensive than empirical models, but can provide more accurate results over a wider range of materials and geometries Dana et al. [1999]. The BRDF of a surface can be used to predict the appearance of an object under different lighting conditions and viewing directions. Ideal diffuse BRDF describes the reflection of light from a perfectly diffuse surface such as a white wall. An ideal specular BRDF describes the reflection of light from perfectly specular surface such as mirror; Cosine lobe BRDFs use cosine lobes, which are curves defined using the cosine function, to account for the haze component of reflection (the haze component is the reflection that occurs between the specular (mirror-like) and diffuse (evenly scattered) reflections.). Schlick [1994] discusses the wide range of BRDF models that have been proposed over years, each of which is based on different assumptions and principles. In the following section, we will review some of the commonly used models. Chart in Figure.1.8 gives an overview of the popularly used BRDF models and their classification.

1.2 Reflectance models

BRDF models are mathematical models that describe the way in which light interacts with a surface. These models are used to predict the appearance of surfaces under different lighting conditions and to understand the physical properties of materials.

Lambertian reflectance model Oren and Nayar [1994], Koppal [2020], Yuille [2012] is named after Johann Heinrich Lambert, a 18th century German scientist who was one of the first to study the reflection of light. The key assumption of the Lambertian model is that the surface is perfectly diffuse, meaning that it reflects light equally in all directions. The opposite case is a specular surface, which reflects light at a specific angle. The reflectance of a Lambertian surface is described by a single parameter, the diffuse reflectance coefficient, which determines the intensity of the reflected light. The Lambertian model is often used to describe the reflectance of matte surfaces, such as paper or painted walls. The model is not suitable for describing the reflectance of glossy or

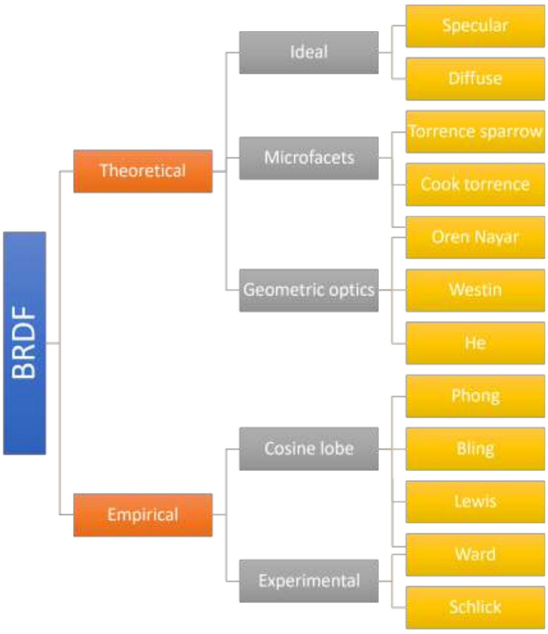


Figure 1.8: Popularly used BRDF models

specular surfaces, as it does not take into account the specular component of the reflection. It is not very accurate in describing the reflectance of real-world surfaces, as most surfaces are not perfectly diffuse. Despite its limitations, the Lambertian model remains a widely used and useful tool for analyzing and simulating the appearance of matte surfaces. The Lambertian reflectance model is described by the following equation:

$$R = k_d \times L \cdot N, \tag{1.5}$$

where R is the reflectance of the surface, k_d is the diffuse reflectance coefficient, L is the incident light energy - a vector that represents the direction and intensity of the incoming light and the surface normal, N, is a vector that represents the orientation of the surface. The diffuse reflectance coefficient is a scalar value that determines the intensity of the reflected light.

The dot product of L and N, $L \cdot N$, is a scalar value that represents the angle between the incident light and the surface normal. The reflectance of the surface is proportional to the dot product of L and N, with a larger dot product resulting in a higher reflectance.

Specular BRDF model is used for describing the reflection of light off ideal specular surfaces. Figure.1.9 illustrates the reflection of light off a mirror. The mirror BRDF model in this case is defined by Eq.1.6 Oren and Nayar [1994]

$$f(\theta_i, \phi_i, \theta_r, \phi_r) = \frac{\delta(\theta_i - \theta_r)\delta(\phi_i + \pi - \phi_r)}{cos\theta_i sin\theta_i}, \tag{1.6}$$

where, f is the BRDF, θ_i, ϕ_i are the azimuth and elevation of the incident ray, θ_r, ϕ_r are the azimuth and elevation of the reflected light ray.

For a perfect mirror, the viewer (camera) receives light only when the viewer angle is same as the reflectance angle, i.e $\bar{v} = \bar{r}$.

Micro-facet models are more complex approach for describing the reflectance first introduced by Cook and Torrance [1982]. While the Lambertian model is simple and easy to use, it is not very accurate in describing the reflectance of rough surfaces, as it does not take into account the microstructure of the surface. The key assumption of the microfacet model is that the surface is composed of a large number of small, flat facets, each of which reflects light in either totally a specular (mirror-like) manner or totally diffuse. The reflectance of the surface is then determined by the distribution and orientation of these facets. The microfacet model is a general model that can be used to describe the reflectance of a wide range of rough surfaces, including surfaces with

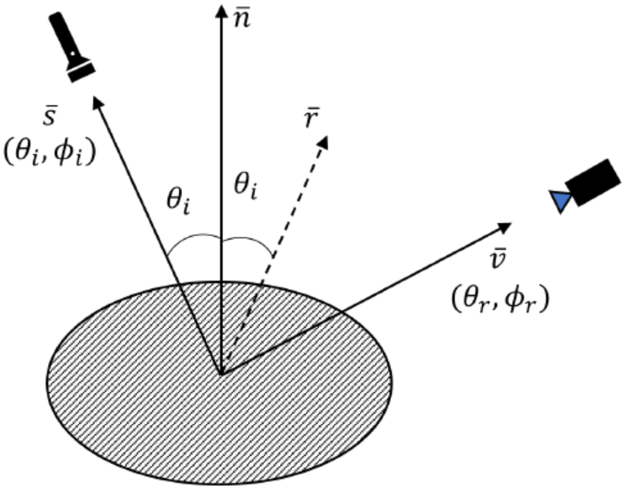


Figure 1.9: Ideal specular reflection

different roughness scales and surface topographies. This model is widely used in computer graphics to create realistic lighting effects and to simulate the appearance of rough surfaces.

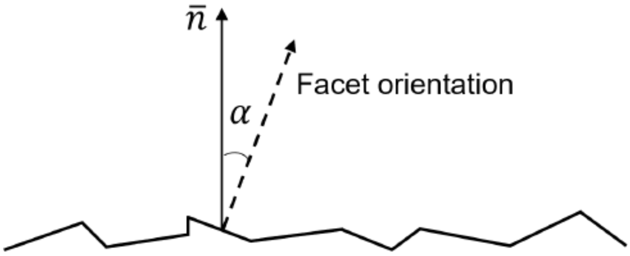


Figure 1.10: Microfaceted surface. $\bar{n}$ is the mean orientation of the facets (surface normal).

The distribution of facet orientations in a rough surface can, for example, be approximated using a Gaussian distribution, which results in a more realistic rendering of the appearance. Figure. 1.10, represents the reflection of a light ray from a micro-facetted rough surface. In this case, the distribution of the facet orientations on the rough surface, p can be modeled using Eq.1.7:

$$p(\alpha, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-\alpha^2}{2\sigma^2}}, \tag{1.7}$$

where σ is the roughness parameter (standard deviation) and α is the orientation of the facet. When $\sigma \to 0$, the surface is perfectly flat and behaves like a mirror. On the other hand, when $\sigma = 1$, the surface is perfectly diffuse. The Gaussian distribution is a simple and effective way to represent a micro-faceted rough surface model. The Beckmann normal distribution Walter et al. [2007] is another popularly used distrubution for micro-facet reflection models. There are several different types of microfacet surface reflectance models each having its own unique set of assumptions and parameters, and they are used to model different types of surfaces and lighting conditions.

Torrance-Sparrow model Knox [1989], Torrance and Sparrow [1967] is based on the assumption that each microfacet on the surface behaves as a mirror reflector, and it uses Fresnel theory Judd [1942] to predict the amount of light that is reflected specularly towards the viewer. The Torrance-Sparrow model consists 1. a facet distribution function, which determines the proportion of facets on the surface that are oriented in the ideal specular direction, 2. a geometric attenuation factor, which takes into account the shadowing of one facet by another, 3. a Fresnel term, which determines the proportion of light that is reflected from those facets that are pointing in the ideal specular direction. It is given by Eq.1.8:

$$f(\bar{s},\bar{v}) = \frac{\rho_s}{(\bar{n}\cdot\bar{s})(\bar{n}\cdot\bar{v})} p(\alpha,\sigma) G(\bar{s},\bar{n},\bar{v}), \tag{1.8}$$

where, ρ_s is the reflectivity of each facet, $\bar{s}$ is the light source direction, $\bar{v}$ is the viewer direction, $p(\alpha,\sigma)$ is the surface roughness distribution, $G(\bar{s},\bar{n},\bar{v})$ is the geometric attenuation factor (masking, shadowing).

Phong reflectance model was developed by Bui Tuong Phong in the 1970s and is a widely accepted model Tan [2020]. The model describes the way a surface reflects light as a combination of diffuse reflection and specular reflection. The Phong model includes several components that contribute to the overall lighting of a surface, including ambient light, diffuse reflection, and specular reflection. Ambient light represents the indirect light present in a scene, and is calculated based on the color and intensity of the light sources in the scene. Diffuse reflection is calculated based on the angle between the surface normal and the light source, while specular reflection is calculated using the Phong specular reflection model, which sets the intensity of specular reflection proportional to the cosine of the angle between the surface normal and the viewer's line of sight. In addition to these components, the Phong reflection model also includes parameters for each light source and material in the scene, which are used to calculate the illumination of each surface point. While the Phong model is widely used for its simplicity and ability to produce realistic-looking images, it has been superseded by more advanced models such as the Cook-Torrance model and the GGX model.

Cook and Torrance model is a widely used method for simulating the reflection of light from rough surfaces in computer graphics. It is based on geometrical optics theory and was developed by Cook and Torrance in 1982 Cook and Torrance [1982]. In this model, the reflection from a surface is described as a combination of diffuse and specular reflection, with the parameters k_d and k_s controlling the fraction of energy that is diffusely or specularly reflected 1.9.

$$f = f_{diff} + f_{spec} \tag{1.9}$$

where, f is the BRDF, f_{diff} and f_{spec} are the diffuse and specular components of the reflectance respectively.

The diffuse reflection is modeled using a classical Lambertian reflection, while the specular reflection is a combination of the Fresnel term and the microfacet distribution and geometric attenuation factors, which account for the masking and shadowing of the microfacets on the surface. The Fresnel term here describes the amount of light reflected by the material, as opposed to the light that is absorbed or transmitted by the polished smooth facet. During the interaction of light with microgeometry, shadowing and masking are significant phenomena that occur. Shadowing happens when the microgeometry obstructs the light source, while masking occurs when one microfacet is obscured from view by another microgeometry (refer to Figure.1.11). The specular term is described by the equation

$$f_{spec}(\omega_i,\omega_o) = \frac{F}{\pi}\frac{GD}{(n\cdot\omega_i)(n\cdot\omega_o)}$$

where, F is the Fresnel term, G is the geometrical attenuation factor, and D is the normal distribution function that represents the fraction of facets that are oriented in the half way vector h.

The basis of the Cook-Torrance model was already established in the prior work of Torrance and Sparrow. However, the key contribution of Cook and Torrance was their proposition that only microfacets that align with h contribute to the specular reflection. One disadvantage of this model is that the parameters are not intuitively set and require experimentation to find good values, and it can also produce results that do not obey the energy conservation law for certain angles.

GGX Model was introduced by Walter et al. [2007] and has since become one of the most popular models for representing rough surfaces in photorealistic rendering. It introduced BSDF (Bidirectional Scattering Distribution Function). BSDFs encompass both BRDF and BTDF (Bidirectional Transmittance Distribution Function), and model both reflection and transmittance. The

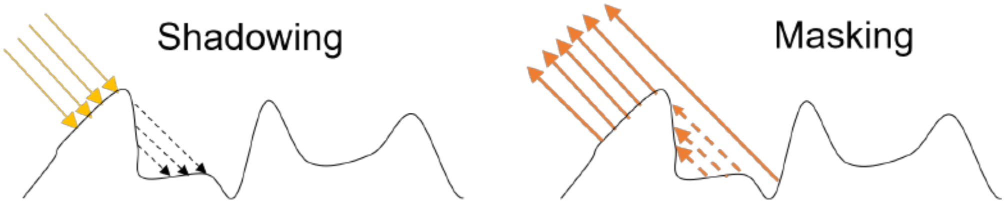


Figure 1.11: The image on the left illustrates shadowing, which is the occlusion of the light source by microgeometry. On the right, the image depicts masking, which is the visibility occlusion of microgeometry.

GGX BRDF is largely influenced by Cook and Torrance's microfacet model. It is expressed as given in Eq.1.10:

$$f(\omega_i, \omega_o) = \frac{FGD}{4(n \cdot \omega_i)(n \cdot \omega_o)}, \tag{1.10}$$

This is identical to the Cook-Torrance specular BRDF, with the exception that the normalization value for the Fresnel term has been adjusted to 4 (as opposed to π).

Measuring actual BRDF

Theoretical models, which are based on the known material properties of a surface, can be applied to predict and simulate the reflectance of that surface. However, in the absence of this information, it is necessary to measure the actual BRDF of a surface in order to accurately model its reflectance.

Actual BRDF measurement can be accomplished using highly specialized instruments such as gonioreflectometers, which are designed to measure the reflectance properties of a material from different angles. These measurements can be used to better understand the optical properties of a material and to improve the accuracy of theoretical models for simulating its reflectance. The measurement of BRDF using the instrument involves making a dense acquisition of a sample point of the surface. As mentioned earlier, the BRDF is a function of 4 dimensions and the wavelength parameter, so the gonioreflectometer typically has 4 degrees of freedom in its mechanical arrangement. Ensuring light source and camera stability, compensating for variations in surface geometry, and dealing with inter-reflection during measurements is a challenge. Therefore, it is not popular to measure BRDF directly using this method. An example of a gonioreflectometer is shown in Figure.1.12. Another approach to approximating the bidirectional reflectance distribution function (BRDF) is to use computational photography techniques that involve capturing images of a surface under different illumination conditions but keeping the camera stationery. They are used to extract visual surface properties such as shape and detail. This group of techniques, known as Multi Light Image Collections (MLIC) techniques Pintus et al. [2019], Dulecha [2021], are commonly used for applications involving surface visualization and appearance analysis, and capturing the global shape of an object Yang and Ahuja [2012], Ackermann et al. [2015]. These MLIC approaches such as Light Field Imaging Ihrke et al. [2016], Wu et al. [2017], Photometric stereo Woodham [1980], Structured light Geng [2011], Ribo and Brandner [2005], Reflectance Transformation Imaging Malzbender et al. [2001b], Pitard et al. [2015] are useful for analyzing and visualizing the surface properties of an object. RTI is the MLIC technique that simplifies the BRDF to model the reflectance of a surface by varying only the light position and keeping the camera fixed at a particular position.

1.3 Reflectance Transformation Imaging

Reflectance Transformation Imaging (RTI) is a computational photographic technique that captures the surface properties of an object, including its shape, color, and texture Imaging. It involves taking multiple images of an object using a stationary camera and varying illumination direction. The resulting images are used to create a high-resolution representation of the object's surface's

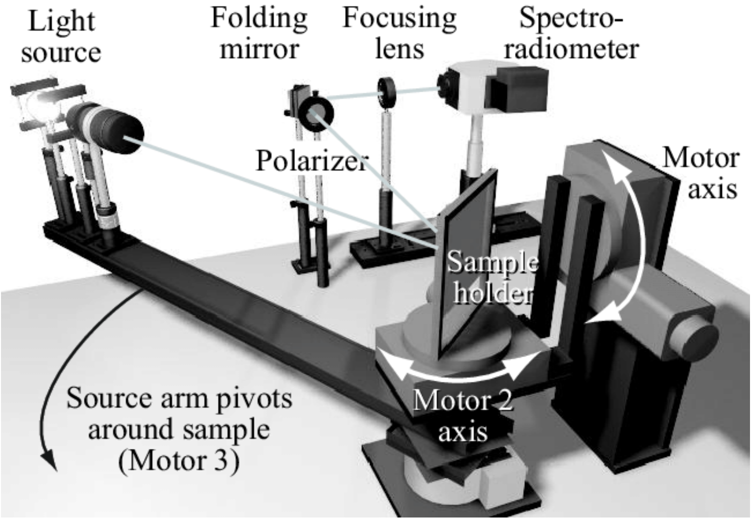


Figure 1.12: An example of BRDF measuring instrument. This is a design published by Li et al. [2006]

appearance and shape, that can be interactively rendered for any light direction. RTI has its roots in the early 20th century, when researchers began studying the way light interacts with surfaces and how this information can be used to reconstruct the surface geometry. In the 1980s, Silver [1980] developed a technique called Reflectance Mapping, which used multiple images of a surface taken under controlled lighting conditions to create a 3D model of the surface. This technique laid the foundation for the development of RTI. In the early 2000s, researchers at the University of Washington and Hewlett-Packard Laboratories developed PTM, which improved upon Reflectance Mapping by using polynomial functions to model the surface reflectance Malzbender et al. [2001b]. This has become widely used in a variety of fields, including cultural heritage conservation, product design, and computer graphics. Today, RTI is a powerful tool for capturing the surface properties of objects. It continues to be an important area of research and has numerous practical applications especially cultural heritage and conseration.

As depicted in Figure.1.13, the RTI process consists of three steps: the acquisition, the modeling, and the visualization. The acquisition phase involves capturing multiple images of an object under controlled lighting conditions. The modeling phase involves fitting the acquired data and creating an RTI model containing the surface information. Finally, the visualization phase involves using the RTI model to interactively re-light the object from different directions.

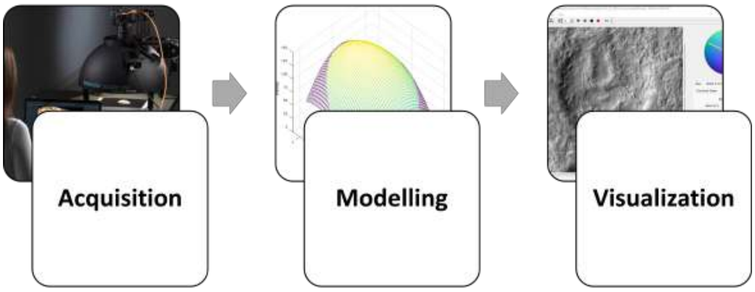


Figure 1.13: The three steps of Reflectance Transformation Imaging

1.3.1 RTI acquisition

RTI acquisition systems differ from BRDF measurement systems in that they have fewer degrees of freedom in their mechanical arrangements. While BRDF systems have four degrees of freedom, allowing for the variation of light position, camera position, and surface position during acquisition, RTI systems only have two degrees of freedom, with only the light source direction (in terms of azimuth and elevation) being varied during acquisition. The camera and surface positions remain unchanged. This simplicity in setup makes RTI systems more affordable and easier to use compared to BRDF measurement systems such as gonioreflectometers. One important factor in RTI acquisition is the distribution of light source positions. The number of light source positions required for good quality RTI acquisition depends on the surface being imaged and the desired level of detail in the resulting model. For simple, flat surfaces, a smaller number of light source positions can be sufficient, while more complex or curved surfaces may require a larger number of positions to capture all of the necessary surface detail. Typically, the light source is moved around the surface being imaged in a circular or hemispherical pattern, with the camera remaining fixed in apex of the hemisphere as illustrated in Figure. 1.14 . There are several strategies for distributing the light source positions in an RTI acquisition. One common approach is to evenly distribute the positions around the surface, with the goal of capturing as much of the surface as possible. Another approach is to focus on specific areas of the angular space that are of particular interest, and place more light source positions in those areas. In general, the more light source positions that are used in an RTI acquisition, the more detailed and accurate the resulting model will be. However, it is also important to consider the trade-off between the number of positions and the time and resources required to acquire them. Balancing these factors can help to ensure that the RTI acquisition is efficient and produces the desired level of detail.

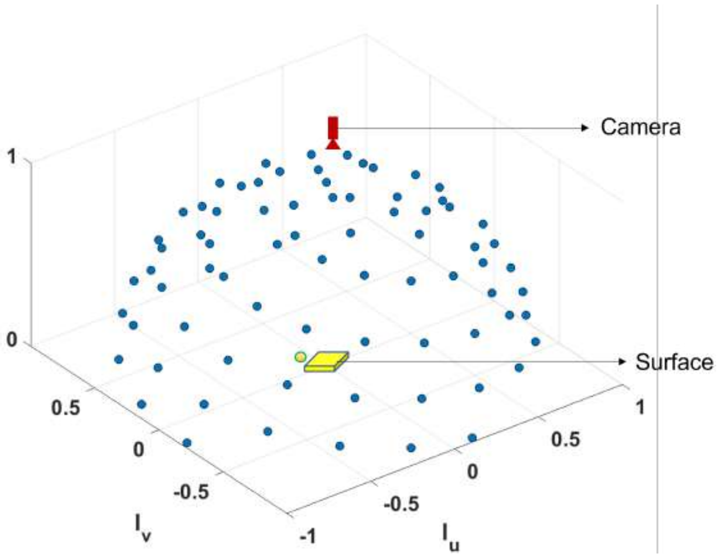


Figure 1.14: Distribution of light positions over an imaginary hemispehere typically used in RTI acquisition. The blue dots represents the light source positions. The green dot represents the reference reflective spheres generally used to compute the incident light directions.

As RTI was first introduced by Malzbender et al. [2001b], various new acquisition systems for performing RTI have emerged over time. A summary of these different acquisition set ups is presented in Table 2.1. The main existing acquisition systems are detailed below.

Hightlight-RTI (H-RTI) acquisition uses a handheld light source and fixed camera. This approach is widely used because of its simpliclity. A handheld light source is used to illuminate the surface of the object being imaged from different directions, while a stationary camera captures the surface luminance associated to each light position Imaging as shown in Figure. 1.15. The distance between the light source and the camera is typically maintained using a thread, which helps to keep the lighting conditions consistent across all images. One advantage of using a handheld light source and fixed camera for RTI acquisition is that it is relatively simple and inexpensive to set up. It does not require any specialized equipment. Additionally, this method is highly portable and can be used to capture RTI images in a variety of different settings, including outdoor environments.

Acquisition set-up	Description
H-RTI	Highlight RTI is carried out with a handheld light source, camera fixed on a tripod and 3 to 4 number of reflective spheres.Imaging
Fixed light dome system	Dome structure having typically between 50 to 100 number of light sources (LEDs) uniformly distributed over the hemisphere and a camera mounted at the apex of the dome. The positions of the light sources are pre calibrated and known.Kinsman [2016]
Mechanized dome system	A dome structure having a light source mounted on a motorized lever that moves the light source independently in azimuth and elevation space. Zendagui et al. [2021]
Robotic acquisition system	The light source is mounted on a robotic arm or a drone that manipulates the position of the light. Luxman et al. [2022]

Table 1.1: Summary of acquisition systems commonly used to perform RTI

However, one issue is that it can be difficult to maintain consistent lighting conditions across all images, especially if the light source is not held steady or if the distance between the light source and the camera changes. This can lead to variations in the intensity and color of the light, which can affect the quality of the final RTI image. Additionally, this method may not be suitable for capturing RTI images of highly specular or reflective surfaces, as the reflections and glare from the surroundings cannot be avoided. Another disadvantage is that it is very tedious, time consuming and generally requires skills to perform the acquisition. Occassionally reflective spheres are used in H-RTI. These spheres are included in the photographs captured in an acquisition to record the lighting angle. When light strikes the reflective sphere, it bounces off and creates a speck. This speck can be used to determine the direction of the light source.

Figure 1.15: Highlight-RTI being performed with handheld light source and camera held by a mechanical support. Picure courtesy: NCPTT

Fixed light dome RTI acquisition systems utilize a fixed number of light sources generally evenly spaced over an opaque hemisphere (dome). During the acquisition process, the object being captured is placed at the center of the dome and images are taken from a fixed camera position as the light sources are sequentially illuminated. This method allows for a large number of light source positions to be captured in a relatively short amount of time. One advantage of using a fixed light dome is that the light sources are always at a fixed distance from the object, resulting in consistent image quality. However, the fixed nature of the light sources is also a disadvantage, as it limits the range of lighting angles that can be captured and the affordable quality of lighting. Additionally, the dome hardware can be bulky and difficult to transport, making it less suitable for certain types of RTI acquisitions. An example this type of acquisition is shown in Fir. 1.16

Figure 1.16: An example of fixed light dome RTI acquisition system. This system is a model Scope D50 system from Broncolor Swiss, containing 50 evenly spaced LED lights. Images from bro

Mechanized dome RTI acquisition systems are a type of RTI setup that use a single light source mounted on a mechanical rig with an actuator, enabling the light source to be positioned at any point on the hemispherical dome. This type of RTI system is often preferred over handheld or fixed dome systems due to its greater flexibility and control over the light source position. This is especially useful for surfaces with complex geometries, as the light source can be positioned to better highlight specific features of the surface. Additionally, mechanized dome systems can typically capture a larger number of images in a shorter amount of time, making them more efficient and less time-consuming to use. There are also some disadvantages to using mechanized dome RTI systems. One of the main drawbacks is the cost, as these systems are often more expensive to purchase and maintain than the previous two set ups. Additionally, mechanized dome systems can be more complex to build, set up and operate, requiring specialized training and expertise to use effectively. Overall, mechanized dome RTI systems offer a high level of control and flexibility for capturing RTI images, making them a valuable tool for many applications in cultural heritage, art conservation, and other fields. An example of the mechanized light dome system is shown in Figure. 1.17

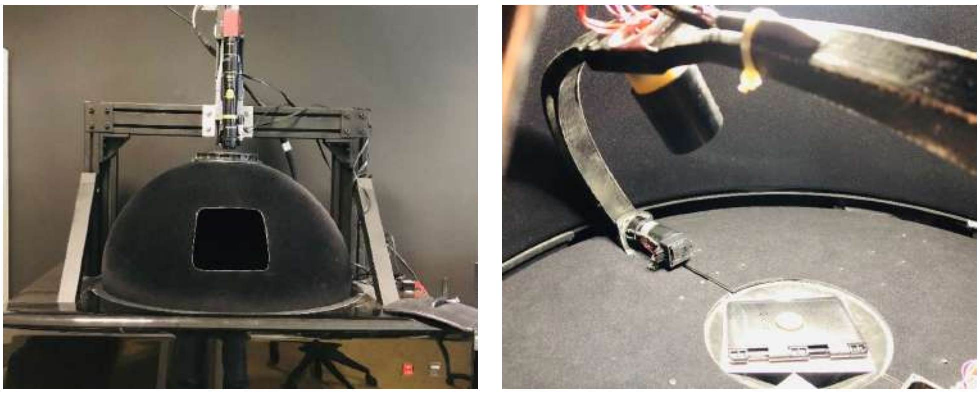


Figure 1.17: An example of mechanized dome RTI acquisition system. In the image on the right, we can see a light source attached at the apex of a semi cirecular rig, that eventually is attached to a motor and a rotating disc. This system is built inhouse in Imvia Laboratory (Universite Bourgogne, Dijon, France).

Robotic RTI acquisition systems There are recent attempts using robotic systems for RTI acquisition. However literature on solutions using robotic systems are very few. Chen et al. [2020] presents robotic arm based acquisition system for large-scale 3D surface imaging. In their approach they built a fixed dome system containing 64 LEDs - 16 evenly distributed at 4 vertical levels and each level has 4 LEDs placed with 90-degrees intervals between every two neighbouring LEDs. This fixed dome is attached to a 6DoF robotic arm as shown in Figure. 1.18

Figure 1.18: Acquisition system by Chen et al. [2020] based on robotic arm and fixed dome for large-scale 3D surface imaging.

While, this is efficient to acquire large surfaces in parts, it does not offer the freedom of choosing any light direction desired adaptive to the surface. Also, the dome radius is fixed and thus limits the option of having variable size dome.

In an another interesting approach by Krátkỳ et al. [2020], Unmanned Aerial Vehicles (UAV/ Drones) are used for positioning the camera and light source for performing RTI acquisitions in locations that are hardly accessible like the ceilings of a tall cathedral. Their design consists of two multi rotor UAVs where the UAV carrying the camera hovers in the air while the UAV carrying the light source flies around to provide the lighting from different directions. Figure. 1.19 shows an example RTI acquisition being performed using the UAV system by them in Church of St. Mary Magdalene in Chlumın

Figure 1.19: UAV based RTI acquisition introduced by Krátkỳ et al. [2020]

Acquisition of buildings is problematic due to its limited accessibility by humans, which led to the introduction the use of autonomous UAVs. The main advantage of the their solution is the ability to perform the RTI scanning procedure in places that are hardly accessible or even inaccessible to humans. The main drawback with this approach is the inability to eliminate any camera motion when the UAV carrying the camera is hovering in the air. Another drawback is that due to lack of reflective spheres, its is not possible to determine the precise direction of the illumination purely by localization of the constantly moving UAV carrying the light source.

Kitanovski and Hardeberg [2021] carried out an objective evaluation of relighting models on translucent materials from multispectral RTI images in which they used an acquisition set up employing a robotic arm to which a light source is attached. Their set up is as shown in Figure. 1.20

In the publication the author claims that their set up can cover elevations from 15° to 65° as well as the whole azimuth range in a fairly uniform manner. The range of the dome size achievable is not

Figure 1.20: Robotic arm based RTI acquisition set up by Kitanovski and Hardeberg [2021]

mentioned. The robot arm used is a 5DoF Haddington Dynamics dexter arm. Using the robotic arm is advantageous in providing flexibility to position the light source in any direction. However Kitanovski and Hardeberg [2021] doesn't have more information on reachability and workspace of their set up.

1.3.2 RTI modeling

The data acquired in RTI can be referred as multi light image collections (MLIC) since it consist of stack of images of the same surface but illuminated from different directions. Once the MLIC data is collected, it undergoes a processing phase where it's used to fit a mathematical model or function. This function is then utilized to virtually illuminate the surface from any desired direction. Various modeling techniques can be used to model the acquired information. The three commonly used methods includes Polynomial Texture Mapping (PTM), Hemi-Spherical Harmonics (HSH), and Discrete Modal Decomposition (DMD). These 3 approaches are detailed below.

Polynomial Texture Mapping Malzbender et al. [2001b] involves storing surface luminance for each input light position for each texel by interpolating the input images (in computer graphics, a texel, texture element, or texture pixel is the fundamental unit of a texture map Glassner [1989]. Just as images are depicted by arrays of pixels, textures are characterized by arrays of texels that represent the texture space). In PTM, a biquadratic polynomial is used to model the reflectance behavior of a surface under varying lighting conditions. The coefficients of this polynomial are stored per texel (texture element), and are used to reconstruct the surface color. This luminance model is represented as shown in Eq. 1.11.

$$L(u, v) = a_0 + a_1 lu + a_2 lv + a_3 lu^2 + a_4 lu lv + a_5 lv^2 \tag{1.11}$$

where (lu, lv) are projections of the normalized light vector into the local texture coordinate system (u, v), L is the resultant surface luminance at that coordinate and the coefficients (a0-a5) of this polynomial are fit to the photographic data per texel and stored as a spatial map called a Polynomial Texture Map. To create a PTM, the best fit for the polynomial coefficients is computed using singular value decomposition (SVD) for each pixel, given an arrangement of light sources. The resulting six floating-point coefficients per texel are stored as 8-bit integers for evaluation speed and compact storage. To eliminate the problem of having several orders of magnitude difference between high and low order coefficients, six scale (λ) and bias (Ω) values are also stored with each PTM, one for each coefficient. During reconstruction, these values are applied to the stored 8-bit coefficients to recover their final values.

Hemispherical Harmonics (HSH) Wang et al. [2009], Gautron et al. [2004] are a set of functions that are defined on the surface of a sphere and are used to represent functions defined on the sphere in a compact and efficient manner. This is done by fitting the captured RTI data to the HSH basis functions. The number of coefficients used to describe the surface normal in HSH is determined by the order that is chosen to create it. The order refers to the degree of the polynomial used in the HSH basis functions. For a first-order HSH, there are 4 coefficients because it uses linear basis functions, which can be thought of as fitting the data to a plane. This is similar to using a 2D plane in 3D space, which has 4 parameters (3 for the normal vector and 1 for the distance from the origin). For a second-order HSH, there are 9 coefficients because it uses quadratic basis functions, which can be thought of as fitting the data to a paraboloid. This is similar to using a 2D parabola in 3D space, which has 9 parameters. For a third-order HSH, there are 16 coefficients because it uses cubic basis functions, which can be thought of as fitting the data to a cubic surface. This is similar to using a 2D cubic curve in 3D space, which has 16 parameters. In general, for an nth order HSH, we choose $(n+1)^2$ coefficients. The higher the order, the more accurately we can represent complex surfaces, but at the cost of increased computational complexity and file size.

The intensity of light reflected from a point on a surface in a particular direction can be represented as:

$$L(\theta_i, \phi_i) = \sum_{l=0}^{n} \sum_{m=-l}^{l} Y_l^m(\theta_i, \phi_i) \cdot L_l^m \tag{1.12}$$

where $L(\theta_i, \phi_i)$ is the intensity of reflected light, $Y_l^m(\theta_i, \phi_i)$ are the spherical harmonic basis functions, L_l^m are the coefficients that we solve for during the fitting process, θ_i and ϕ_i are angles defining the direction of incident light, the double summation is over all orders l from 0 to n, and all degrees m from $-l$ to l.

This equation essentially says that we can represent the light reflected off a point on a surface as a sum of spherical harmonic basis functions, each multiplied by a certain coefficient. These coefficients are what we solve for when we do the fitting process.

Discrete Modal Decomposition (DMD) method, proposed by Pitard et al. [2015], is an innovative approach used for surface appearance modeling and rendering. It employs vectors that are derived from dynamics, specifically the natural modes of a structure, to model the angular reflectance of a surface. To better understand this concept, consider the analogy of a song. When we listen to a song, we hear a blend of various instruments playing together. However, if we were to break down the song into its individual components - such as the guitar, drums, and vocals - we could listen to each part separately. This would give us a clearer understanding of each instrument's contribution to the overall song. The DMD method does something similar, but with light instead of sound. It breaks down the light reflected from a surface into its individual components or modes.

This process is similar to performing a discrete Fourier transformation. The decomposition allows us to evaluate the spectral content of a discrete signal from a projection space that is derived from vibratory mechanics. The projection base is composed of eigenvectors that result from solving a structural dynamics problem.

This problem begins with the geometry of a hemispherical surface where the base circle is constrained to zero displacement. Mathematically, this is defined as:

$$M \cdot \ddot{q} + K \cdot q = 0, \text{ with } q = q(\theta, \phi, t) \tag{1.13}$$

Here, M represents the mass matrix, K represents the stiffness matrix, and $q(\theta, \phi, t)$ is the displacement vector that characterizes the modal shapes.

The solution to this problem is given by:

$$q(\theta, \phi, t) = \sum_{k=1}^{inf} Q_k cos(\omega_k t) \tag{1.14}$$

In this equation, Q_k is the amplitude vector associated with the frequency ω_k. To determine the eigenmodes defined by (Q_k, ω_k), we solve the linear system of equations defined below:

$$(M^{-1}K - \frac{1}{\omega_k^2}I)Q_k = 0 \tag{1.15}$$

In this equation, I is the identity matrix and $M^{-1}K$ is assumed to be diagonalizable. The discrete solution is calculated using finite element analysis. The solution forms what we call the projection base or modal base.

These modes are designed to capture non-periodic components of a surface as they are inherently linked to the concept of periodicity. The resulting modal basis represents a set of elementary forms that can be used to describe the visual appearance of any surface.

When compared with other methods based on a second-order polynomial (PTM) and hemispherical harmonics (HSH) vector basis, it was found that the DMD-based reflectance model accurately describes complex local reflectance in terms of angular and intensity luminance variations Pitard et al. [2015]. It provides a more reliable approximation of specular lobes and glossy surface areas compared to other methods. An example of modelling the angular reflectance of a surface point using PTM, HSH and DMD is shown in Figure. 1.21.

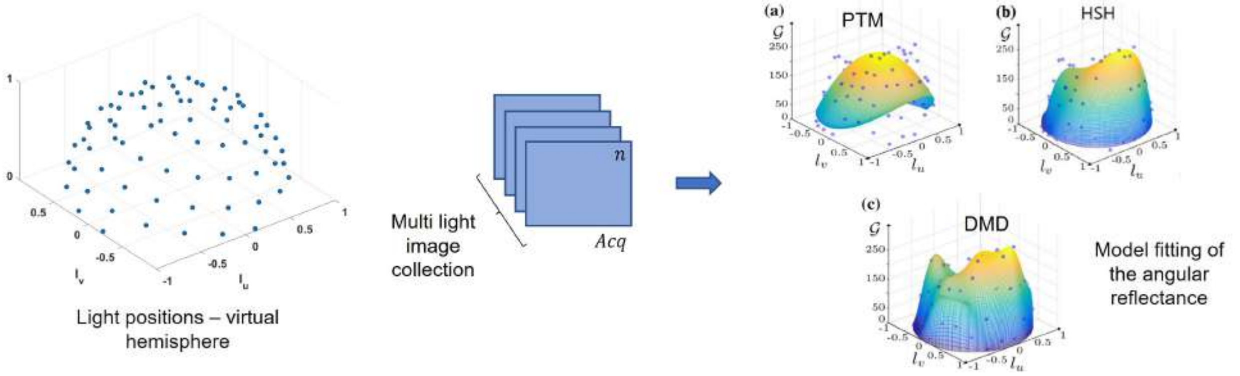


Figure 1.21: RTI model fitting to the data acquired - PTM, HSH and DMD. DMD modelling is observed to be more accurate than the other two.

Neural RTI is another approach for generating Reflectance Transformation Imaging (RTI) models using deep learning techniques. The NeuralRTI method proposed by Dulecha et al. [2020], involves training a deep neural network to predict the RTI model of an object from a set of input images captured under different lighting conditions. The network is trained using a combination of supervised and self-supervised learning, with the goal of accurately predicting the RTI model. In this thesis, we did not employ the Neural RTI technique, and as a result, an in-depth discussion of the method is not included.

1.3.3 RTI visualization

RTI enables the visualization of surface information that may not be apparent through direct examination of the physical object. One of the key features of RTI is the ability to interactively relight the surface from any directions. By changing the direction of the virtual light source, the surface's micro-geometry is highlighted, allowing the user to visualize and analyze the surface in great detail Mytum and Peterson [2018]. Another key aspect is the ability to build various maps of the surface, including normal maps, directional slopes, and specular enhancement maps from the RTI model data Pamart et al. [2019], Ponchio et al. [2018]. These maps can provide valuable insights into the micro-geometry and appearance of the surface, and have a range of applications in fields such as archaeology Mytum and Peterson [2018], Florindi et al. [2020], cultural heritage Manfredi et al. [2013], Earl et al. [2011] and materials science Coules et al. [2019]. By mathematically enhancing the subject's surface shape and color attributes, RTI reveals surface information that may not be apparent through direct examination of the physical object. RTI software and methodologies are widely used for analyzing and visualizing surface properties in fields such as archaeology, cultural heritage, and materials science. Figure. 1.22 shows an example of the visualization of different maps obtained from RTI data.

(a) Relighted image (b) Specular enhancement (c) Normal map (d) Directional slope

Figure 1.22: Example of visualization of various maps obtained from the RTI data

1.3.4 Challenges in RTI

While RTI has numerous applications in fields such as cultural heritage, there are still several challenges that need to be addressed to fully realize its potential.

A significant challenge in RTI is the time-consuming data acquisition process. RTI requires multiple images to be captured from different angles and lighting conditions, which can be a time-consuming process, particularly when imaging large objects. To address this challenge, there is a need to streamline the data acquisition process. Further research is needed to develop automated RTI systems that can handle a wide range of object sizes, shapes, and surface properties. It is desired to build a novel system that addresses the limitations of the existing systems discussed above. Ideally, the novel acquisition system must have the following characteristics:

- Capable of positioning the light source at any desired direction
- Capable of achieving range of dome radius
- Capable of translating the surface being acquired to cover larger area
- Enable automation of the acquisition process

For larger surfaces, after the data acquisition, another challenge is the stitching of the acquired data. In the curret state of the art, there is no pipeline to perform stitching of RTI data for large surfaces. This is particularly challenging as well as important since there are active lights being used in the RTI acquisition and hence regular image stitching methods cannot be directly applied to RTI data.

Another most significant challenge in RTI is the problem of choosing the best light positions adaptive to surface. The surface of an object can be complex and irregular, which can make it difficult to determine the optimal lighting angles to capture the most detailed and accurate RTI data. Adaptive lighting approaches have shown promising results in some cases, but there are still some challenges that need to be addressed. For example, the algorithms used for adaptive lighting can be computationally expensive, requiring significant processing power and time.

In conclusion, Reflectance Transformation Imaging (RTI) is a powerful imaging technique with numerous applications in various fields. However, there are still several challenges that need to be addressed to fully realize its potential. The problem of choosing the best light positions adaptive to surface, time-consuming data acquisition, and acquisition of large surfaces while maintaining high image resolution are some of the most significant challenges in RTI. By addressing these challenges, we can further enhance the capabilities of RTI and expand its applications in various fields.

1.4 Conclusion

In this chapter, we have provided a comprehensive overview of the state of the art in reflectance imaging. It begins by introducing the basic concepts of reflectance, including radiometry and BRDF. Then we discussed various reflectance models, including the Lambertian model, mirror BRDF model, and microfacet models such as Torrence Sparrow, Phong, Cook and Torrence, and GGX. Finally, the chapter introduces the topic that we focus on this research - the Reflectance Transformation Imaging (RTI) technique, including RTI acquisition systems, modeling, and visualization, and highlights the challenges associated with RTI. This chapter provides valuable insights into the fundamentals of reflectance imaging and highlights the importance of RTI in various fields, such as cultural heritage and art conservation.

CHAPTER 2

A robotic system for automated and adaptive RTI acquisitions

> ***Chapter overview***
>
> ***Utilization of RTI techniques in the field of cultural heritage is increasing . Presently, the systems employed for acquiring the RTI data primarily utilize free-form or dome-based methods. However, these systems have limitations in terms of repeatability, the maximum size of objects that can be captured, as well as speed and mobility. The main drawback of these systems are that they are not able to adapt to the various sizes and shapes of the objects being acquired. In this chapter, we describe a new system we developed, called LightBot, which utilizes a robotic arm. Our proposed approach enables automated, reproducible capturing of large or complex scenes in two dimensions while maximizing image resolution.***

2.1 Introduction

RTI involve capturing a series of images where only the light source varies in its spatial position. This technique involve taking pictures of an object from a fixed viewpoint while changing the light direction in each image captured. Regardless of the ultimate goal of a multi-light image collection process, whether it be creating a 3D view or relighting, this technique have been proven to provide a reliable description of important parameters related to the surface appearance and geometry. As a result, they have found widespread application in the characterization of cultural heritage surfaces.

Free-form H-RTI systems offer greater flexibility in terms of setup as they typically involve using portable equipment and a handheld light source with an adjustable camera position. This allows for multi-scale acquisitions, but have several limitations including

- Time needed to perform the acquisition
- Lack of rigor
- Repeatability and reproducibility not guaranteed

- Achieving the accurate distance and orientation of the light is not guaranteed.

These limitations significantly affects the quality of the RTI data. Dome-based systems, on the other hand, offer a higher level of repeatability but have limitations in terms of the size of objects that can be captured and a limited range of angles for positioning the light source. Additionally, these systems tend to be less portable compared to free-form systems. We present a novel robotic arm-based system design that addresses the limitations of both dome-based systems and manual H-RTI systems. The proposed system addresses the problem of surface adaptivity that is present in dome-based systems, as well as the issues of accuracy, repeatability, and longer acquisition time that are associated with manual H-RTI systems. The system allows for more robust automation and reproducibility of series of acquisitions of large or complex scenes in two-dimensional space while optimizing for surface adaptivity and pixel resolution.

The size of the dome used in Reflectance Transformation Imaging (RTI) acquisition should be adapted to the size of the surface being captured. As shown in Figure. 2.1, using a larger dome for a larger surface and a smaller dome for a smaller surface ensures that the entire surface is captured with the best possible image quality.

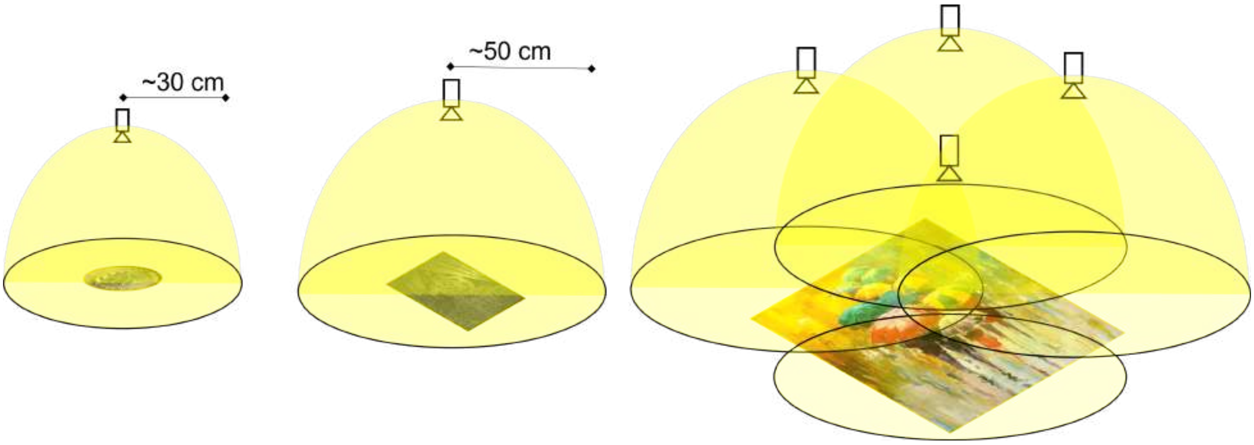


Figure 2.1: While a typical RTI dome system is suitable for capturing small artifacts like coins, it may not be able to handle larger surfaces effectively.

However, using a dome that is too large for a small surface can result in a less focused and less detailed image. This is because the light captured by the dome is scattered over a wider area than the surface being imaged, resulting in a less focused image. Imagine the dome as a camera lens, and the surface being imaged as the object in the frame. If the lens is too wide, more of the background will be included in the frame, which can reduce the sharpness and clarity of the image of the object. Similarly, if the dome is too large for a smaller surface being imaged, the light that is captured will be coming from a wider area than the surface, resulting in a less focused image - where we end up capturing more ambient light and less light coming from the surface itself, leading to a less detailed image with less contrast. For larger surfaces, it is often necessary to divide the surface into multiple regions and perform multiple acquisitions, which can then be stitched together to form a complete image. This is to not lose the required level of magnification to capture the details on the surface.

In the article by Kitanovski and Hardeberg [2021] , they propose a method similar to ours, in which they use a robot arm for performing RTI acquisition. However, our system, called LightBot, is the first to utilize both a robot arm and an XY stage to control both the light position and the surface position, thereby enabling RTI stitching. Other robot-based systems for BRDF data acquisition have also been developed, such as the one by Santos et al. [2017] in which the robot arm is used to manipulate the camera position, and a turntable is used to manipulate the position of the object being scanned. In contrast, our LightBot system keeps the camera stationary and uses the robot arm to manipulate the light pose. This combination of a robot arm and XY stage enables a wide range of possibilities, particularly in the case of large-scale or complex-shaped Cultural

Heritage objects that cannot be moved, such as paintings and manuscripts. Furthermore, this system enables automation, which is particularly useful when large numbers of acquisitions are required, such as in documenting series of coin collections or monitoring the condition of an object over time. Our proposed LightBot system addresses the challenges of performing RTI on large-scale surfaces, which are not well addressed by current state-of-the-art acquisition systems. LightBot is a movable, high-speed, efficient, cost-effective, and adaptable RTI acquisition system. To the best of our knowledge, this is the first fully automated approach for the RTI acquisition process. The system utilizes a robotic arm and addresses several important aspects of the RTI acquisition process such as reducing acquisition time, enabling stitching of RTI data, and increasing the repeatability and reproducibility of acquisition setups. The robot arm offers flexibility by allowing for a variable distance between the light source and the captured surface, allowing for adjustments based on the surface geometry. LightBot uses a standard Light Positions (.lp) file (simple text file format with 4 columns - image filename, x, y, z, which maps captured images with their respective light directions in Cartesian space.) and manipulates the robot arm such that the light source is positioned at those locations one at a time and maintain the pose of the light source such that it points towards the center of the surface.

2.2 System design

Previously we described the advantages and limitations of the existing acquisition systems. Table 2.1 gives an overview of the limitations of the existing RTI acquisition systems and the list of characteristics of a desired novel system to overcome these limitations.

Our goal is to develop a versatile system capable of adaptive surface acquisition. This system would include features such as a dome of variable size and continuous spatial light positioning. For surfaces larger than 30 cm, the system should be able to capture the details by acquiring the surface in sections, each providing a reasonable level of magnification. For surfaces smaller than 30 cm, a single acquisition covering the entire surface would typically suffice. To facilitate the acquisition of a surface in sections, the system should have the capability to move the surface in the XY plane, positioning different sections within the camera's field of view for each acquisition. Our novel robotic acquisition system, LightBot is composed of a collaborative robot arm, a light source attached to the end effector of the arm, a high-resolution camera, and a cartesian platform for precise translation of the object in the X-Y space with respect to the camera.

Using robotic arm to manipulate the light source enables us to achieve:

- Variable size light dome.
- Continuous spatial light positioning.

Using XY translation table enable us to achieve:

- Surface size adaptive acquisition without compensating the magnification.
- Batch acquisition.

Finally, we also developed a web based UI and an API for the system to enable:

- Full automation of the RTI acquisition.
- Efficient acquisition pipeline.

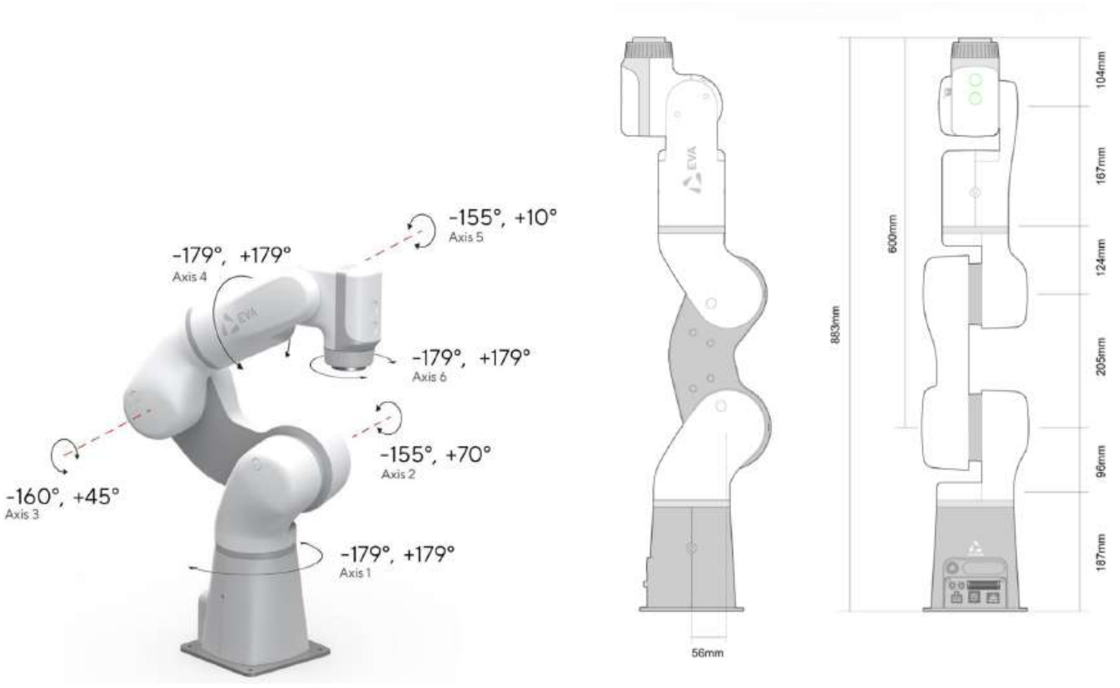
We built two prototypes of the robotic acquisition system - LightBot1 and LightBot2.

Acquisition system	Limitations	Novel system
Freeform highlight RTI	1. Tedious and time consuming 2. Repeatability not guaranteed 3. Highly vulnerable to human errors	- Efficient and enable full automation - Repeatable - Variable size dome - Handle range of object sizes - Transportable
Fixed lights dome	1. Limited only to a particular set of light directions configured during the construction. 2. Limited to small sized objects. 3. Radius of the dome is fixed.	
Mechanized light dome	1. Limited to small sized objects 2. Radius of the dome is fixed.	
Drone system	1. Inability to make the camera not move during the acquisition 2. Determining the direction of illumination is not accurate.	

Table 2.1: Limitations of the existing RTI acquisition systems and the need for new system

2.2.1 LightBot 1

In the first prototype, we have used a 6 DoF lightweight robotic arm - Automata Tech's EVA robot arm (Figure. 2.2) which has a maximum reach of 600 mm and a workspace as shown in Figure. 2.3.



(a) Joint limits of Eva robotic arm (b) Link lengths of the Eva robotic arm

Figure 2.2: Automata.Tech Eva 6DoF robotic arm. Source: Automata tech, aut

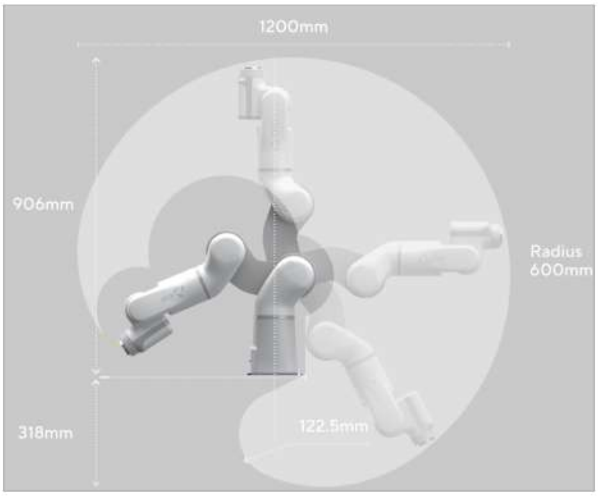


Figure 2.3: Work space of Eva robot arm. Source: Automat tech, aut

The camera used in this setup is an Allied Vision's Manta G504B (sensor size 2/3) with a suitable objective of focal length 3mm to 6mm to achieve an FoV of 150 mm width and 100 mm height at a working distance of 450 mm to 550 mm.

In LightBot1 we used a normal household white LED spotlight. If we must use a LED spotlight for RTI acquisition, it is important to choose one that has high CRI (Color Rendering Index) which is a measure of how well a light source can accurately render the colors of objects to make it appear more natural and realistic under light. The ideal temperature of a light source for RTI is 5500K. This is because 5500K is the color temperature of daylight, which is the most common light source used for photography and videography. This temperature make the object appear more natural. Further more, for RTI it is important to choose a light source that has a wide spectral distribution (SPD) that describes the way that the light is distributed across the visible spectrum. A good spectral distribution for RTI will have a wide range of wavelengths, from the blue end of the spectrum to the red end of the spectrum. This will help to ensure that all of the details of the object being imaged are captured.

We chose the Feit Electric PAR30 LED bulb for LightBot1 because it has a high CRI of 95. The beam angle and luminance of the light source are illustrated in Figure.2.4. As we see, the beam angle of the bulb is 36 degrees. The luminance of the bulb is 1200 lux at a distance of 1 meter. Overall, the Feit Electric PAR30 LED bulb is a good choice for LightBot1. It has a high CRI, a wide beam angle, it is very bright, commonly available and costs lower price.

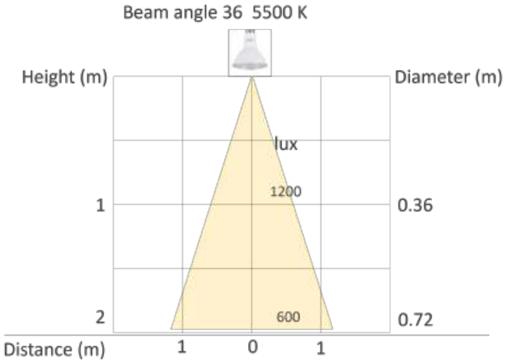


Figure 2.4: LED lamp used in LightBot1

We used a motor actuated XY platform which is a device that can be used to translate objects in two dimensions. The platform consists of two axes, one for the X direction and one for the Y direction. Each axis is driven by a motor, which allows the platform to move in the desired direction to desired point precisely. The XY platform used in LightBot 1 has the stroke length of 40 cm. Stroke length is the maximum distance that it can move in either direction and thus the XY platform can move objects up to 40 cm in either direction.

As shown in Figure.2.5, the system is set up on an optical table. In this particular configuration, the XY platform is mounted in a vertical position. This allows for scanning to be performed with the surface attached to the XY platform while it is in an upright position. To maximize the workspace and ensure the camera fits within it, the surface is positioned vertically in front of the robot arm when it is in an upright position. As seen in the figure, the light source is attached to the end effector of the robotic arm with an offset from the axis of the end effector. This allows us to create a wider range of lighting angles and positions thereby offering increased flexibility.

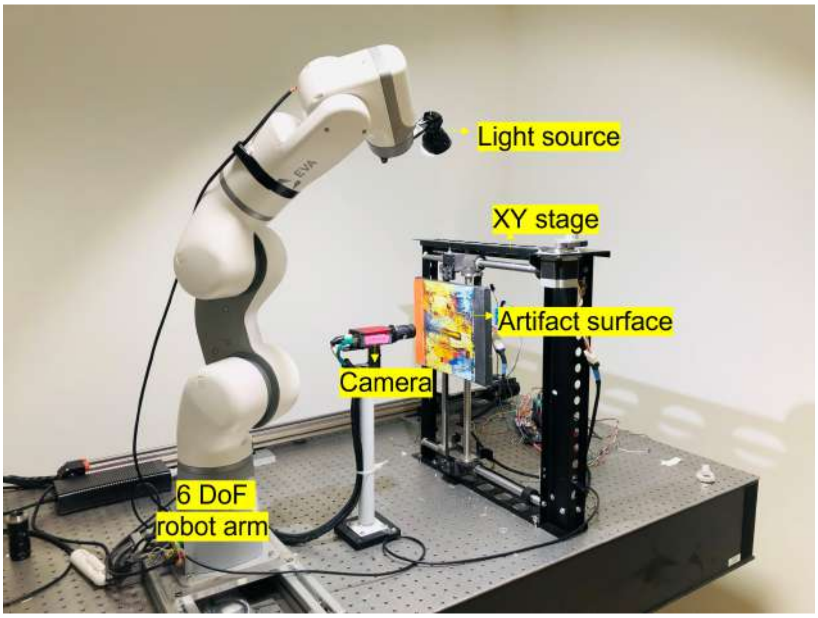
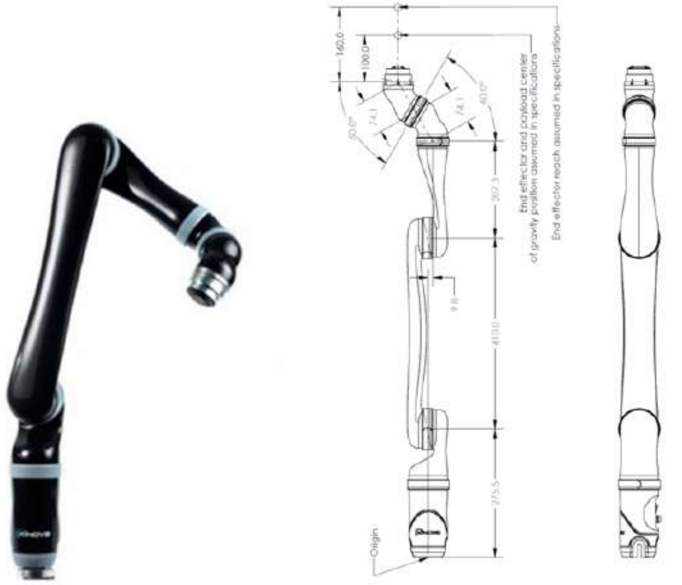


Figure 2.5: LightBot prototype

2.2.2 LightBot 2

The LightBot 1 system is a proof of concept, and subsequent to that we developed an upgraded version of the system with an advanced and longer robotic arm, a high-resolution camera. In this iteration of the LightBot development, the robotic arm and camera are mounted in a top-down configuration, allowing for the surfaces to be placed in a horizontal position as opposed to the previous configuration having vertical positioning. This change will enhance the versatility of the system and provide even more flexibility in the acquisition process. This design allows for more efficient and accurate RTI acquisitions, and makes it possible to acquire images of large-scale surfaces with minimal manual interventions.

In this upgraded version of the system, we use Kinova Jaco gen 2 robot arm, which offers greater reach and dexterity compared to the previous version. This allows for a wider range of dome sizes to be used in RTI acquisitions. This robot arm has seven degrees of freedom (7 joints), which allows it to move in a wide range of directions. It has a length of 1.06 meters and a reachability of 0.985 meters. The payload of the kinova robot arm is upto 2.4 kg and hence it is possible to attach superior light source with own cooling system that weighs heavy to the robotic arm. Another important advantages of this robot arm is its lightweight design. The arm weighs just 5.5 kg, which makes it easy to transport and install in any configuration (including top down or side mount). This lightweight design also makes the arm more maneuverable, which is important for applications that require precise movements. The arm has isolated power lines running through it from the base to the end effector. This means that the power lines are electrically seperated from the rest of the arm, which enables us to attach any light source to the robot arm without addressing the cabling of light source's power and control cables. In summary, the Kinova Jaco Gen 2 7DOF robot arm is a powerful and versatile robotic arm which makes it an ideal choice for our application. Figure.2.6a , 2.6b shows the robotic arm and its link lengths. In Figure.2.7, we have shown the workspace of the robotic arm which is quite huge. Another benefit of the Kinova Jaco gen 2 robot arm is its high repeatability and precision, which makes it ideal for capturing high-resolution geometry data of CH objects.



(a) Robot arm

(b) Link lengths of the Kinova robotic arm.

Figure 2.6: Kinova Gen 2 Jaco 7DoF robotic arm. Source: Kinova kin

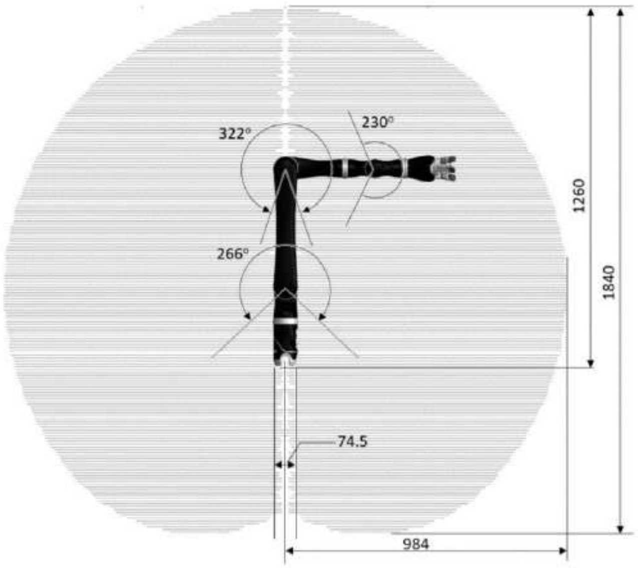
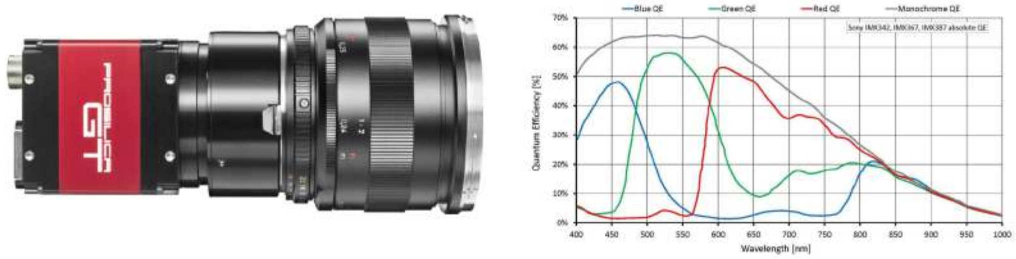


Figure 2.7: Usable workspace of Kinova Gen 2 7DoF robotic arm. Dimensions in mm

For LightBot 2, we have used the Allied Vision Proscilica GT6400 high-performance camera that is designed for a variety of industrial applications. The camera features a 6 megapixel sensor with a power over ethernet option. This model has the F-mount option and hence is compatible with a wide range of lenses including professional DSLR camera lenses. The EF mount lens control option of the camera enables us to control the focus, aperture, and other features of an EF mount lens from a remote device. This is useful to achieve automation of adjusting the aperture, focus to the surface being acquired. This camera model has good Quantum efficiency (QE) which is a measure of how efficiently a camera sensor converts light into an electrical signal. A higher QE means that the sensor is more efficient at capturing light, which results in better image quality. In general, a QE of 50 or higher is considered good for most types of photography. DSLR camera have QE beteen 50 to 90, smart phones between 30 to 55 and for astrophotography between 70 and 95. As seen in the Figure.2.8b, the LightBot 2 camera has good QE overall in the visible spectrum of light and that is suitable for RTI.



(a) Proscilica GT6400 camera (b) Quantum efficiency plot

Figure 2.8: Camera used for LightBot 2

LightBot 2 is equipped with a high-performance LED light source, the EFFI-spot by EffiLux. This light source is specifically designed for short-range working distances, with an optimal range of up to 50 cm. It features an integrated cooling fan to ensure optimal performance during prolonged use. One of the key features of the EFFI-spot light source is its color temperature of 5500 K ± 500 K, and is ideal for RTI acquisition. The light source is mounted on the end effector of the robot arm as illustrated in Figure. 2.9. This allows for maximum flexibility.

(a) EFFI spot light source (b) Light source attached to the end effector of the robot arm

Figure 2.9: Light source used in the Box set up

Figure.2.10 illustrates the working distance and the illuminance of the light source used. As seen in the illustration, the light source used in LightBot 2 has much higher illuminance than the one used in LightBot 1. The diameter of the illuminance cone is smaller than that of the LED light source used in the LightBot 1. Hence incremental region of the surface (area covered per acquisition) is lower than the LightBot 1.

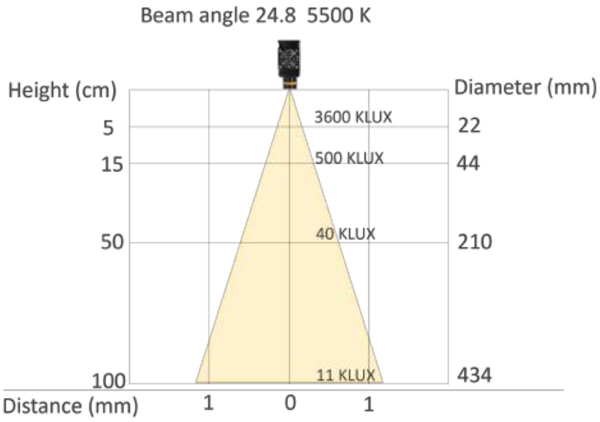


Figure 2.10: EFFI spot light working distance and illuminance

We used the same XY platform in LightBot 2 as in LightBot 1. As seen in Figure. 2.11, the XY platform (cartesian platform) is composed of a 1mm steel plate on which the surface is placed, secured by small L-shaped magnets that are equipped with Aruco markers Fiala [2004]. These aruco markers are used to automatically center the surface within the camera's field of view, as well as to accurately estimate the size of the surface being imaged. This allows for a fully automated manipulation of surface position, removing the need for manual adjustments and increasing the overall efficiency of the acquisition process.

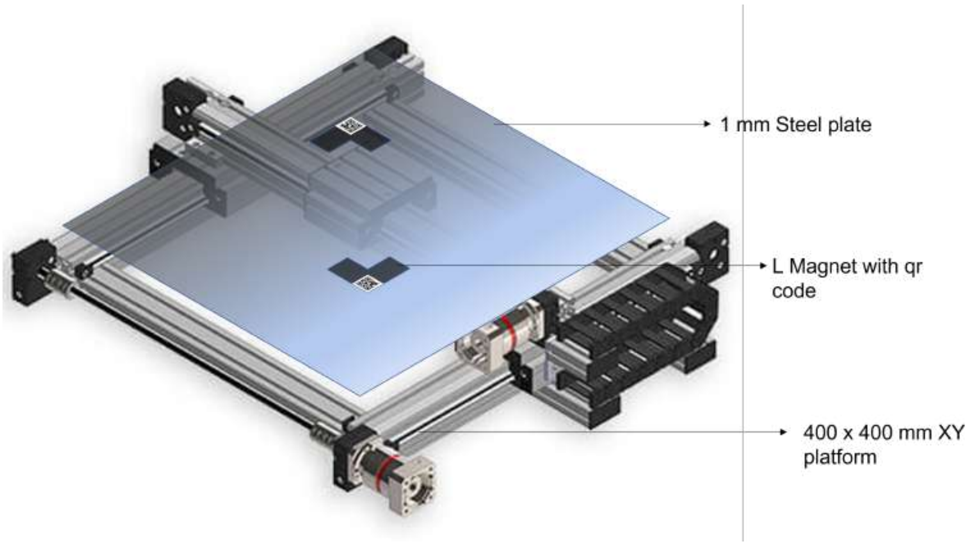
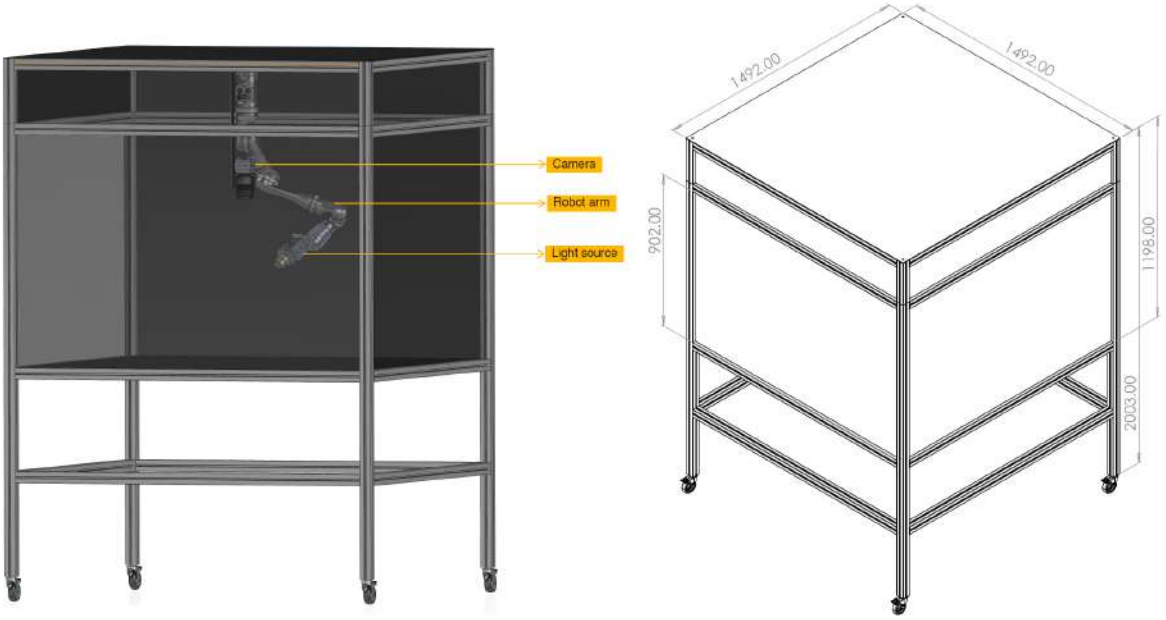


Figure 2.11: XY platform to translate the surface(s)

The LightBot 2 system is designed with a focus on versatility and adaptability. This arm used in this version is longer, more dexterous, and has a higher reachability than the previous version of the LightBot system. LightBot 2 is designed as a simple, black box to create an enclosed space for the acquisition of surfaces. The box system structure is built using aluminium profiles with camera and robotic arm are assembled top-down. This new design allows for dome sizes ranging up to 600 mm. The interior and exterior of the box are black to prevent any external light from entering or internal light leaving the box, creating a completely dark environment for RTI acquisition. This design allows for optimal light control and eliminates the need for additional light-proofing measures. With these new features, the LightBot 2 is designed to be a powerful, next generation tool for RTI acquisition, which can be used for a wide range of applications. Figure.2.12 illustrates the LightBot 2 assembly.



(a) The box set up assembled (b) Dimensions of the box

Figure 2.12: LightBot2 - The box set up

Some of the components for example, a motor-actuated camera objective for auto-magnification and focus, Region of Interest (ROI) finder using fiducial markers are still a work in progress. With these additional features, we aim to realize a fully automated surface adaptive RTI acquisition system.

2.3 System architecture and engineering

The LightBot system features a modular architecture, as illustrated in Figure. 2.13.

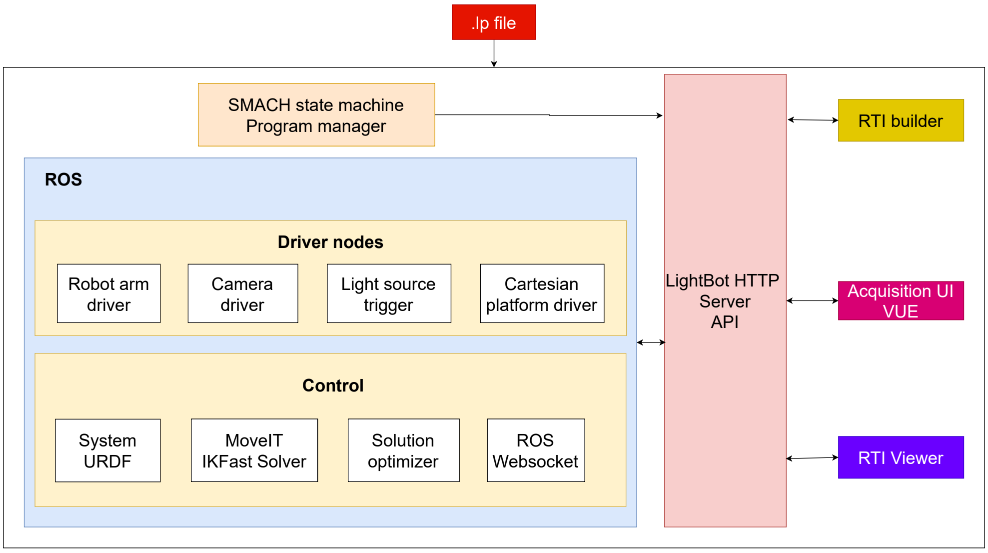


Figure 2.13: System architecture of the LightBot.

This architecture allows for easy integration of various components and facilitates communication between them. The system is built on the Robot Operating System (ROS) framework [Quigley et al. [2009], Koubâa et al. [2017], Quigley et al. [2015]], which is an open-source software framework for robotics. It provides a set of libraries and tools that enable developers to easily create and control robots. ROS provides a common interface for various hardware and software components, which makes it easy to add new components or replace existing ones.

To control the system, a HTTP server is built, and an API documentation is also developed for the same. An API (Application Programming Interface) provides a set of predefined functions and protocols that allow different software components to communicate with each other. In this case, the API allows developers to control the system remotely and automate tasks such as data acquisition.

To make the system user-friendly, a simple web-based user interface is built using the VUE frontend framework. The Vue.js, Filipova [2016], Kyriakidis et al. [2016] is a popular JavaScript framework for building user interfaces. This interface allows CH conservators to easily carry out RTI acquisitions by just uploading the *.lp* file and setting a few parameters like the surface size, etc. This eliminates the need for developers to have specialized knowledge of the system and allows them to focus on the preservation and analysis of the cultural heritage objects. Images of the user interface is provided in the appendix of this thesis.

2.3.1 System modelling and transformations

We modelled the robot system by first creating the URDF (Unified Robot Description Format) of the system from the 3D model of the links and components. URDF is a XML file format that is used to describe the geometry and kinematics of a robot. URDF is used by ROS to represent robots. The URDF consists of number of different elements that includes - links, joints, visual elements, collision elements. Links are the physical components of a robot. Joints are the connections between links which can be fixed, prismatic or rotational. Visual elements are used to represent the appearance of a robot and can be meshes, textures or colors. Collision elements are used to represent the collision geometry of a robot. Collision elements can be meshes, spheres, or boxes.

Figure.2.14 illustrates the coordinate systems we have defined for LightBot and the transformations involved. We use the same model for both LightBot 1 and LightBot 2.

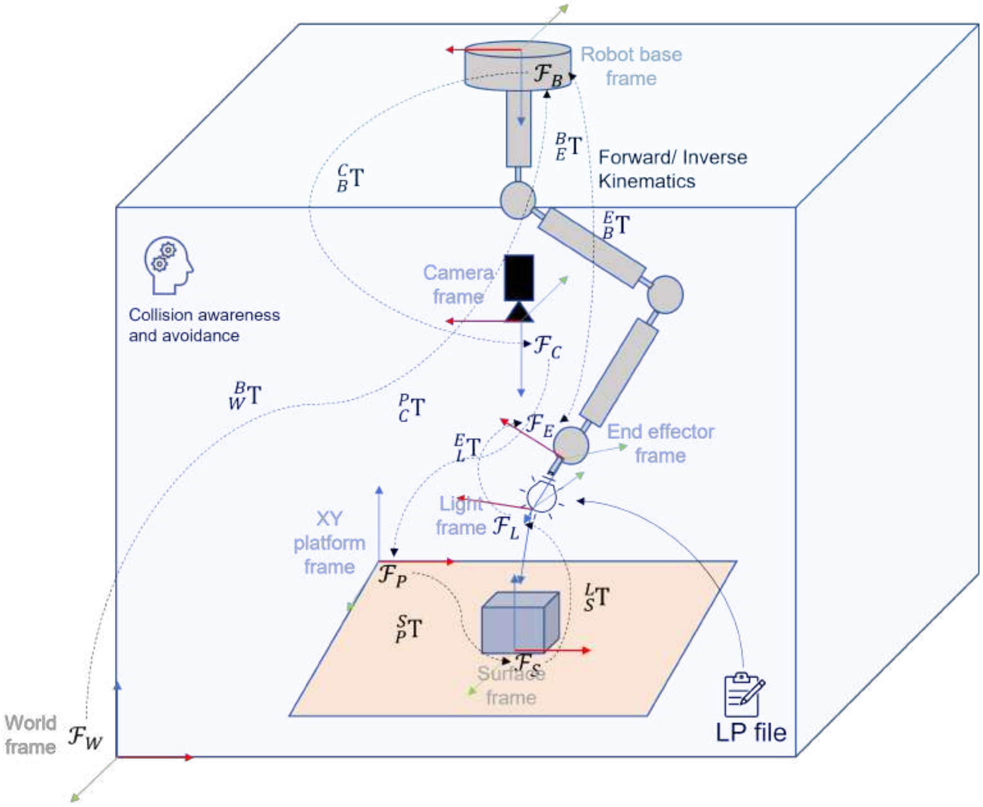


Figure 2.14: Coordinate systems and the transformations in the LightBot system

As seen in the figure, we use a .lp file to specify the light positions with respect to the surface frame. The lp file contains the light positions in the form of x, y, z coordinates. To find the full 6D pose of these light positions, we point the z-axis towards the origin of the surface frame. A 6D pose refers to the position and orientation of an object in 3D space. The position is given by 3 coordinates (x, y, z) and the orientation is given by 3 angles $(roll, pitch, yaw)$. These 6 values together define the complete pose of the light source in the 3D space. We denote the transformation matrix from the surface frame to the light position frame as ${}^{L}_{S}T$. For a light position $p = [xyz]^T$, we construct a unit vector pointing from the light position towards the origin of the surface frame as follows:

$$z_{SL} = \frac{p}{\|p\|} \tag{2.1}$$

This vector defines the z-axis of the light position frame with respect to the surface frame. Since there are an infinite number of solutions for the x and y axes, we can randomly select two orthogonal unit vectors that are also orthogonal to z_{SL} to define the x and y axes.

In robotics, there are two spaces that are used to describe the position and orientation of a robot arm: task space and joint space. Task space is the space of all possible positions and orientations of the robot's end-effector. Joint space is the space of all possible configurations of the robot's joints, or the angles of the joints. The robot's end-effector can be moved to any point in task space by specifying the desired position and orientation of the end-effector. However, the robot's joints can only be moved to a limited set of angles. The process of converting a point in task space to a set of joint angles is called inverse kinematics. Inverse kinematics is a complex problem, and there is no single solution that works for all robot arms. However, there are a number of methods that can be used to solve the inverse kinematics problem. In our case, the robot arm joints are manipulated by solving the inverse kinematics problem of the robot arm to position the light source in a LP direction. For computing the inverse kinematics [D'Souza et al. [2001], Kucuk and Bingul [2006]] of the robot, an IKFast [Diankov [2010], Coleman [2014]] plugin from the URDF (Unified Robot Description File) [MONICA [2016], Lu [2016]] of the robot arm was created using the IKFast kinematics solver in OpenRave [Diankov and Kuffner [2008]]. This plugin allows for the computation of inverse kinematics solutions quickly and efficiently. All the transformations are carried out using the ROS frame transformation server, which is a part of the ROS framework. This server provides a convenient way to perform coordinate transformations between different frames of reference.

2.3.2 Motion planning

Once the light positions have been transformed to the robot base frame and the IKP has been computed, a motion plan can be generated for the robot arm. To generate the motion plan, the rapidly-exploring random tree star (RRT*) method [Noreen et al. [2016], Karaman and Frazzoli [2011], Karaman et al. [2011], Nasir et al. [2013]] is used. The RRT* method is a sampling-based motion planning algorithm that is known for its efficiency and ability to handle complex environments. This method generates a sequence of robot arm trajectory points that connect the current position of the arm with the desired target pose.

For every point in an identified valid trajectory path, there exists a list of inverse kinematics solutions. Each trajectory point in the task space is represented as a layer. Hence for N trajectory points generated, there are N layers. The valid IK solutions in a layer are represented as nodes. Every node in an i^{th} layer is connected to all the nodes in the $(i+1)^{th}$ layer.

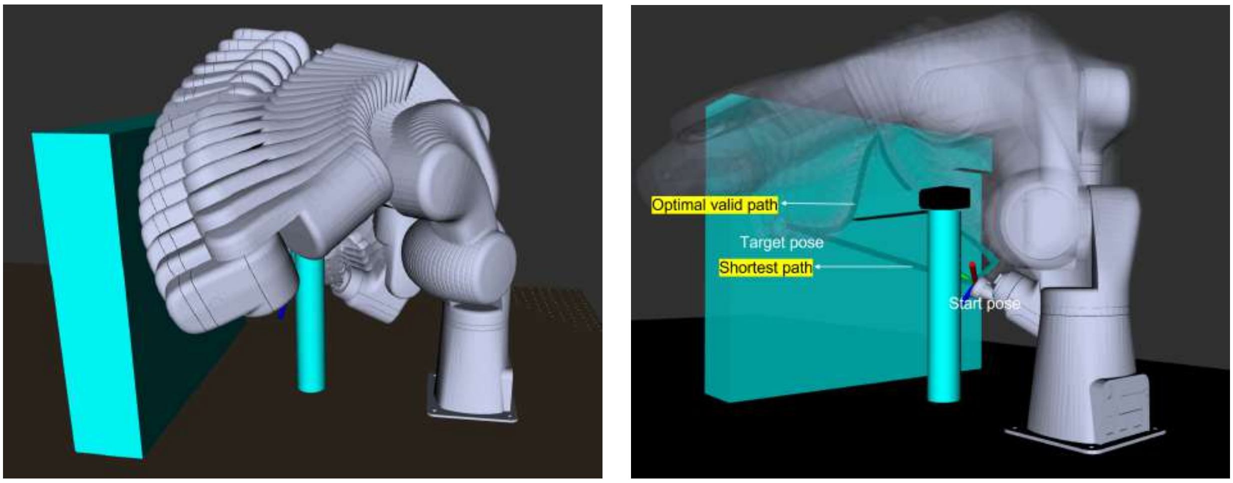
The optimality of a path is represented by the travel cost from the start point to the endpoint in the joint space. The optimal path is chosen such that the sequence of nodes that connects a path from layer 1 to layer N has the lowest possible total cost defined in eq. 2.2. The cost function takes into account various factors such as the distance between the nodes, the smoothness of the trajectory, and the energy consumed by the robot. By minimizing this cost function, the robot arm can move in the most efficient and safe way possible while still achieving the desired light positions.

$$C_i = \sum_{i=0}^{N-2} |p_i - p_{i+1}|^T.w, \tag{2.2}$$

$$\text{where} \begin{cases} i \text{ represents a layer,} \\ p_i \text{ is the joint position in layer } i, \\ w \text{ is the weights assigned to individual joints.} \end{cases}$$

w is a 6 length vector determined empirically. It is the same for all position p_i. For example, the base joint of the robot arm have the highest weight (the probability of collision is high), while the end-effector joint as the lowest weight (probability to hit an object almost null).

Figure.2.15 illustrates an example of the results obtained using the method used. As seen in the figure, the system finds the optimal valid path to reach the target pose which is both collision free and efficient in terms of distance traveled.



(a) Planned sequence of the robot poses during it's motion from starting pose to the target pose.

(b) Optimal valid path having the lowest cost vs the shortest path.

Figure 2.15: Planning of motion between two points with collision avoidance and optimal trajectory.

2.3.3 System calibration

In order to ensure the precision and accuracy of the system, self-calibration is performed using fiducial markers. This involves determining the relative pose of each system component, including the robot arm, camera, and XY platform. LightBot 1 setup of the system is on an optic table, where the relative pose of each component is fixed and known. However, LightBot 2 is configured to work in the absence of an optic table and hence requires calibration.

One important calibration that is performed is the hand-eye calibration [Horaud and Dornaika [1995], Strobl and Hirzinger [2006], Rémy et al. [1997]]. This is done by attaching a fiducial marker [Kalaitzakis et al. [2021]] to the end effector of the robot arm (where the light source is attached). This allows the system to accurately determine the relationship between the robot arm base coordinate system and the camera coordinate system. This is crucial for ensuring that the light source is positioned correctly in relation to the surface being imaged.

Another important calibration is the pose calibration between the camera coordinate system and the XY platform coordinate system. This is done by attaching a fiducial marker to the XY platform within the camera's field of view. This calibration allows the system to accurately determine the relative position of the surface being imaged with respect to the camera and the light source.

The use of fiducial markers and these calibration techniques ensures that the LightBot system is precise and accurate in its RTI acquisitions. This is particularly important for applications in Cultural Heritage, where high-quality imaging of valuable and delicate objects is crucial. Additionally, this system can be used in the field of industrial metrology, 3D scanning and other applications where precise and accurate measurements are needed.

2.3.4 System specifications

We built LightBot 1 and LightBot 2 acquisition systems and demonstrated its use to perform RTI acquisitions. Table.2.2 summarizes and compares the specifications of the two systems.

	Specification	
Parameter	**LightBot 1**	**LightBot 2**
Robot arm		
DoF	6	7
Repeatability	+/- 0.5mm	+/- 0.1mm
Max payload	1.25 kg	2.4 kg
Weight	9.5 kg	5.5 kg
Installation	Upright	Top down
Ingress rating	IP20	IP42
Max reach	600 mm	985 mm
Camera		
Model	Allied Vision Manta G504B	Allied Vision Proscilica GT6400
Resolution	2452 (h) 2056 (v)	6480 (h) 4860 (v)
fps	~20	~12
Lens	C mount, Focal length - 3mm / 6mm	F mount, Focal length - 28mm
XY Platform		
Stroke	40 cm	40 cm
Installation	Vertical	Horizontal
Overall system		
Dome size range	~5cm to 28 cm	~5cm to 60 cm
Surface size limit	40 cm	40 cm
Acquisition time	A typical acquisition having 50 light positions takes around 5 minutes to complete.	A typical acquisition having 50 light positions takes around 3 to 4 minutes to complete.

Table 2.2: LightBot system specifications

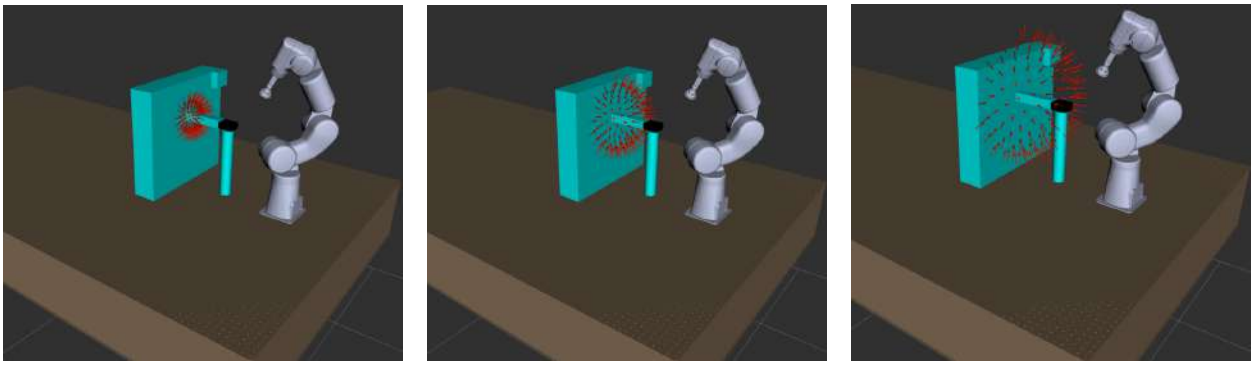
In summary, the LightBot system uses a combination of frame transformations, motion planning, and inverse kinematics to accurately position the light source and capture high-quality RTI images. The use of the ROS framework and other tools such as the RRT* algorithm and IKFast solver enable the system to generate efficient and safe motion plans while also providing a user-friendly interface.

2.4 Applications

Museum custodians and conservators continue to face challenges in digitizing high-resolution geometry data of cultural heritage (CH) objects for both documentation and analysis of artifacts. The built LightBot system is focused on providing efficient and scientifically reliable acquisition of RTI data. In this section, we demonstrate three key applications - surface adaptive RTI acquisition with an adjustable radius of the light dome, RTI acquisition of medium and large surfaces using data stitching methods, and automated batch processing for efficient large-scale RTI acquisitions. RTI is primarily used for relighting surfaces, so we have included only relighted images to show the results.

2.4.1 Surface adaptive virtual dome with adjustable radius

Conventional RTI acquisitions typically use a fixed number of homogeneously distributed light positions, regardless of the size and complexity of the surface being acquired. However, as discussed earlier in chapter 2, this approach can lead to shortcomings in capturing the full range of surface details. In order to fully realize the power of RTI in revealing surface phenomena, it is crucial to adapt the size of the virtual dome and the distribution of light positions to the specific surface being imaged. The LightBot system addresses this issue by providing the option for a virtual dome of variable sizes, which can be adjusted according to the size and complexity of the surface being imaged. This is illustrated in Figure.2.16 where we can see that we are able to achieve variable sizes RTI dome using LightBot depending on the size and complexity of the surface being acquired. This feature opens up new possibilities for automating the RTI imaging process and developing new methods for surface-adaptive RTI acquisitions. The ability to adjust the dome radius in the proposed LightBot system allows for greater flexibility in capturing surface details of Cultural Heritage (CH) objects. This is a significant departure from traditional RTI acquisition methods, which use a fixed dome size and light position distribution regardless of the surface being captured.



(a) Hemisphere radius = 15 cm. (b) Hemisphere radius = 20 cm. (c) Hemisphere radius = 25 cm.

Figure 2.16: Dome configurations with variable radius. The radius of the dome can be adjusted from a few cm to 30 cm with this robot arm.

To demonstrate the impact of dome size on RTI quality, we conducted multiple acquisitions of a 20th century metal print plate using LightBot with dome radii of 15 cm, 20 cm, and 25 cm. As shown in Figure. 2.17, the acquisition made with a dome radius of 25 cm resulted in a more evenly illuminated surface when relighted from different positions. This uniformity of illumination directly affects the quality of the various maps extracted from the RTI data. It's important to note that the way surface phenomena are highlighted can vary not only with the direction of the light source, but also with the distance between the light source and the surface. This highlights the importance of being able to adjust the dome radius to best suit the surface being captured, and the non-linear relationship between dome radius and surface illumination.

2.4.2 RTI of large surfaces using data stitching methods

The proposed LightBot system addresses the challenges of digitizing large-scale cultural heritage (CH) objects by enabling the acquisition of high-resolution geometry data for documentation and analysis. Another important application of the LightBot system is its ability to acquire images of large-scale surfaces without compromising resolution. This is achieved by breaking down the acquisition into multiple smaller parts, and using the XY positioning table to translate the surface. The system can perform RTI acquisition with a set of homogeneously distributed light positions at each translation, and later stitch the data together to create a complete RTI of the surface. This process is demonstrated using a canvas painting, where the surface is translated along the X and Y directions and RTI acquisition is performed at each translation. The results of the stitching are

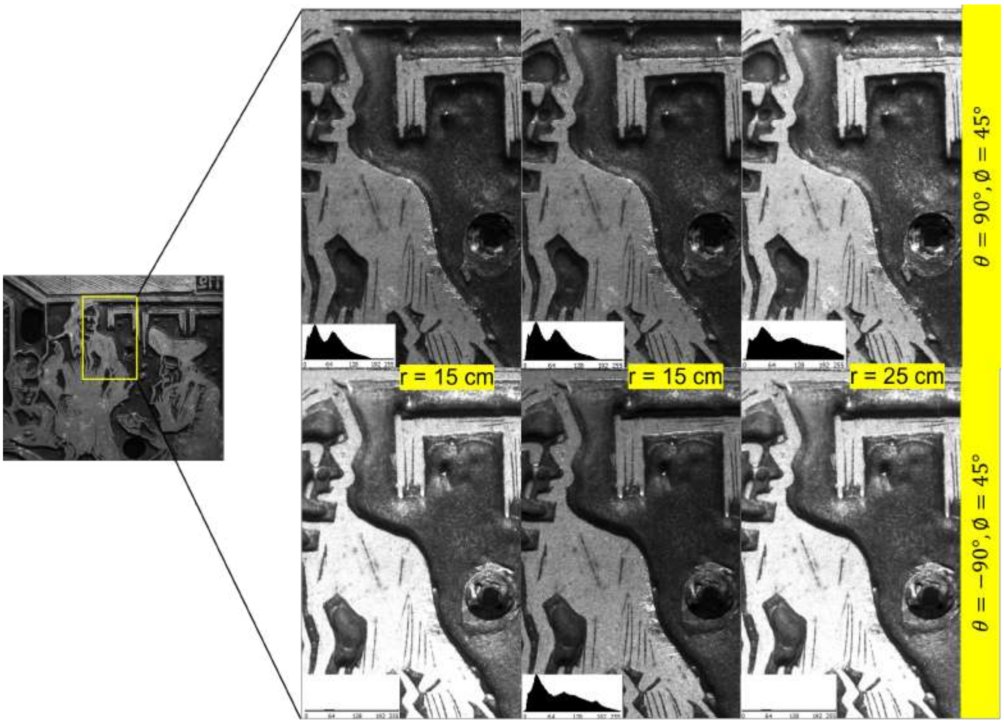


Figure 2.17: Relighting of the surface from RTI of the surface captured with dome of different sizes. In this example the surface is relighted from $\theta = 90°$, $\phi = 45°$ and from $\theta = -90°$, $\phi = 45°$.

illustrated in Figure. 2.19, where a sample of the relighted surface and its normal map are shown.

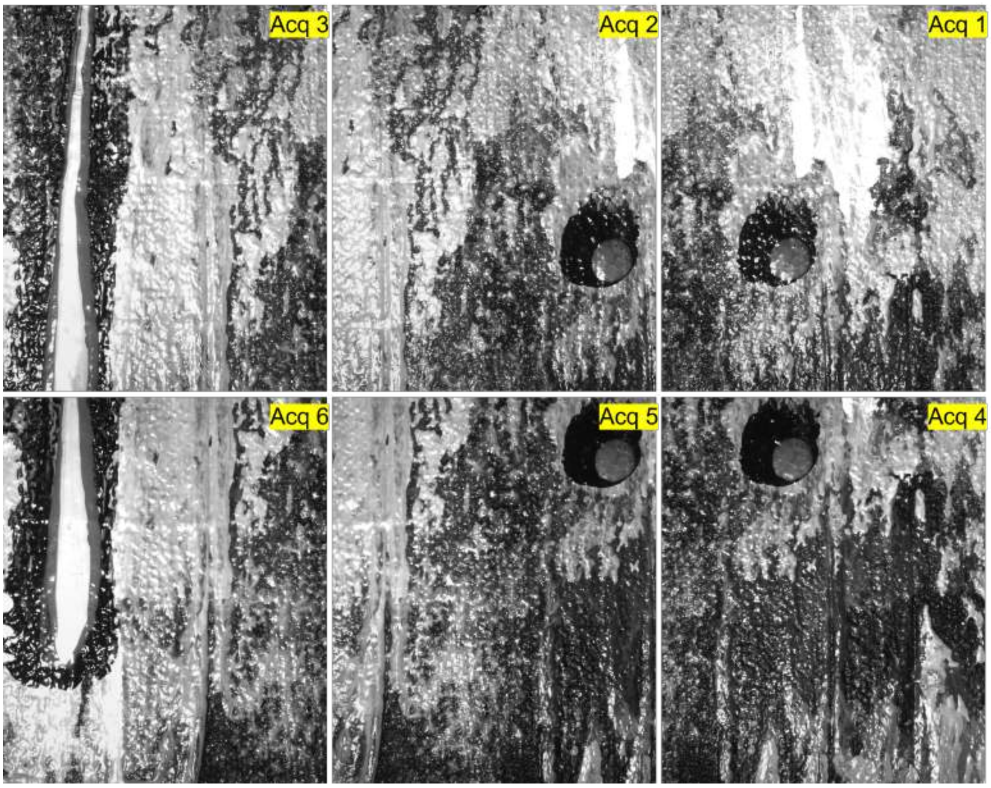


Figure 2.18: RTI acquisitions of a canvas painting (24.5 cm x 20 cm) in parts. There are totally 6 acquisitions each covering an area of 9.3 cm x 7.6 cm with 30 % overlap between each pairs of consecutive acquisitions.

Stitching RTI data raises several issues and are addressed in chapter 3. In this example, the stitching was performed using the method presented in the chapter for illumination correction and registration. In Figure.2.19, we have shown the result of the stiched RTI data representing the whole canvas painting surface and the normal map obtained from the stitched RTI data. Thus, the LightBot system offers a powerful tool for efficient and scientifically reliable acquisition of RTI data for large-scale cultural heritage objects, making it a valuable resource for museum custodians and conservators.

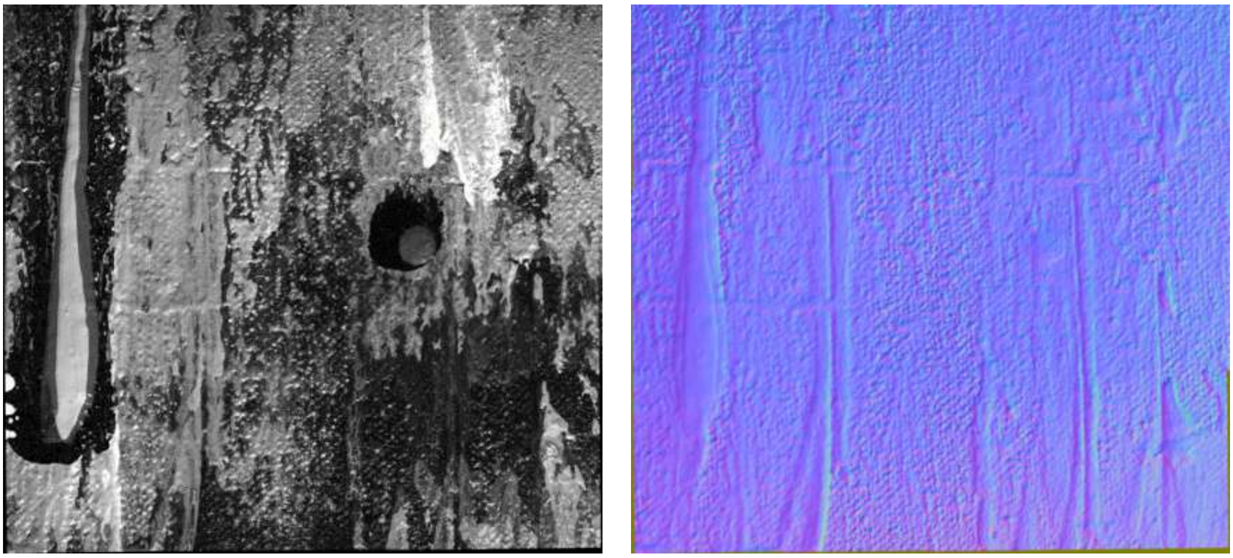

(a) Relighted image. (b) Normal map.

Figure 2.19: The acquired data are stitched to reconstruct the whole canvas painting. Visualization of the relighted image, normal map and directional slope obtained by processing of the RTI data using DMD Pitard et al. [2015] model fitting.

2.4.3 Batch acquisition

Another key features of the system is its ability to perform batch processing of RTI acquisition workflows, which is particularly useful in museums where manually performing RTI acquisitions on a large number of artifacts is both time-consuming and prone to errors.

Figure. 2.20 illustrates an example of how batch acquisition can be carried out using the LightBot system.

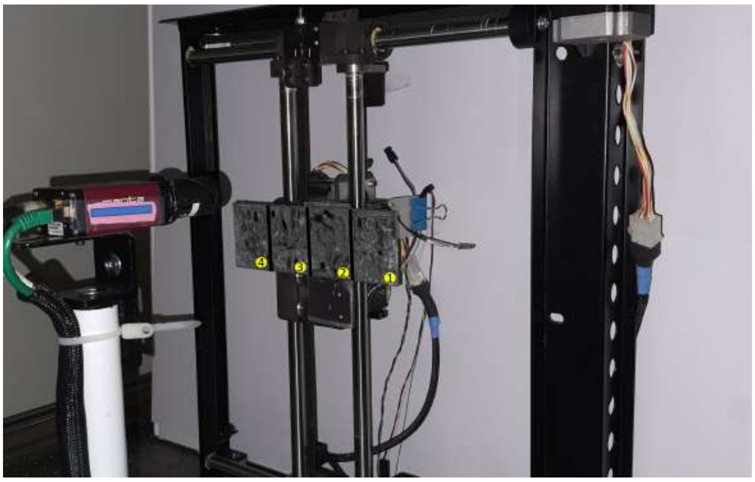


Figure 2.20: An example of batch acquisition where the system executes acquisition of 4 surfaces one after the other.

The image shows four surfaces attached to the XY stage, which can be acquired with minimal manual intervention by simply setting the initial position, final position, and intermediate steps along both the X and Y directions as batch acquisition parameters in the LightBot WebAPI. The light positions can be configured to be the same or distinct for each acquisition, depending on the surfaces being acquired.

The system then executes the series of acquisitions, generating a new acquisition directory and the corresponding .lp file for each. Table 2.3 compares the acquisition time and number of human interventions required when the acquisition of these 4 surfaces are performed using a conventional RTI system and that using the LightBot system. Although the acquisition time per surface is almost the same, avoiding intervention between each surface acquisition effectively reduces the overall acquisition time. This simple yet powerful automation of the bulk acquisition process saves a lot of time for the conservators and makes the acquisition process much easier and reliable overall.

	Acquisition time	Number of interventions
Conventional systems	$4 \times T_m$	4
LightBot	$T_b \approx T_m$	1

Table 2.3: Comparison of number of manual interventions and the acquisition time for acquiring 4 metal print plate samples using conventional system and that with the LightBot system.

The batch acquisition feature in LightBot is a direct offshoot of a specific need in museums as well as industries to perform RTI acquisition of metal coupons of sizes around 5 cm × 5 cm like in the study of corrosion and change Degrigny et al. [2007]. Additionally, the system's ability to perform batch acquisition of large-scale surfaces without compromising resolution is also a significant advantage. By breaking the acquisition into multiple acquisitions, each capturing a smaller part of the surface, it is possible to address the problem of tedious manual alignment and lack of existing systems for performing such acquisitions in the case of RTI. The system's ability to perform automated batch processing for efficient large-scale RTI acquisitions is a major advantage that makes the acquisition process much more efficient, reliable, and easier for conservators.

2.5 Conclusions

The LightBot is a state-of-the-art robotic system designed for the efficient and scientifically reliable acquisition of Reflectance Transformation Imaging (RTI) data. It is built on a modular architecture and is based on the widely-used Robotics Operating System (ROS) framework, which facilitates easy integration of various components and seamless communication between them.

One of the key features of LightBot is its ability to perform surface-adaptive RTI acquisitions, which allows for the selection of an optimal size and distribution of light positions based on the size and complexity of the surface being acquired. This is a major advancement in RTI acquisition as it addresses the shortcomings of conventional RTI acquisitions that use homogeneously distributed fixed light positions.

The system is also equipped with a simple and user-friendly web-based interface, built using the VUE frontend framework, that enables museum custodians and conservators to easily carry out RTI acquisitions by simply uploading the .lp file and setting a few parameters such as the surface size. This eliminates the need for manual intervention and makes the acquisition process more efficient and reliable.

Furthermore, LightBot is capable of performing batch processing of RTI acquisition workflows for large-scale objects, saving a significant amount of time for conservators and making the acquisition process more efficient and reliable. Additionally, the system is equipped with a pipeline for automated acquisition of multiple surfaces in batches, leveraging the degrees of freedom of the robot arm and the XY platform.

In addition to these features, LightBot also has the capability of performing data stitching for large-scale surfaces, making it possible to acquire images of large-scale surfaces without compromising resolution. The system also enables the option of having virtual dome of variable sizes, making it a powerful tool for automation of RTI imaging process and the development of new methods towards surface adaptive RTI acquisitions.

CHAPTER 3

A new framework for stitching multi-light image data

Chapter overview

To acquire large surfaces at high resolution, it is necessary to divide the surface into smaller parts and then stitch them together as a post-processing step. The objective is to have a better resolution of large fields acquisition while preserving the metrological aspect of the data. In this chapter we present a new framework for stitching large surfaces captured using MLIC techniques. The main challenge associated with this process is the need to correct for non-homogeneity of illumination in order to merge the data, as well as relighting the stitched surface without any artifacts. Our method combines multiple techniques and models of surfaces and illumination to efficiently stitch and relight large surfaces. This results in reduced reflectance inconsistencies in the overlapping regions and improves the quality of stitching. The effectiveness of this method is demonstrated through experiments on two virtual and two real surfaces showing different luminous behaviors.

3.1 Introduction

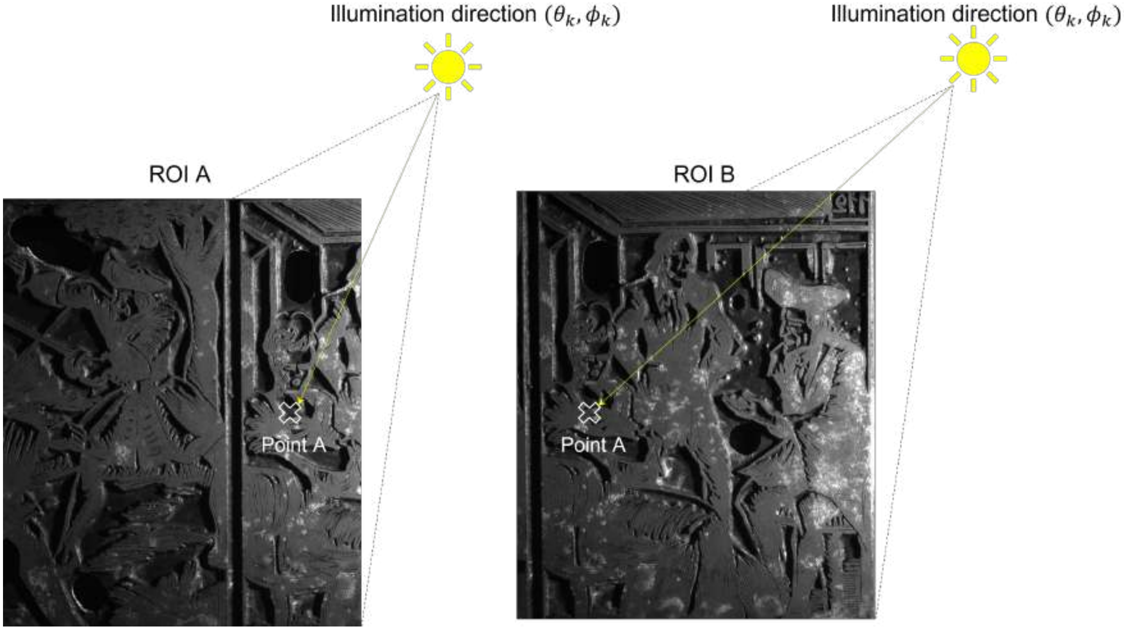
Image stitching is the process of combining multiple images with overlapping fields of view to create a high-resolution panorama Adel et al. [2014]. Image stitching is often applied to images with ambient light, which can be considered as uniform on the different subimages to be stitched. However, this is not the case in MLICs, where active light sources are used.

The active light used, are in this case not collimated. This can lead to uneven illumination, where the region of the surface closer to the light source being more illuminated than the region farther away as illustrated in 3.1. As a result, traditional image stitching techniques may not work effectively for RTI.

We observe the following three problems associated with the stitching of RTI data.

(a) A metal print plate surface under ambient lighting. Capture of the full surface in 1 image.

(b) Illumination across the surface. The left image shows the captured image of the illuminated surface, the right image illustrates the variation in energy distribution on the surface points through a heat map representation, with brighter areas indicating higher levels of energy and darker areas indicating lower levels.



(c) To maintain high resolution in RTI, we capture larger surfaces in parts. However, the overlapping region like point A receives different energy from the same light position in the two ROIs. As the surface translates, point A moves from a position closer to the light source to a position farther from it. This creates an inconsistency in the measurements which needs to be corrected prior to performing the RTI.

Figure 3.1: Illustration of RTI data stitching problem.

1. Misalignment of the images due to movement of surface (or camera).
2. Uneven distribution of light across the surface.
3. Presence of artifacts in the overlapping regions. These artifacts can manifest as visible seams.
4. Updated light positions representing a larger virtual dome to relight the whole surface after performing stiching.

The goal thus is to solve these problems. We stitch RTI images with overlapping fields of view in order to create continuous and seamless images, addressing the differences in illumination across the surface. To achieve this, we focus on correcting the illumination in each image rather than using blending methods during the stitching process.

3.2 Related work

3.2.1 Light direction interpolation and illumination correction

Illumination inconsistency across the surface is a common issue in computational photography rendering techniques such as photometric stereo Ackermann et al. [2015], Schlick [1994], RTI Malzbender et al. [2001a], and multilight imaging. It refers to the gradual change in the amount of light that reaches different points on a surface as the distance from the light source varies. This change in light intensity results in variations in luminance across the surface. A physical approach to achieve consistent illumination is to use a light source at a sufficiently large distance to be considered as a point. This ensures that all points on the surface receive equal amounts of light and thus produce consistent illumination. However, these methods can be impractical as they require very large light source and a large workspace. In addition to the light source, there are a variety of other factors that can cause illumination inconsistencies in photography or imaging, including the optics of the camera Debevec and Malik [2008], the geometry of the surface being imaged, internal reflections, and/ or shadowing. There are relatively few published works addressing the problem of correcting for illumination errors for RTI. Le Goïc et al. presents a new pixel-wise illumination correction method for reflectance transformation maging.

Huang et al. [2015] proposed a near-light illumination model for image relighting and 3D shape recovery, which compensates for each pixel's intensity based on its distance to the light source, removing the spot-light effect often seen in RTI captures. However, this method does not correct for errors in the incident ray angle at each pixel. McGuigan and Christmas [2020] presented an algorithm that corrects RTI images for non-uniform illumination and improves the determination of light direction. Their method fits a bi-quadratic polynomial to each input image and uses it to increase the intensity of underlit regions. While this method can flatten the intensity uniformity of the image, it does not guarantee the accuracy of the correction. This method can make the image's intensity levels more consistent, but alters the data and the estimation of the physical quantities(luminance) of the surface.

3.2.2 Image stitching

Image stitching is a process of combining multiple images to create a single, panorama. The process of image stitching involves detecting and matching distinctive features in the images, and then aligning and blending the images based on these correspondences. The technique has been widely used in various fields such as photography, robotics, virtual reality, and computer vision. Image stitching is a well-established technique with a rich history of research. One of the earliest and most well-known approaches for image stitching is the traditional feature-based method Zitova and Flusser [2003]. This method involves extracting features such as corners, edges, and textures from the images, and then matching these features across the images to determine the relative positions of the images. Examples of feature extraction methods include SIFT (Scale-Invariant Feature Transform) Lowe [2004], SURF (Speeded-Up Robust Features) Bay et al. [2006], ORB (Oriented FAST and Rotated BRIEF) Rublee et al. [2011]. These methods detect and describe local features in an image based on their scale, orientation, and shape. The transformation between two images is estimated from the correspondence between the set of features extracted from both images. The features are matched using algorithms like Brute Force Matcher (BFMatcher), K-Nearest Neighbours (KNN) and the incorrect matches (outliers) are removed using Random Sample Consensus (RANSAC) technique Fischler and Bolles [1981]. With a minimum of 8 matched pairs, homography matrix is calculated by solving Eq. 3.1.

$$I_1 = H * I_2, \tag{3.1}$$

where, H is the homography matrix and I_1 and I_2 are the pair of images.

The general form of homography matrix is

$$\begin{bmatrix} h_11 & h_12 & h_13 \\ h_21 & h22 & h23 \\ h31 & h32 & h33 \end{bmatrix} \tag{3.2}$$

where, h_{11} and h_{22} represent scaling factors along x and y directions respectively, h_{13} and h_23 represent translations along the x and y directions respectively, h_{12} and h_{21} represent shearing factors and h_{31} and h_{32} represent perspective transformations.

Given a pixel (x_1, y_1) in image I_1 and its corresponding pixel (x_2, y_2) in image I_2, the relationship between these pixels under the homography H can be written as:

$$\begin{bmatrix} k_{ij}x_1 \\ k_{ij}y_1 \\ k_{ij} \end{bmatrix} = H \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} \tag{3.3}$$

where k_{ij} is a scale factor that makes the equation homogeneous.

The alignment between images can be rigid (translations and rotations), affine (scale, shears), homographies, or complex large deformations models. This transformation warps one image to the coordinate frame of another image.

Another approach is to use intensity-based methods where the transformation is estimated by maximizing a similarity metric between two images formulated as an optimization problem. Sum of absolute difference (SAD), correlation ratio (CR) , correlation coefficient (CC), sum of squared difference (SSD) and mutual information (MI) are some of the common similarity measures used in this approach Goshtasby [2012], Roche et al. [1998]. These methods performs extremely well for images of same scenes taken in different time intervals and is ideal for coarse to fine registration.

Once the images are aligned, the overlapping regions are blended together using a blending technique such as feathering Porter and Duff [1984], averaging Adelson et al. [1984] or Poisson blending Pérez et al. [2003].

With the rise of deep learning, the performance of image stitching has been greatly improved. However, a survey on deep learning based approaches is beyond the scope of this thesis as we focus on the correction part more than the alignment of the images. Haskins et al. [2020] discusses various deep learning approaches for the image registration (alignment) problem.

3.3 Proposed method

Acquiring RTI of large surfaces can be challenging as it requires capturing the entire surface in high resolution. One approach to achieve this is by setting up a camera in a fixed position and translating the surface between each acquisition. Another approach is to use multiple RTI domes with identical specifications to perform multiple acquisitions that cover the entire surface. Figure.

3.2 illustrates the process of acquiring RTI of a large surface by dividing it into parts (ROI A and ROI B). The relative position of the overlapping region in ROI A and ROI B can vary greatly with changes in light direction, which can lead to inconsistencies in illumination and result in visible seams and artifacts.

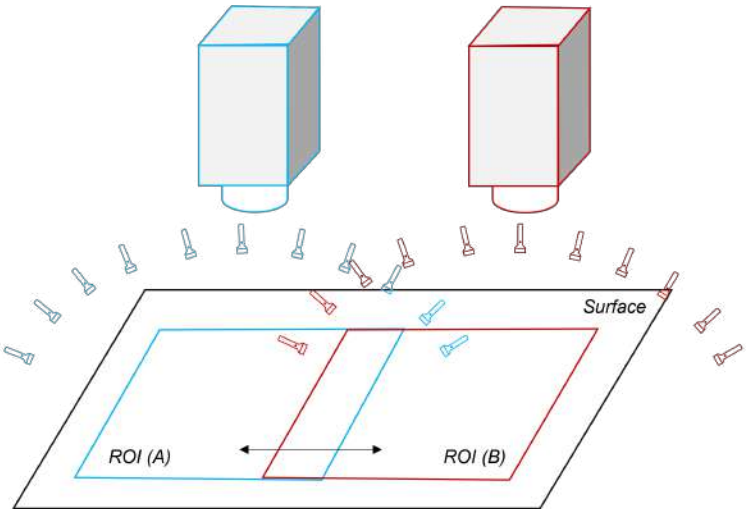


Figure 3.2: The diagram illustrates the setup for acquiring RTI of large surfaces. The surface is divided into two regions of interest (ROI A and ROI B), with an overlapping region in between. RTI is acquired separately for each of the regions of interest.

Figure 3.3 illustrates the process involved in stitching together data from two acquisitions, referred to as Acq 1 and Acq 2. The process consists of four steps:

- Proper alignment of images between different light positions is crucial in RTI acquisition. This is because even the slightest movement of the camera or the surface during the acquisition process can lead to misalignment of the images, affecting quality of the final output. An initial registration of images between different light positions in the same acquisition is carried out to ensure the alignment.

- Pixel-wise light corrections are applied to the images, having the light positions known. These corrections help to compensate for variations in brightness in the overlapping region between the two acquisitions that occur due to relative differences in the light position.

- The light-corrected images from the two acquisitions are registered and spliced together. Light corrections does not completely remove the illumination inconsistencies. There can still be a subtle seam visible which are removed by using feather blending technique. This results stitching images together smoothly to create a seamless composite image.

- The light positions for modelling are updated to cover a larger hemisphere. This step helps to improve the accuracy and reliability of the resulting image by taking into account a wider range of lighting conditions. We update the light positions to a larger virtual dome with a radius of $r + \frac{d}{2}$, where r is the original hemisphere radius and d is the distance, the surface was translated as illustrated in Figure. 3.4.

3.3.1 Alignment of images between the light positions

Ensuring the alignment of images between different light positions is crucial in RTI, as it is essential for the camera and surface to remain completely stationary during the acquisition process. Every pixel in an image captured must correspond to the exact same surface point in every other image. In a traditional RTI acquisition system, the surface is placed on a stationary platform. However, in

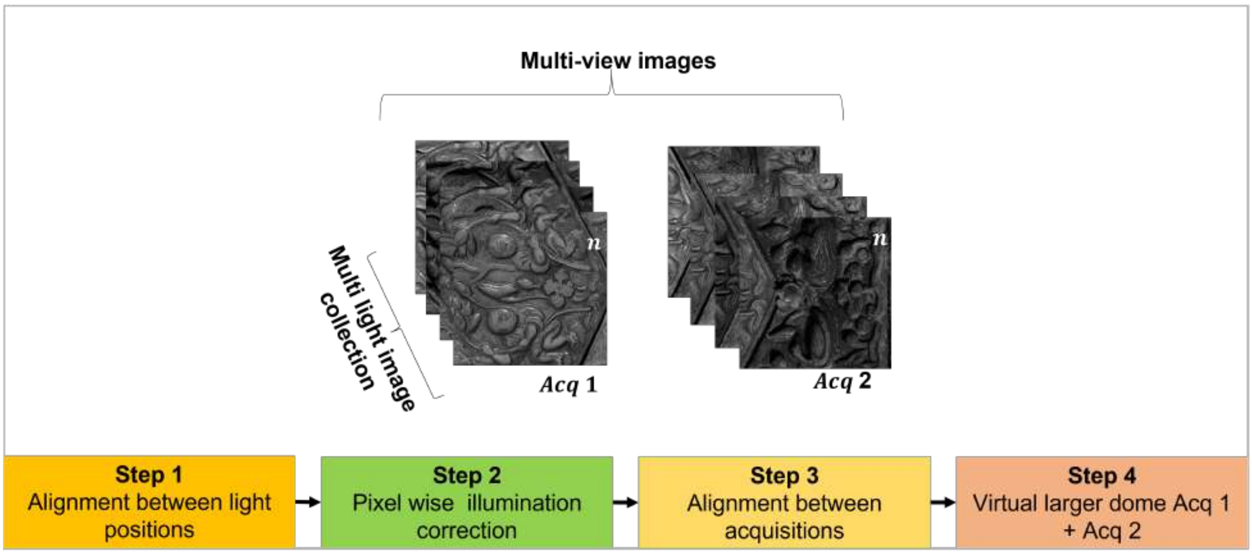


Figure 3.3: Overview of the framework of the proposed approach

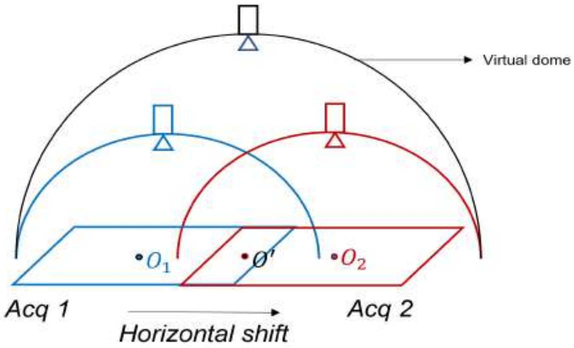


Figure 3.4: Larger virtual dome with a radius of $r + \frac{d}{2}$ used for relighting the stitched data.

our set up, as the acquisition of a surface is carried out in parts, the surface is attached to a moving platform, and the light source is attached to a robot arm. The introduction of these mechanical components can cause vibrations, difficult and expensive to eliminate them completely. To address this issue, we employ a technique called subpixel level alignment - where images captured at all light positions in an acquisition are aligned with respect to the first image of the acquisition, effectively compensating for any slight displacement of the surface.

To achieve this, we adopt a similarity measure called *Enhanced Correlation Coefficient* (ECC) Evangelidis and Psarakis [2008] for performing intensity based registration in the gradient domain. We used gradient domain as it is more invariant to the changing photometric conditions. Figure. 3.5 shows this alignment process with ECC co-efficient. Most intensity based registration algorithm assumes constant illumination condition between the images which is violated in our case. ECC addresses this by including an additional photometric transformation $\psi(I, \alpha)$ to account for the photometric changes. The optimization problem is defined as eq. (3.4).

$$\underset{\min}{E}(p, \alpha) = \underset{p,\alpha}{min} \sum_{x \epsilon \tau} \left| I_0(x) - \psi(I_n(\phi(x;p)), \alpha) \right|^2 , \tag{3.4}$$

where, p represents the parameters of the geometric transformation (translation, rotation, scaling, affine) that aligns the input image I_n with the reference image I_0, α represents the parameters of the photometric transformation that adjusts the appearance of the input image to match the reference image (brightness, color) , $E(p, \alpha)$ is the error function with respect to the parameters p and α. The error function is defined as the sum of squared differences between the reference image $I_0(x)$

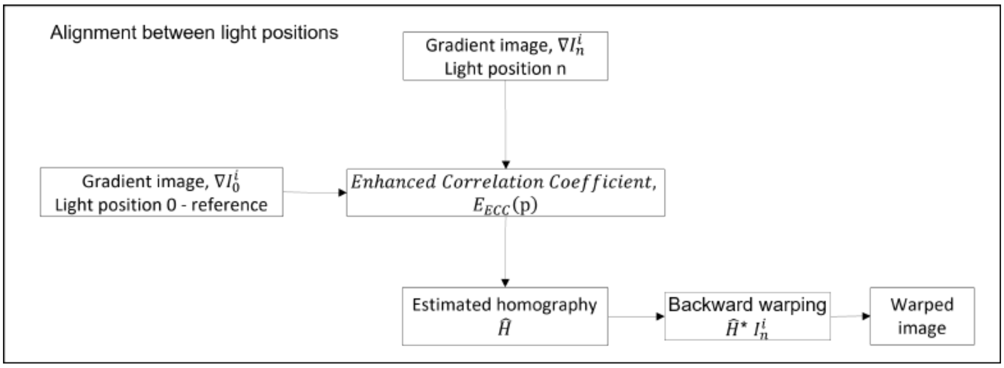


Figure 3.5: Alignment between the light positions in an acquisition.

and a warped version of the input image $I_n(\phi(x;p))$, where $\phi(x;p)$ is the geometric transformation. The function $\psi(I_n(\phi(x;p)), \alpha)$ represents the photometric transformation function. x represents the pixel coordinates in a template region τ in the image. The template region τ is a subset of pixels in the reference image that is used to guide the alignment of the input image. This subset contains distinctive features or object that can be used to match and guide the alignment

Evangelidis and Psarakis [2008] proposes the criterion below to quantify the performance of the warping transformation in their algorithm

$$E_{ECC}(p) = \left\| \frac{\bar{i}_0}{\|\bar{i}_0\|} - \frac{\bar{i}_n(p)}{\|\bar{i}_n\|} \right\|^2, \tag{3.5}$$

where $i_0 = [I_0(x_1)I_0(x_2)...I_0(x_k)]$ is the reference vector and $i_n(p) = [I_n(x_1)I_n(x_2)...I_n(x_k)]$ is the warped vector.

3.3.2 Illumination correction

Illumination correction is the process of removing or compensating for variations in brightness or intensity in an image that are due to factors like differences in the distance and angles of incidence of the light across a surface. It is necessary for achieving more accurate and reliable reflectance modelling of larger surfaces.

Many approaches can be employed to correct variations of illumination in an image. Some approaches involve modelling the illumination inconsistency across the image, while others involve using machine learning techniques such as neural networks. In this section, we describe our illumination correction approach by modelling the non-homogeneity, the implementation of this method to adjust the image intensities. Modeling the non-homogeneity in illumination in multi light acquisition data can be a complex task, as it often requires taking into account the reflectance properties of the surface being imaged, as well as the characteristics of the light source and the imaging system. In our approach, we first compute the surface normals and albedo of the surface from their respective multi-light photometric stereo acquisition data. These surface properties, along with the light positions, are then used to model the brightness falloff across the surface.

Surface normals and albedo

Lambertian Photometric Stereo based on Sparse Regression by Satoshi Ikehata and Aizawa [2012], Satoshi Ikehata and Aizawa [2014] allow to estimate surface normals and albedo in photometric stereo using L1 residual minimization. The surface normals and albedo are estimated by minimizing the L1 residual between the observed and the estimated pixel intensities under a given lighting condition. The estimated pixel intensities are given by the dot product of the surface normals and the lighting directions, multiplied by the albedo. The L1 residual is defined as the sum of the absolute differences between the observed and estimated pixel intensities. The surface normals and albedo are obtained by solving the optimization problem:

$$\min ||Ax - b||_1, \tag{3.6}$$

where A is a matrix containing the lighting directions, x is a vector containing the surface normals and albedo, and b is a vector containing the observed pixel intensities. This approach is particularly effective for Lambertian surfaces, and has been shown to yield accurate results. In our method, we employ this technique to obtain surface normals and albedo estimates for Lambertian surfaces.

Robust Photometric Stereo (RPS) via Low-Rank Matrix Completion and Recovery by Lun Wu and Ma [2010] is another approach which is particularly useful for handling noise and outliers in the data, and is able to handle surfaces in general rather than just Lambertian surfaces. In this method the photometric stereo data is represented as a matrix, where each column corresponds to a pixel on the surface and each row corresponds to a different lighting condition. This matrix is then decomposed into two low-rank matrices: one representing the surface normals and the other representing the albedo. To estimate the surface normals and albedo, the authors use an optimization problem that seeks to minimize the rank of the matrix while simultaneously enforcing constraints on the surface normals and albedo.

Mathematically, the optimization problem is expressed as follows:

$$\text{minimize } \operatorname{rank}(N) + \operatorname{rank}(A), \tag{3.7}$$

subject to:

$$N * A = M, \tag{3.8}$$

where N is the matrix of surface normals, A is the matrix of albedo values, and M is the matrix of photometric stereo data. The rank of a matrix is a measure of its complexity, and the goal is to find the lowest rank possible while still accurately representing the data. To solve this optimization problem, we use an iterative algorithm that alternately updates the estimates of N and A. At each iteration, the estimates of N and A are refined using a low-rank matrix completion method, and the rank of the resulting matrices is minimized using a nuclear norm minimization algorithm. This process is repeated until convergence, the estimation of N and A are obtained.

Overall, this method provides a robust way to estimate surface normals from photometric stereo data, even in the presence of noise and outliers. It is well-suited for handling surfaces in general and not just Lambertian surfaces, however it is computationally expensive. In our method we employ this technique for complex and specular surfaces that cannot be properly handled by the previously discussed Satoshi Ikehata and Aizawa [2012] approach.

Modeling the non-homogeneity in illumination

We propose two innovative methods for modeling the the non-homogeneity in illumination across the surface of an object in order to correct for illumination in the images for stitching. These approaches provide new ways to splice MLIC data without the need of relying on blending the overlapping region alone to eliminate the artifact, allowing for more accurate restoration of the captured information. Once the surface normals and albedo are computed, we use them along with the light positions to model the illumination inconsistency.

1. *Using inverse square law, Lambert's cosine law and surface properties*

Le Goïc et al. proposed a new method for correcting uneven illumination in images acquired using Reflectance Transformation Imaging (RTI) to perform pixel wise reflectance modeling. We implemented this approach and applied for stiching. The method uses the Lambert's cosine law Weik [2001], and calculate the intensity of the light at each pixel using the following equation:

$$I = albedo \times \frac{cos(\theta_i)}{\pi}, \tag{3.9}$$

where I is the intensity of the light at a given pixel, albedo is the reflectance of the surface at that pixel, and θ_i is the angle of incidence at that pixel.

However, this equation only accounts for the angle of incidence and does not take into account the distance between the light source and the surface. To account for this, we modified the equation to include the inverse square law Voudoukis and Oikonomidis [2017]. The inverse square law states that the intensity of the light at a given point on a surface is inversely proportional to the square of the distance between the light source and the surface.

Thus, the modified equation for calculating the intensity of the light at each pixel becomes:

$$I = albedo \times \frac{cos(\theta_i)}{\pi} \times \frac{1}{d^2}, \tag{3.10}$$

where d is the distance between the light source and the surface at a given pixel.

Finally, we need to take into account the amount of light absorbed and reflected by the surface at each pixel. We can do this by modifying the equation once again:

$$I = albedo \times \frac{cos(\theta_i)}{\pi} \times \frac{1}{d^2} \times (1 - absorption - reflection), \tag{3.11}$$

where absorption and reflection are the amount of light absorbed and reflected by the surface at a given pixel, respectively. They are calculated as:

$$absorption = albedo \times abs(N[2]), \tag{3.12}$$

$$reflection = (1 - albedo) \times (1 - abs(N[2])), \tag{3.13}$$

where N is the surface normal and N[2] is the z-component of the surface normal at the pixel.

Using the corrected intensity, we create a new image that represents the corrected illumination across the surface. This corrected illumination is then used to correct the original images acquired in the RTI acquisition.

2. Fitting the MLIC data to a BRDF model

We propose correcting the illumination by fitting a non-spectral BRDF Montes and Ureña [2012]) model to the acquired MLIC data. The BRDF model defines the relationship between the incident and outgoing light directions and the reflectance of the surface. It can be used to predict the intensities of the light sources for any given surface normals and incident and outgoing directions. As explained in chapter 1 there are many models that can be used to describe the BRDF of a surface. In our approach we use the BRDF model based on Beckmann distribution Walter et al. [2007] based on the microfacet theory, which describes the reflectance of a surface as the sum of contributions from many small, randomly oriented microfacets. The Beckmann model is a widely used model that accurately captures the physical properties of surfaces such as the roughness and anisotropy of the surface. It is also computationally efficient and has only a small number of parameters, making it a good choice for fitting RTI data.

The Beckmann BRDF model, is defined by the following equation:

$$D(\mathbf{N} \cdot \mathbf{H}) = \exp\left(-\frac{(\mathbf{N} \cdot \mathbf{H})^2 - 1)^2}{\text{roughness}^2 (\mathbf{N} \cdot \mathbf{H})^2}\right), \tag{3.14}$$

$$F(\mathbf{N} \cdot \mathbf{L}, \mathbf{N} \cdot \mathbf{V}) = \text{reflectance} + (1 - \text{reflectance}) \cdot (1 - \mathbf{N} \cdot \mathbf{L})^5, \tag{3.15}$$

$$\text{BRDF}(\mathbf{N} \cdot \mathbf{L}, \mathbf{N} \cdot \mathbf{V}) = \frac{D(\mathbf{N} \cdot \mathbf{H}) \cdot F(\mathbf{N} \cdot \mathbf{L}, \mathbf{N} \cdot \mathbf{V})}{4 \cdot \mathbf{N} \cdot \mathbf{L} \cdot \mathbf{N} \cdot \mathbf{V}}, \tag{3.16}$$

where $\mathbf{N}$ is the surface normal, $\mathbf{H}$ is the halfway vector, a vector that lies halfway between the incident and outgoing directions, $\mathbf{L}$ is the incident direction, $\mathbf{V}$ is the outgoing direction, and roughness and reflectance are parameters of the model.

The term $D(\mathbf{N} \cdot \mathbf{H})$ in the BRDF equation is the roughness term, which describes the roughness of the surface. The term $F(\mathbf{N} \cdot \mathbf{L}, \mathbf{N} \cdot \mathbf{V})$ in the BRDF equation is the reflectance term.

To fit the BRDF model to the MLIC data, we use an optimization algorithm to minimize the objective function that measures the difference between the observed intensities and the intensities predicted by the BRDF model. The objective function is defined as follows:

$$f(\mathrm{BRDF}, \mathbf{i}, \mathbf{O}, \mathbf{N}, \mathbf{S}, \mathbf{I_{meas}}) = \sum \left(\mathbf{I_{meas}} - \mathbf{I_{pred}}\right)^2 , \tag{3.17}$$

where, $\mathbf{i}$ is the set of incident directions, $\mathbf{O}$ is the outgoing direction, $\mathbf{N}$ represents the surface normals for each pixel, $\mathbf{S}$ represents the surface positions for each pixel, and $\mathbf{I_{meas}}$ and $\mathbf{I_{pred}}$ are the measured intensities and predicted intensities for each image and pixel.

The optimization algorithm can be used to find the values of the roughness and reflectance parameters that minimize the objective function. Once the optimal values are found, the BRDF model can be used to predict the intensities of the light sources for any given surface normals, incident and outgoing directions. By using the corrected incident and outgoing directions for each pixel, we apply the BRDF model to obtain illumination corrected images.

3.3.3 Registration of data between the acquisitions

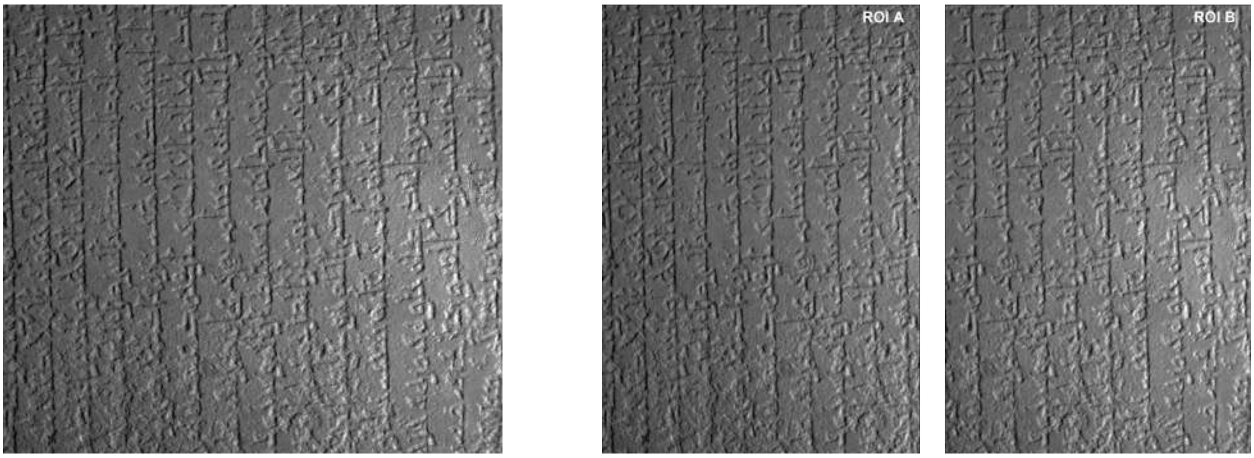
Stitching a pair of images requires both the images to have a common overlapping region. Our acquisitions are sequenced such a way that, there is at least 30% overlap. Ideally the nature of the displacement between each acquisitions is either horizontal or vertical shift. However due to mechanical noise, there is always a small deviation. We used SIFT feature based image registration is used for our application.

For a pair of acquisitions containing n number of light positions, there are n number of registrations carried out. Ideally the n number of homographies obtained from these registrations should be identical. Any deviation in the homographies is either due to improper alignment of images within an acquisition or due to error in the feature extraction and matching. We choose the homography from the image pair that has the highest number of good matches and lowest cost as the common homography for all the light positions in the acquisition pair.

3.4 Experiments

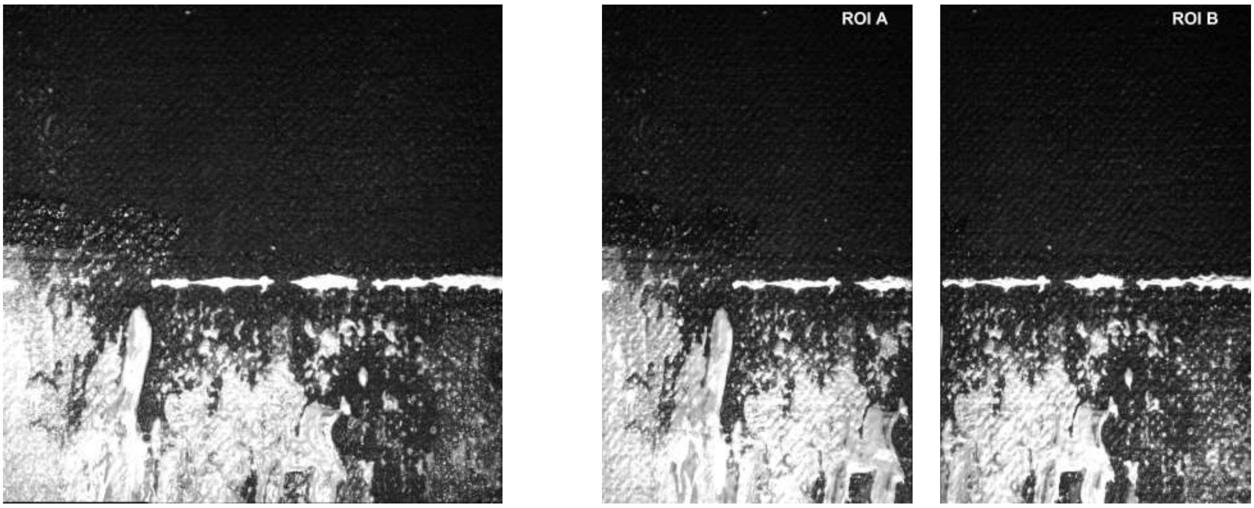
3.4.1 Dataset

To validate our approach, we first tested the methods on virtual surfaces using the Blender rendering software. We then tested the method on real surfaces acquired using the robot arm-based acquisition setup, as described in chapter 2. We used two virtual surfaces - a tombstone für Kunst und Gewerbe Hamburg [2021] (diffuse) and Sutra container, a Gilt bronze with incised and high relief decoration of Art [2022] (specular) and two real surfaces - a canvas painting (diffuse) and a zinc metal print plate (specular) to test and validate our methods. RTI aquisitions were performed having 50 light directions homogeneously distributed over the hemisphere. The surface was acquired in parts translating using a XY platform so that there is atleast 30 % overlap between the acquisitions. Table. 3.1 summarizes the surfaces we used for our experiments and testing along with their sizes. Figure 4.6 shows the full size of the images of the surfaces in ambient light and the respective ROIs used for acquiring them in parts.



Tombstone

Sutra container



Canvas painting

Zinc print plate

Figure 3.6: Surfaces used in the experiments and testing of the proposed methods. Tombstone and sutra container are virtual surfaces. Canvas painting and the zinc print plate are real surfaces. On the left are the full size images of the surfaces captured under ambient light and the right are the subimages (ROIs) used for capturing the surfaces in parts

Table 3.1: Surfaces acquired and used for the experiments and testing of our methods

Surface	**Size**	**Reflectivity**
Virtual surfaces		
Tomb stone	71 cm H x 49.5 cm W	Diffuse
Sutra container	10.2 cm H x 6.5 cm W	Semi specular
Real surfaces		
Canvas painting	22 cm H x 13 cm W	Diffuse
Zinc print plate	4 cm H x 10.2 cm W	Semi specular

3.5 Results and discussions

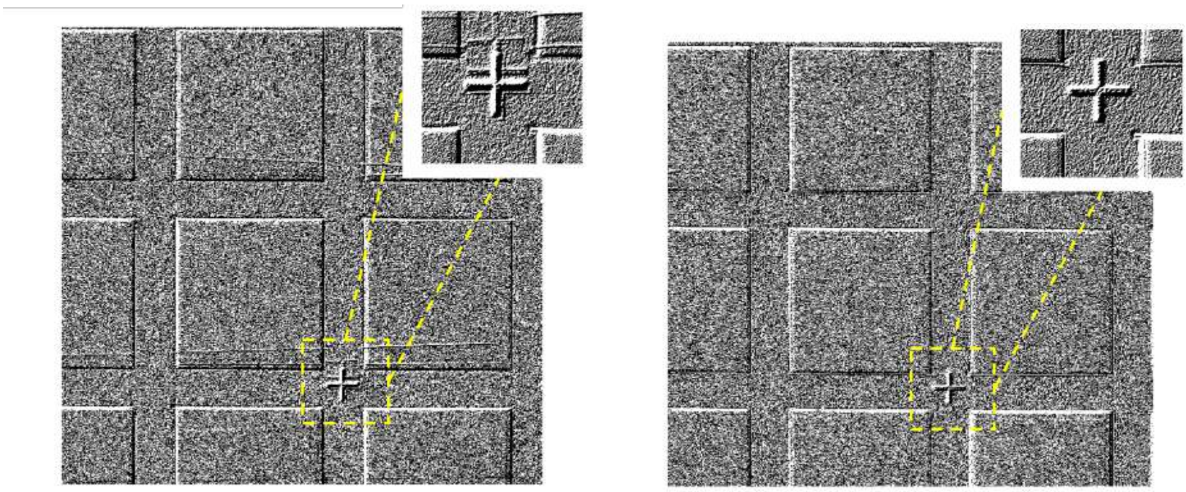
3.5.1 Alignment of images between the light positions

After the acquisition of data, the next step is to make sure the images in an acquisition are completely aligned using the Enhanced Correlation Coefficient approach described previously. To illustrate better the alignment of images, we have used a color checkerboard that has a cross mark. To visualize the alignment between the images, we add the gradient of all the images in an acquisition together whereby the edges and other prominent features appears distinct:

$$\sum_{i=0}^{n}\left(0.5\frac{\partial I_i}{\partial x}+0.5\frac{\partial I_i}{\partial y}\right), \tag{3.18}$$

where, n is the total number of images in an acquisition.

The figure (3.7) illustrates the comparison of the acquisition data before and after sub-pixel realignment. On the left image, we can see that the cross position varies across the images. On the right, after alignment, only one cross is visible. It can be observed that after applying sub-pixel registration, the images are properly aligned. This step addresses the misalignment due to minute mechanical vibrations.

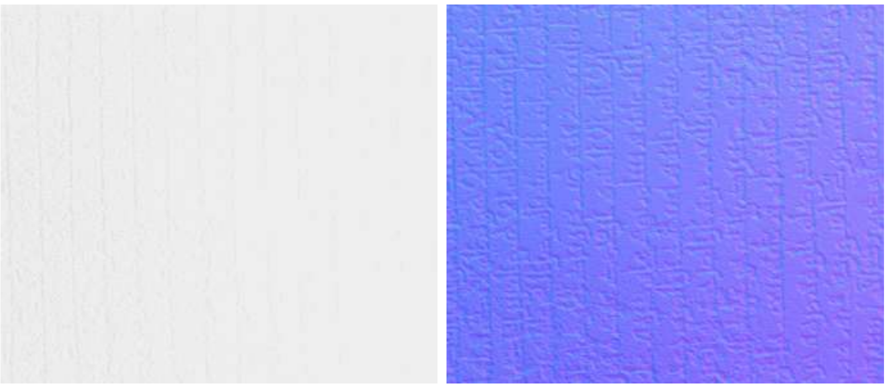


(a) (b)

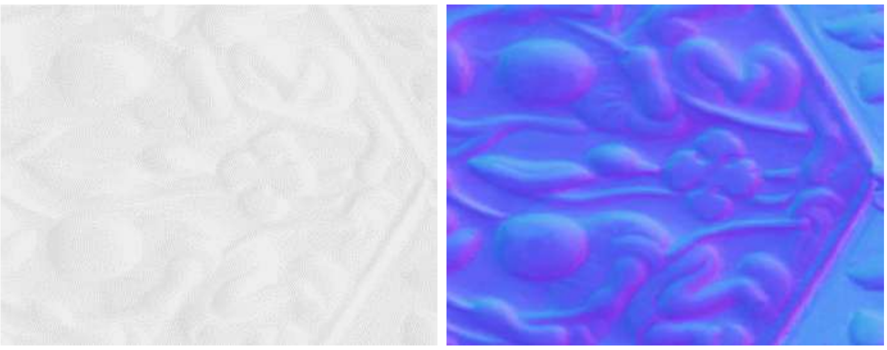
Figure 3.7: Re-alignment of images within an acquisition. (left) represents the acquisition before sub-pixel realignment and (right) represents the acquisition after sub-pixel realignment.

3.5.2 Normal map and albedo estimation

We obtained the normal maps and albedos of the surfaces using the methods detailed previously. For the tomb stone (diffuse), L1 residual method is applied and for the sutra container (specular) RPS method is applied. Figure. 3.8 represents the albedo and surface normal maps obtained for the virtual surfaces and 3.9 represents that of the real surfaces. In the case of real surfaces, there is no ground truth to allow the estimation of the MSE errors. Table 3.2 presents the mean squared error of the normal maps obtained.



(a) Tomb stone (L1 residual method)



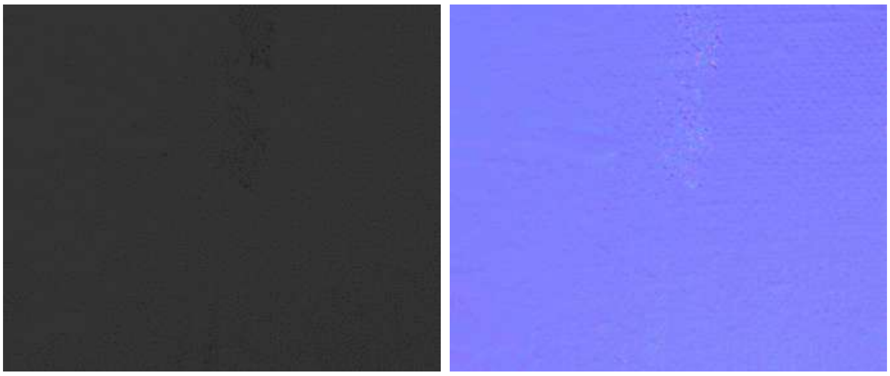
(b) Sutra container (RPS method)

Figure 3.8: Estimated albedo and normal map of the virtual surfaces.

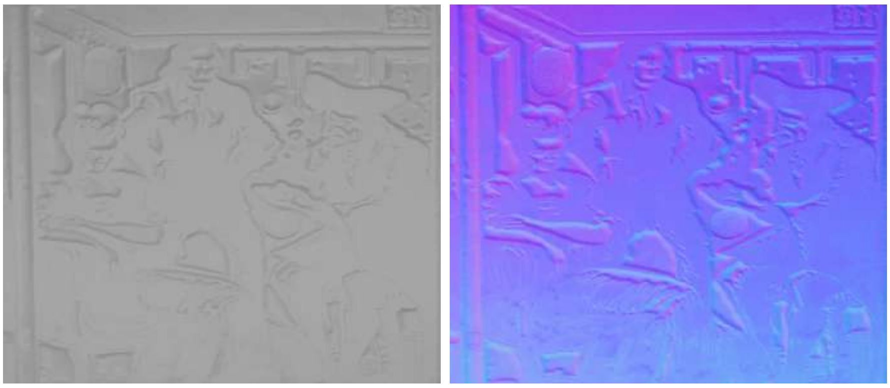
Table 3.2: Mean squared error of normal maps obtained using the L1 residual minimization and RPS methods for two virtual surfaces, compared to their respective ground truth normals.

	Tomb stone		Sutra container	
Method	L1 residual	RPS	L1 residual	RPS
MSE	6.12	7.91	49.13	8.18

We tested the approaches for accurately obtaining normal maps and albedos of surfaces using RTI acquisition of virtual surfaces in Blender. The ground truth normals and albedos can be obtained in the software, which allows us to compare the performance of the different methods. Our results showed that the L1 residual method is efficient and well-suited for diffuse surfaces, while the RPS approach is slower but handles specular surfaces more effectively.



(a) Canvas painting (L1 residual method)



(b) Metal print plate (RPS method)

Figure 3.9: Estimated albedo and normal map of the real surfaces.

3.5.3 Illumination correction and stitching

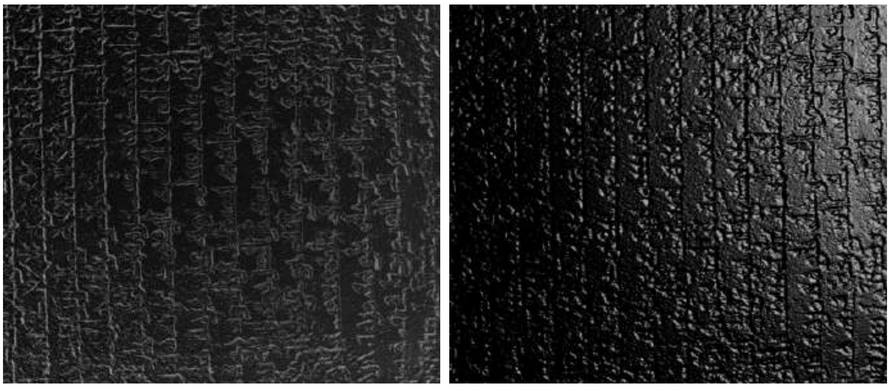
Illumination corrections of every single image in each acquisition set were carried out using both the inverse square law and Lamberts consine law approach and the BRDF model fitting approach. Figure. 3.10 shows examples of the corrected images of the simulated surfaces.

Figure 3.11 shows examples of corrected images of real surfaces. Unlike virtual surfaces, real surface acquisition contains noise that must be taken into account. It has been observed that illumination correction using Lambert's cosine law and the inverse square law can work reasonably well in achieving consistent illumination and removing artifacts in the overlapping region. However, for specular surfaces, this correction is not sufficient.

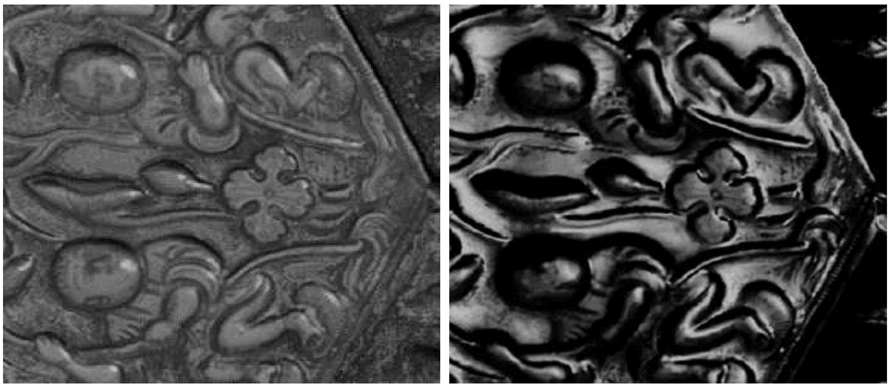
3.5.4 Stitching and relighting

In order to evaluate the performance of our proposed illumination correction methods, we compare the results of stitching of original images, images corrected using the inverse square law method, and images corrected using the BRDF fitting method. The results of stitching are shown in Figure 3.12.

Images on the left shows the result of stitching the original images without any correction applied. The seam between the overlapping regions is clearly visible in the stitched images of the surfaces before corrections, as indicated by the white arrow. Our proposed method of correction using the inverse square law and the Lamberts cosine law have shown to effectively address the seam and artifact problem in virtual surfaces with low noise levels. However, it has been observed that these methods are less effective in real surfaces, although they do reduce the visibility of seams and artifacts significantly. On the other hand, the correction using the BRDF fitting approach has been found to be effective in removing the seams completely, regardless of the surface type. The figure 3.14 displays a graph of the mean squared error of the differences between pixels in the overlap region



(a) Tomb stone



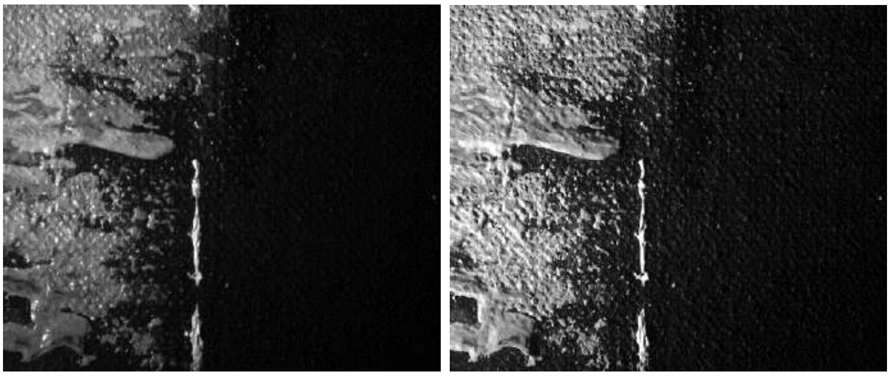
(b) Sutra container

Figure 3.10: Examples of the corrected images of the virtual surfaces. Left represent the image corrected using the inverse square law and Lamberts cosine law, and the right represents the image corrected using the BRDF fitting approach.

of the stitched images with and without correction. The first approach is effective for correcting virtual acquisition data, but did not perform as well when applied to real acquisition data with noise, resulting in visible seams after correction. While the second method is more computationally expensive, it effectively corrects the uneven illumination of the surface. The BRDF fitting approach shows promising results for specular surfaces. Although uniform illumination and artifact removal are achieved, it is to be noted that there is a significant change in the appearance of the surface. The Beckmann model, a microfacet distribution model used to simulate surface reflections, is based on a theoretical understanding of how light interacts with rough surfaces rather than being empirical. The appearance of the surface is sensitive to the roughness parameter in the model, making it difficult to obtain the right parameters to reproduce the ground truth appearance.

In Figure 3.13, we present examples of relighted images to illustrate the effectiveness of light correction in RTI. The figure compares the relighted image obtained from the RTI modeling of the original stitched images, the RTI modeling of the stitched images with traditional alpha blending in the overlapping region, and the RTI modeling of the corrected stitched images. It can be observed that in the case of the original and traditional stitched dataset, the relighting is not robust and contains artefacts and missing information, resulting in holes and grains in the relighted image. However, the RTI modeling of the corrected image provides a more faithful representation of the relighted surface.

To further demonstrate the effectiveness of our proposed methods, we apply RTI model to the stitched data using the DMD modelling. We compare the relighted images obtained from a model fit using the original stitched images, the images stitched with traditional blending, and the images stitched after correction. An example of relighting the Sutra Container surface from a light direction of ($\theta = 110^{\circ}, \phi = 60^{\circ}$) is shown in Figure 3.13. We use the modified light positions with a larger



(a) Canvas painting

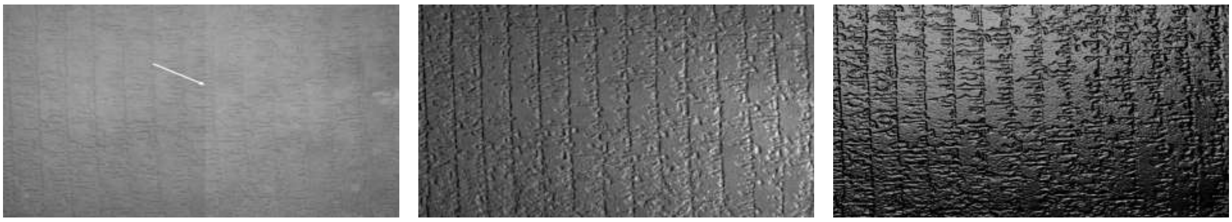
(b) Metal print plate

Figure 3.11: Examples of the corrected images of the real surfaces. Left represent the image corrected using the inverse square law and Lamberts cosine law, and the right represents the image corrected using the BRDF fitting approach.

virtual dome to relight the stitched images.

3.6 Conclusions

In this chapter, we presented two methods for correcting the illumination in photometric stereo data. Both these methods requires normal and albedo estimation in order to model the illumination correction. We used two approaches to obtain the normal map and albedo of the surfaces and presented the performance of the both. For the illumination correction the first method utilizes the inverse square law and Lambert's cosine law to correct the illumination. This approach performs well when applied to virtual acquisition data, but may not handle noise as effectively in real acquisition data, resulting in visible seams after correction. The second method involves fitting a BRDF model to the MLIC data. While this method is more computationally expensive, it effectively corrects the uneven illumination of the surface. However, the appearance of the surface may be affected by the choice of the BRDF model, as the surface is reconstructed entirely by parameterizing the BRDF model. To achieve even more accurate correction of illumination, more sophisticated and accurate reflectance models may be explored in future.



(a) Tomb stone

(b) Sutra container

(c) Canvas painting

(d) Metal print plate

Figure 3.12: Results of stitching images with and without illumination correction. In the left, the original images were stitched without any correction. The middle and right images show the results of stitching images using method 1 and method 2, respectively.

(a) Relighted from the model fitted with the stitched original data

(b) Relighted from the model fitted with the original images stitched and alpha blending applied to the overlapping region.

(c) Relighted from the model fitted with the illumination corrected images stitched.

Figure 3.13: RTI modelling of the stitched data is done using DMD and using the fitted model, the surface is relighted from virtual direction ($\theta = 110^{\cdot}, \phi = 60^{\cdot}$ here)

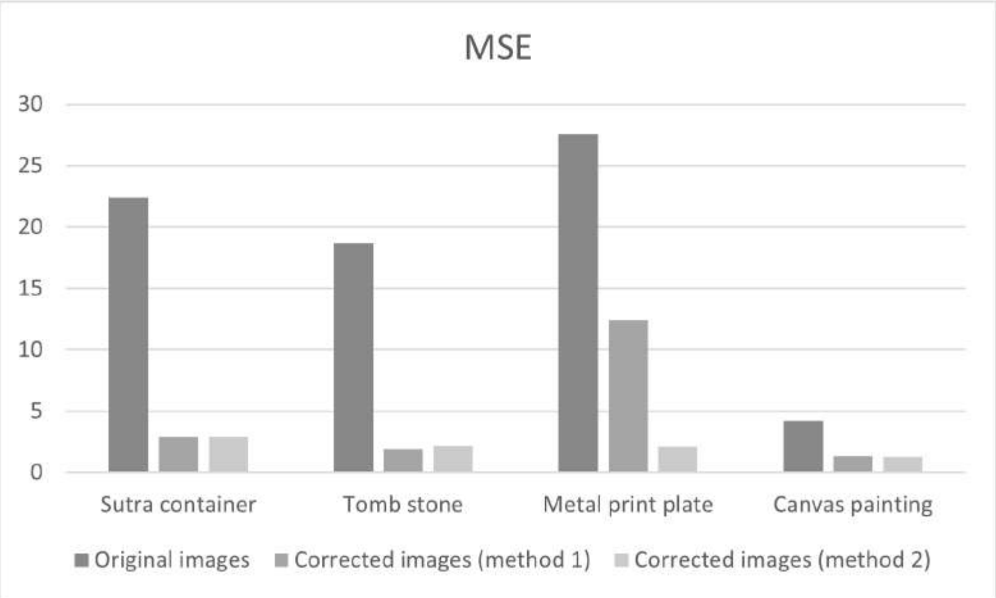


Figure 3.14: Graph showing the mean squared error of the differences between pixels in the overlap region of the stitched images with and without correction.

CHAPTER 4

A benchmark dataset and evaluation for best light positions in Reflectance Transformation Imaging

Chapter overview

There is a lack of appropriate benchmark data-set and reference light configurations for RTI acquisitions. Due to this, quantitative comparison and evaluation of RTI data acquisitions is difficult to achieve. In this chapter, we present a dataset that can be used to evaluate the effectiveness of different surface light configurations for RTI acquisition. We also introduce methods for deriving a good reference light configuration that is adapted to a surface from its dense RTI acquisition. This dataset provides a standardized set of dense RTI acquisitions, along with the corresponding reference light configurations that were derived using our methods. By using this dataset, researchers and developers can more easily compare the performance of their approaches for solving the best light positions problem in RTI acquisition. This dataset can help to improve the accuracy and efficiency of RTI acquisition, and make it more widely applicable in a variety of fields.

4.1 Introduction

RTI has emerged as a simple yet powerful tool for visual analysis and surface characterization particularly in the field of cultural heritage. It has been extensively researched in post-acquisition steps such as modeling and visualization, leading to numerous advancements in these areas. However, the acquisition process itself has not been studied extensively, and there are currently no widely accepted methods available to ensure quality RTI acquisition that is adaptive to the object being captured. There are only limited number of studies carried out in choosing light positions in an RTI acquisition. A general analysis of the RTI acquisition pipeline is presented in [Dellepiane et al., 2006]. The article addresses concerns regarding object size and acquisition conditions, as well as quality assessment considerations. These assessments examines the degradation of PTM quality in relation to the number and position of lights used during acquisition. Pintus et al. [2016] presents an automated light calibration pipeline with exponential fall-off depending on the distance from the light source in a RTI acquisition but does not discuss the selection of light positions.

To investigate the impact of light position selections on the performance and quality of Reflectance Transformation Imaging (RTI), we conducted an evaluation of RTI data quality on various surfaces using different light configurations. As part of this quality evaluation, we use brushed metal with a dent and canvas painting as case studies, and performed very dense RTI acquisitions of these surfaces (1000 homogeneously distributed Light Positions (LPs)) which were considered as the ground truth. We then performed sparse acquisitions with homogeneously distributed light positions incrementing from 40 number of positions to 55, 70, and so on. We then created DMD model for each acquisition.

We analyzed the degradation of RTI quality in these sparse acquisitions through comparison of the normal maps of the dense DMD and each of the subsampled ones. Calculating the dihedral angle between two normals (from dense RTI and sparse RTI) provides a measure of how different the surface orientations are at that point on the object. Comparing two normal maps through dihedral angles provides a more meaningful measure of the difference in surface orientation between the two maps than simply comparing pixel values. The dihedral angle between two normals represents the angle between the two planes that are perpendicular to those normals, providing a measure of how different the surface orientations are at that point on the object. For each pixel in the normal maps obtained from parse RTI and dense RTI, we extract the normal vectors. The dihedral angle θ between the two normal vectors can be calculated using the dot product:

$$\theta = \arccos\left(\frac{n_{dense} \cdot n_{sparse}}{(\|n_{dense}\|\|n_{sparse}\|)}\right), \tag{4.1}$$

where θ is the dihedral angle, n_{dense} and n_{sparse} are the normal vectors of the pixel in the normal maps obtained by RTI of the dense acquisition data and sparse acquisition data respectively.

Another analysis was done by comparing the relighted images from the sparse acquisition with that from the dense acquisition. The mean squared error (mse) of each pixel between the dense relighted image and the sparse relighted images was used as a metric to make this comparison. Figures 4.1 and 4.2 illustrate these comparisons made on the brushed metal with a dent surface and a canvas painting surface, respectively. Figures 4.3 and 4.4 represent the plots of the two analysis for the respective surfaces.

The analysis conducted on the surfaces using different light configurations clearly indicates that the quality of RTI acquisition is significantly influenced by the choice of light positions, and this impact varies for each surface. Therefore, using a common set of light positions for every surface does not guarantee quality RTI of the surfaces. With regards to the number of lights, the analysis shows that it is possible to reduce the number of lights without experiencing excessive degradation of quality, compared to a very dense acquisition. For instance, for the brushed metal surface, using 250 homogeneously distributed light positions can achieve RTI quality close to that achieved with 1000 light positions in a dense acquisition. Similarly, for the canvas painting surface, using more than 55 images is sufficient to achieve quality comparable to that obtained from a dense acquisition.

Reducing the acquisition size can be advantageous in various aspects, such as stitching RTI data of large surfaces, performing High Dynamic Range RTI [Nurit et al., 2019], focus variation RTI [Lewis et al., 2021], multi-spectral RTI [Giachetti et al., 2017], etc. This can help to improve efficiency by lowering computational costs while still achieving desirable results.

The best light positions for RTI will vary upon the specific object being imaged and the desired outcome. The light sources should be positioned to create a range of highlights and shadows on the object, allowing for the capture of a wide range of surface details. Experimenting with different light positions and adjusting the settings on the RTI system can help to produce the best possible results. Uniformly distributing the light positions in dome-based systems or manually positioning

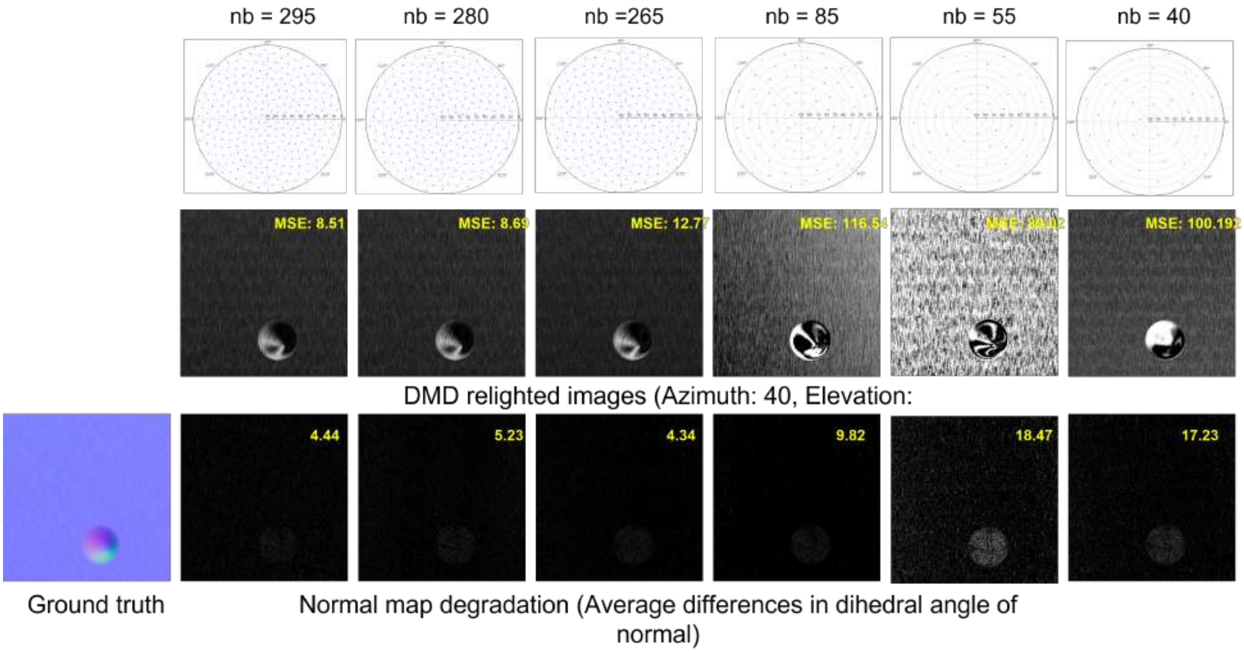


Figure 4.1: Degradation in the quality of RTI of the brushed metal with dent surface degrades with different light configurations. The top row displays the positions of lights projected on a 2D plane. The middle row shows maps of the differences between the dihedral angles of ground truth normals and the normals obtained from corresponding sparse RTI acquisitions. The bottom row displays the reconstructed images from the respective sparse acquisitions (relighted from an elevation of 45˙ and an azimuth of 40˙)

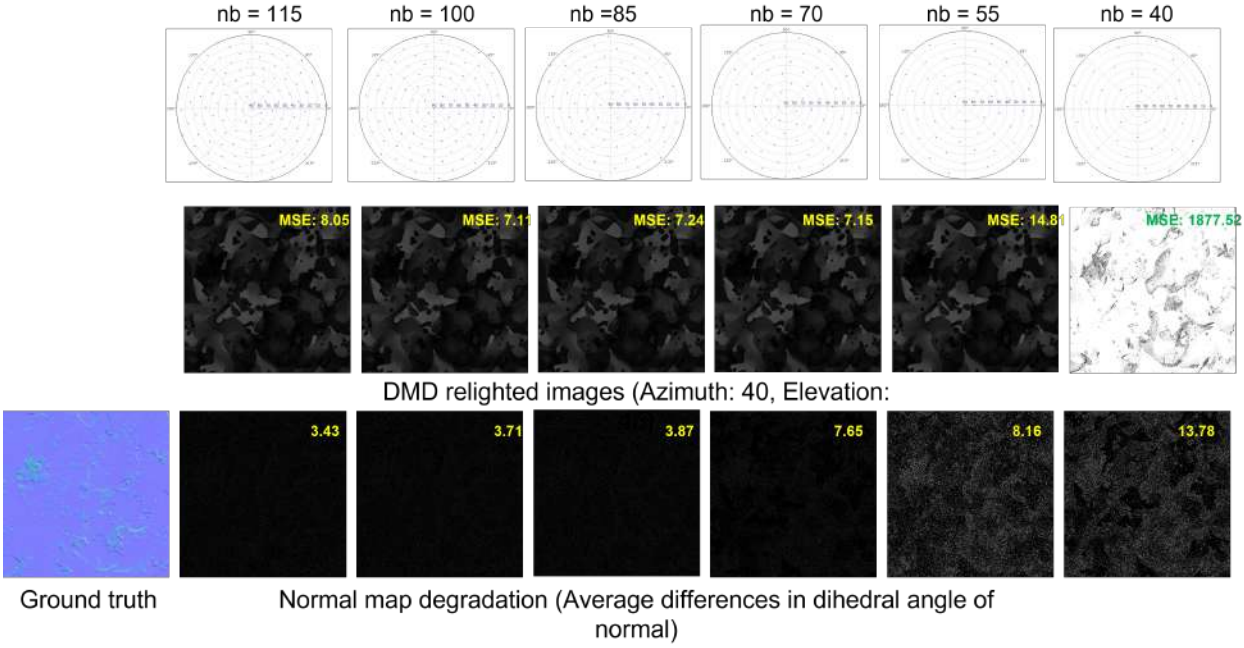
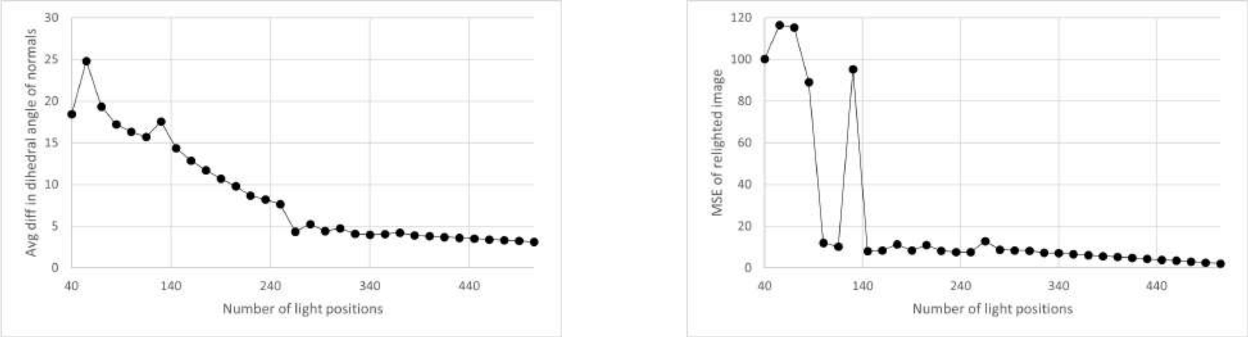
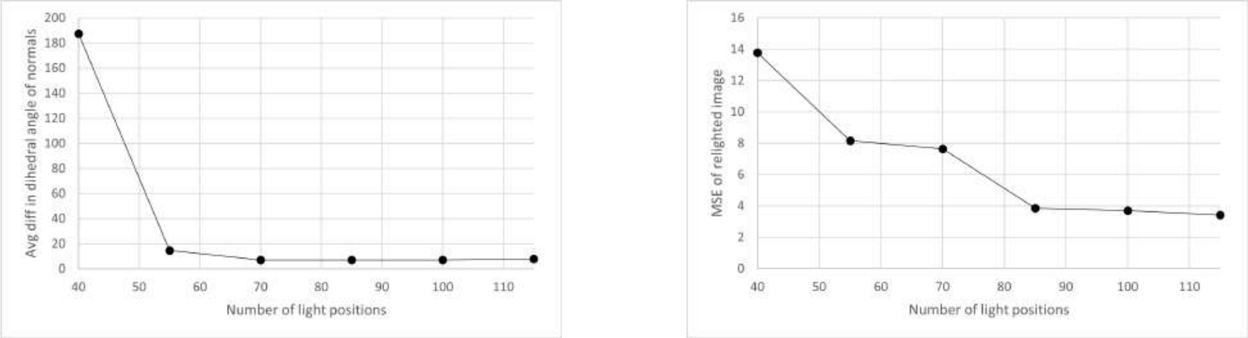


Figure 4.2: Degradation in the quality of RTI of the canvas painting surface with different light configurations. The top row displays the positions of lights projected on a 2D plane. The middle row shows maps of the differences between the dihedral angles of ground truth normals and the normals obtained from corresponding sparse RTI acquisitions. The bottom row displays the reconstructed images from the respective sparse acquisitions (relighted from an elevation of 45˙ and an azimuth of 40˙)



(a) (b)

Figure 4.3: RTI quality of the brushed metal with dent surface with the number of homogeneously distributed light positions (a) Plot of mean difference of the dihedral angle between the ground truth normals and the sparse acquisition normals with the number of homogeneously distributed light positions. (b) Plot of MSE of the relighted images with the number of homogeneously distributed light positions.



(a) (b)

Figure 4.4: RTI quality of the canvas painting with the number of homogeneously distributed light positions. (a) Plot of mean difference of the dihedral angle between the ground truth normals and the sparse acquisition normals with the number of homogeneously distributed light positions. (b) Plot of MSE of the relighted images with the number of homogeneously distributed light positions

light sources in free-form settings do not guarantee the quality of an acquisition, leading to problems such as skipping pertinent light positions and capturing redundant information.

Figure. 4.5 illustrates the importance of choosing best light positions in RTI acquisition. In the figure, a surface point signal measured from very dense acquisition (1000 LPs), acquisition containing best light positions estimated for the surface (34 LPs), and evenly spaced acquisition (28 LPs) are plotted as continuous curves using spline fitting. For illustration, we only show one acquisition ring (fixed elevation, variable azimuth in the hemisphere). It can be observed that the acquisition performed using a fixed set of light positions risk losing the relevant light directions. This problem is common on semi-gloss, anisotropic surfaces. It can be observed that conventional acquisition using a fixed set of light positions completely misses the specular orientation and thus loses key information. One could argue that making dense acquisition every time will guarantee capturing of all information. However when the acquisition is made with estimated best light positions, that minimizes information loss without severely impacting the size of acquisition. Acquiring numerous high-resolution images and processing them to obtain different feature maps presents multiple challenges due to high computational cost. This is particularly true when developing new methods like HDR [Nurit et al., 2021], Focus Variation Lewis et al. [2021] in RTI, where handling large data is difficult before the method is optimized.

The two major goals in finding best light positions in an RTI acquisition adaptive to the surface

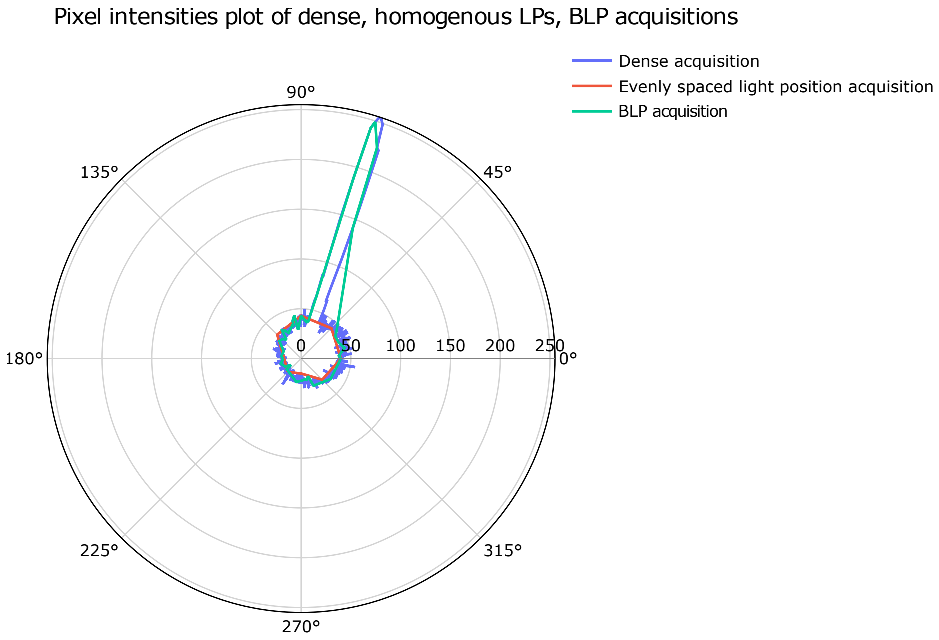


Figure 4.5: Comparison of specular lobes obtained from regular uniformly distributed light position acquisition, dense acquisition, and acquisition with optimal light positions. For better illustration, this representation only considers one ring of the acquisition.

are:

- it must ensure the highest quality of data (adaptive to the object) with an optimal number of captures in an acquisition,
- determine the termination of the acquisition through a stop-criterion.

Determining the optimal light positions and estimating when an RTI acquisition should be terminated are challenging problems that have not been extensively studied. To date, there are no established metrics for quantitatively evaluating the quality of a light configuration in RTI. This remains an open area of research, and further work is needed to develop effective techniques for adapting light positions and determining when an RTI acquisition is complete.

4.2 RTI dataset

The RTI dataset was created to evaluate methods for estimating the optimal light configuration for a surface under standard RTI imaging set-up, where the camera is fixed orthogonally to the surface. This dataset focuses on anisotropic, non-Lambertian, semi-glossy, non-Lambertian diffuse surfaces. There are some publicly available datasets for RTI and RTI-like techniques Aanæs et al. [2012], Jensen et al. [2014], Shi et al. [2016]. But none of them are suitable for evaluating methods for detecting good lighting directions. We therefore created a dataset of RTI acquisitions of virtual and real surfaces, accompanied by very dense RTI acquisitions and ground truth 3D data. For real surfaces, RTI acquisitions were made using a carefully calibrated mechanized RTI dome setup built in-house, and 3D shapes were acquired using a structured light scanner. Virtual surfaces were created using Blender and to perform RTI acquisitions, we built a RTI plugin that can be directly used in the Blender software.

This dataset contains 20 surfaces, as shown in Figure.4.6. Surfaces 1 to 13 are virtual surfaces created in Blender software using physically based rendering methods Pharr et al. [2016], and surfaces 14 to 20 are real objects. The physical dimensions of the surfaces range from 2 cm to 4 cm wide and 3 mm thick. With regards to BRDF, the dataset covers anisotropic surfaces (brushed metals, 1 to 7), isotropic metallic surfaces (8,9), Lambertian surfaces with strong details and nearly homogenous specular lobes (11,12,13) and non-Lambertian diffuse surfaces with strong specular spikes (14 to 20). This is summarized in Table. 4.1.

RTI dataset surfaces

Virtual surfaces			
Brushed metal	Linear brushed	Simple[1]	Anisotropic
		With crack defect[2]	
		With dent defect[3]	
		With scar defect[4]	
		With random defects[5]	
		With shear defect[6]	
	Circular brushed	Simple[7]	
Rust metal	Rusted specular white shinny metal[8]		Isotropic metallic
	Diffused green[9]		
Concrete wall[10]			Lambertian with strong details
Wood	Planck wood[11]		
	Old painted wood[12]		
Canvas painting[13]			
Real surfaces			
Ancient coins	Coin1 front[14]		Non-lambertian diffuse
	Coin1 back[15]		
	Coin2 front[16]		
	Coin2 back[17]		
	Coin3 front[18]		
	Coin3 back[19]		
Zinc print relief plate[20]			

Table 4.1: Surfaces in the RTI dataset and the range of BRDF covered.

Imaging set-up: Acquisitions of the real surfaces were performed using Allied Vision model Manta-G-1236B monochrome camera. The resolution of the images captured are 3008 x 4112 pixels. The RTI dome is black so there are no internal reflections. The light source is mounted on a motorized stage that allows the light source to be positioned at any desired position in angular space between 0 to 360 degrees in azimuth (θ) and 0 to 75 degrees in elevation (ϕ). The radius of the hemisphere in our set up is 32 cm. The light source is a high power white LED, brand - SmartVision Light, model - ODSX30. The light source positioning device is well calibrated using reflective spheres. The dataset follows the standard RTI acquisition format and includes images in portable networks graphic (png) format with light positions saved in .LP files. For the 3D shapes, we used Shining 3D's Einscan INSPEC 3D scanner. It is a structured light 3D scanner with an accuracy of $\leq 10\mu m$. The 3D shapes are available in commonly used formats.

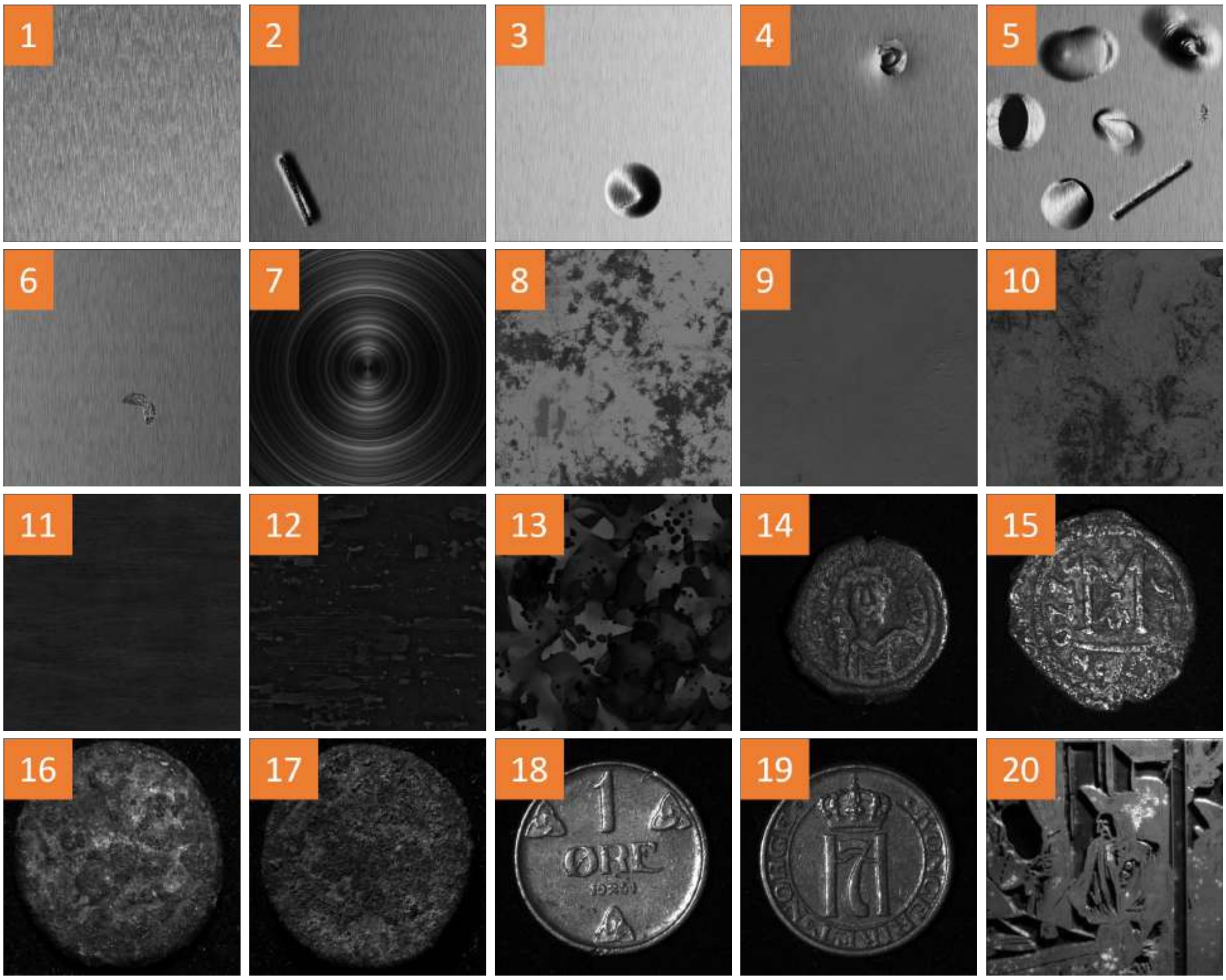


Figure 4.6: RTI dataset surfaces.

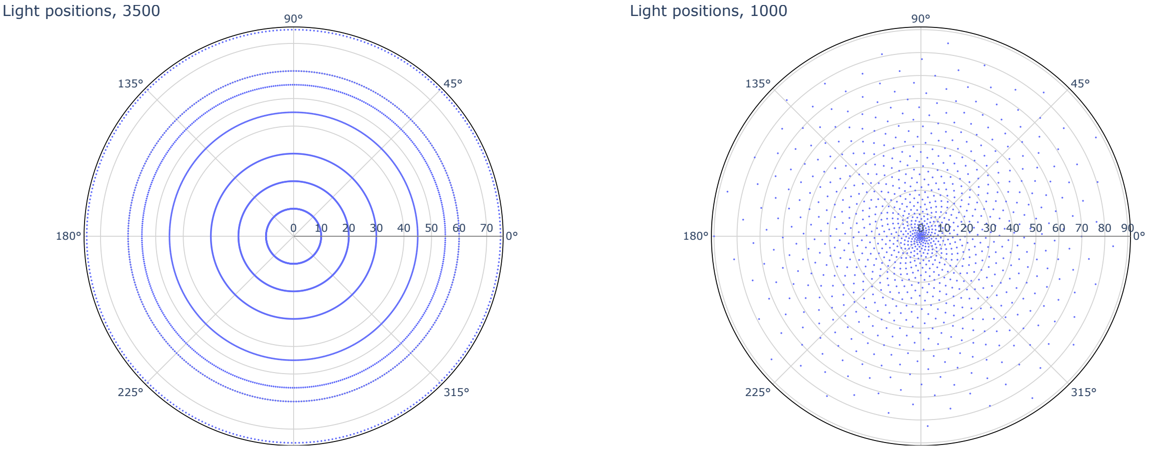
Dense RTI acquisitions: In our approach, we derive reference good light directions for a sparse RTI acquisition of a surface from a very dense acquisition of the same surface. This may seem like an indirect approach, but it is justified because there are no ground truths for RTI acquisitions. We evaluate the performance of a light selection method by comparing it with the reference sparse RTI acquisition. In other words, we compare the results of a sparse with a known "good" sparse acquisition result, in order to assess the accuracy and reliability of the planning method.

Dense ring acquisitions of surfaces were carried out to focus on analyzing the behavior of surfaces when illumination changes in azimuthal space. The importance of azimuth space versus elevation space will be discussed in a later section. For each surface, the dataset includes ring acquisitions, as shown in Figure. 4.7a (projected on a 2D plane) and Table 4.2. For higher elevations, the acquisition size is relatively small because the circumference of the ring decreases with increasing elevation.

Dense acquisitions of surfaces in azimuth-elevation space is made for analyzing the behavior of

Elevation (ϕ)	10°	20°	30°	45°	55°	60°	75°
Number of points	600	600	550	550	450	400	350
Avg spacing between adjacent points	0.328 cm	0.317 cm	0.317 cm	0.259 cm	0.258 cm	0.251 cm	0.1487 cm

Table 4.2: Lighting configurations for dense ring acquisitions carried out for each surfaces. The number of points are evenly spaced between 0° to 360° azimuths.



(a) Dense rings - each ring contains LPs evenly spaced in azimuth space.

(b) Dense light positions evenly spaced in azimuth-elevaiton space.

Figure 4.7: Distribution of light positions in.

surfaces when illumination changes in both azimuthal and elevation space. For each surface, the dataset includes dense acquisitions in uv space (this space is typically used to map three-dimensional points on to a a two-dimensional image. We captured the surface with changing illumination direction to 1000 distinct points uniformly distributed over an imaginary hemisphere. The light positions are shown in Figure. 4.7. The homogeneous distribution of points over an encompassing hemisphere is dome by mapping the Fibonacci lattice (golden Spiral) onto the surface of the sphere Marques et al. [2019].

4.3 Creating a reference for a good acquisition by using information from a dense acquisition

For a given object, it is interesting to determine the optimal light positions that result in high-quality reflectance transformation imaging. To achieve this, we developed methods that utilize information from dense acquisition to derive the optimal light positions for our dataset. Figure 4.8 and table 4.3 represents the notations used in this section.

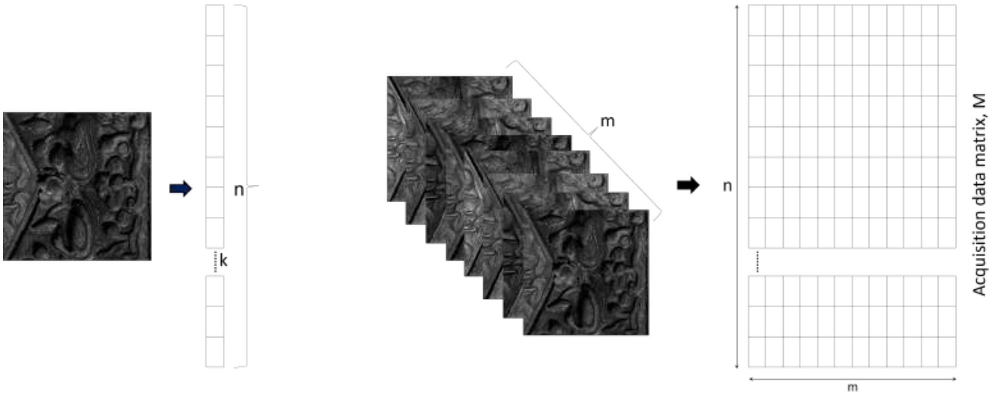


Figure 4.8: Stack of images captured in an acquisition. Here, m and n represents the height and width of an image, i and j represent the row and column index of the imaage, l represents the size of an acquisition (number of images), k represents the index of an image in the acquisition. Note: In ring acquisition, the captured images are ordered (0 to 360), but, in azimuth-elevation space there is no order.

Here, we have a stack of m images captures in an acquisition. From this stack, we form a matrix

Notation	Description
m	Size of an acquisition - Number of images in an acquisition
n	Size of each image (vectorized)
k	Index of a surface pixel
S	Set of all points (directions) in the dense acquisition
s_i	Individual point (direction) in the dense acquisition
$d(s_i, s_j)$	Distance between points s_i and s_j
$g(s_i)$	Gradient of the signal at point s_i
$v(s_i)$	Value of the signal at point s_i
T_d	Threshold distance for sampling
T_g	Threshold gradient for sampling
T_v	Threshold difference in signal value for sampling
S'	Set of sampled points

Table 4.3: Notation table.

with m columns and n rows, where n is the size of image when it is vectorized by flattening it.

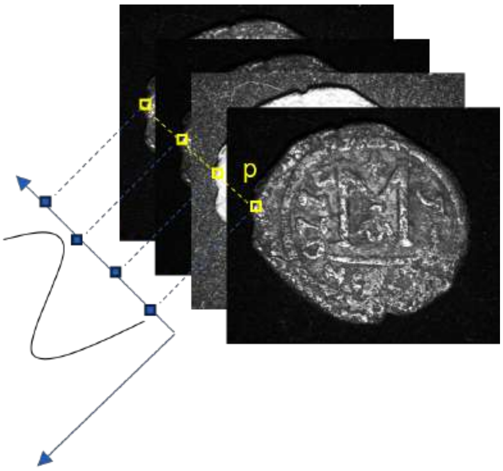


Figure 4.9: A pixel p and its corresponding signal across the stack of images

Let's say we are interested in a specific pixel p as illustrated in Figure.4.9. The signal v corresponding to this particular signal is expressed mathematically as:

$$v = [M_{p1}, M_{p2}, ..., M_{pm}]^T \tag{4.2}$$

where, M is the acquisition data matrix.

The gradients of this signal is then expressed as:

$$g = [(v(s_2) - v(s_1)), (v(s_3) - v(s_2)), ..., (v(s_m) - v(s_{m-1}))] \tag{4.3}$$

here the points $S = s_1, s_2, ..., s_m$ are in regular grid and sorted.

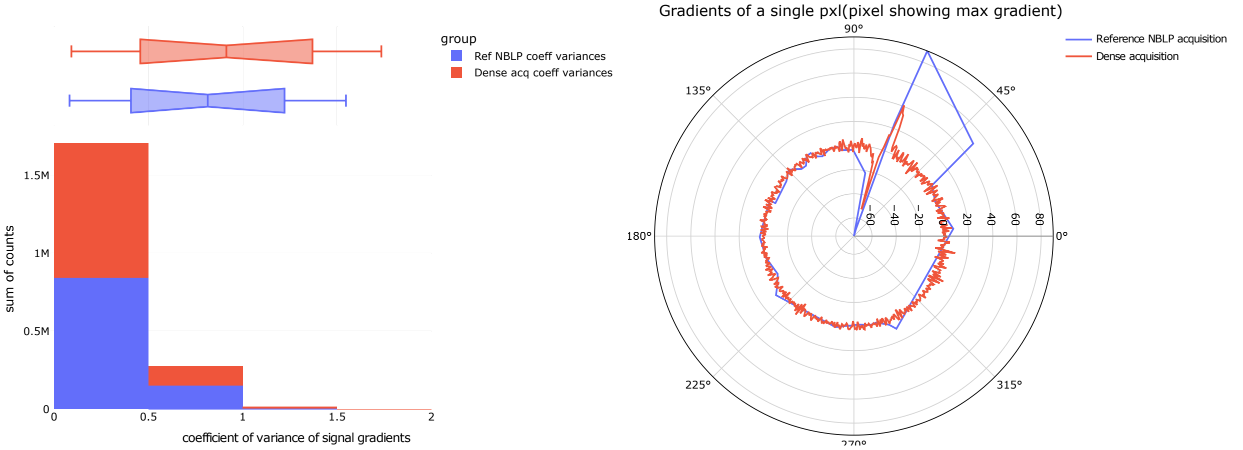
4.3.1 Aim

We propose to derive reference good sparse acquisition by analyzing very dense acquisition that is guaranteed to capture all information. The aim is to ensure that reference acquisitions have the following characteristics:

1. The estimated reference good light directions must be a unique set of light positions and must have no or least number of redundant measurements in it. The acquisition must capture the

luminance behavior of all points on the surface as close as possible to a dense acquisition, but with the fewest number of light directions.

2.The gradients of the signals in a reference acquisition should be as uniform as possible, that is, the light position is densely distributed in the direction where the luminance of a surface point changes rapidly, and sparsely distributed in the direction where the luminance change is small. This is shown in Figure. 4.10a and Figure. 4.10b . This example corresponds to a ring acquisition for an ancient coin. The first plot represents the histogram of gradients of *reflectance signals* of all the surface points in the reference acquisition and that of the dense acquisition. The second plot is a radial polar plot of the gradients of the reflectance signal of a single surface point for the reference acquisition versus the dense acquisition (in this example, we have shown a surface point that exhibits the highest gradient of the reflectance value among the entire surface). It can be observed that the reference acquisition gradient tends to be uniform.



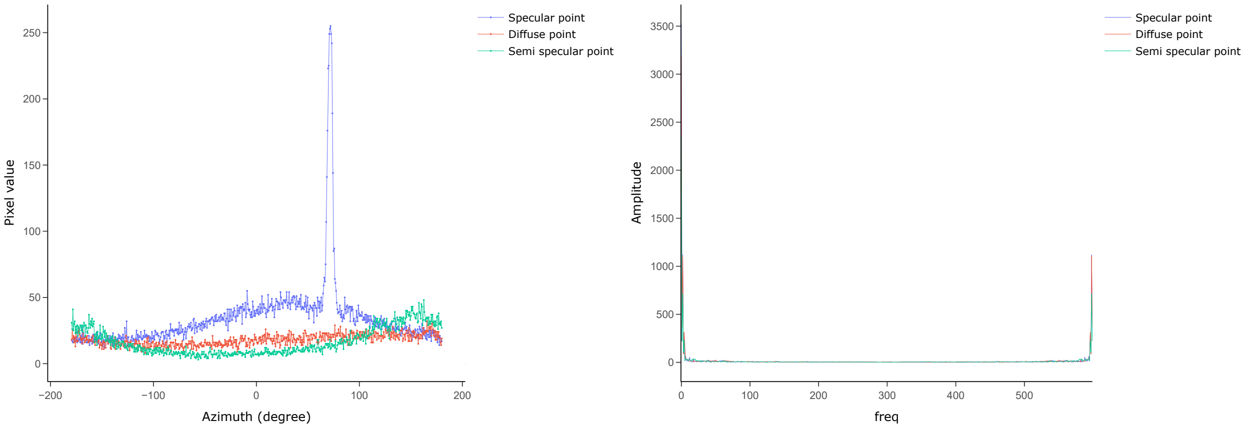
(a) Distribution of pixel gradient values of a azimuth space acquisition.

(b) Gradient of the pixel having the maximum gradient value in the azimuth space.

Figure 4.10: Gradients pertaining to a azimuth only space RTI acquisition.

RTI acquisitions typically involve measuring the luminance behavior of surface points at high resolution, which can result in the acquisition of millions of points. Traditional signal sampling methods cannot be directly applied to this high-dimensional problem. Instead, we must find a way to sample the reflectance signals from the surface points in a way that captures their non-linear behavior while also being efficient. Simultaneously sampling such a large number of reflectance signals is a challenging problem. Depending on the material, albedo, and specularity, large differences in reflectance behavior at various points on the surface can be observed. Figure. 4.11a illustrates such changes observed in a ring RTI acquisition, where we show measurements for 3 surface points. It can be observed that the reflected signal is strongly nonlinear in the azimuth space. Fig. 4.11b is the fourier transform of the three signals.

First, the presence of peaks in the frequency range of 5-10 suggests that there are periodic components in the signal that repeat at a frequency of 5-10 cycles per sample. This implies that the signal was sampled at a rate that is high enough to accurately capture these periodic components. Nyquist sampling theorem states that to avoid aliasing (the phenomenon where high-frequency components in a signal are misinterpreted as low-frequency components), a signal must be sampled at a rate that is at least twice the highest frequency present in the signal. The presence of peaks in the frequency > 590 indicates that the the signal is not sampled at a high enough rate even in the dense acquisition. Therefore FFT based signal sampling methods cannot be directly applied in our case.



(a) Signal.

(b) FFT of the signal.

Figure 4.11: Examples of reflectance signals measured of 3 different points belonging to the same surface.

4.3.2 Proposed methods

Azimuth only space

For characterizing the reflectance of a surface, the azimuth space is often more important than the elevation space. Many analysis techniques use ring acquisitions to create maps based on directional slopes and curvature. On the other hand the reflectance behavior of the surface in the elevation space can be interpolated using a second-order polynomial with just a few sample measurements.

To efficiently capture the non-linear nature of reflectance in the azimuthal space, we introduce a method for with best azimuth angles for a surface from its dense RTI acquisitions. The goal of the sampling strategy is to reduce a dense acquisition of surface reflectance signals to a sparse representation without losing much information. The dense acquisition comprises surface points illuminated from different directions in the azimuth space from 0˙ to360˙.

Firstly the surface points are classified as diffuse, semi-specular, or specular based on the maximum gradient observed in its reflectance signals. If the maximum gradient of any signal is greater than 10 (the average minimum difference in pixel intensity that the human eye can detect is 10), the surface is considered to have some specular or semi-specular points. Otherwise, the surface is considered diffuse.

To cluster the signals, we use the K-means temporal signal clustering technique proposed by Grønlund et al. [2017], as shown in Figure. 4.12. For diffuse reflectance signals, we simply take 8 to 10 evenly spaced light positions as suitable for the RTI acquisition. For specular and semi-specular signals, we apply an approach as described below to identify their pertinent light directions.

To achieve this, we traverse from 0˙ to 360˙ in the dense acquisition and sample a point if it satisfies any of the following three criteria:

- The current sampling point is spaced higher than a threshold distance (arc length) to the previous sampled point. This criterion ensures that we sample points at regular intervals, rather than sampling points too close together.
- The gradient of the signal at a point is higher than a threshold. This criterion ensures that we sample points where there is a significant change in the signal, as these points are likely to be more informative than points with little or no change.
- The difference between the current signal value and the previously sampled signal value is higher than the threshold. This criterion ensures that we sample points where there is a significant change in the signal, even if the gradient at that point is not very high.

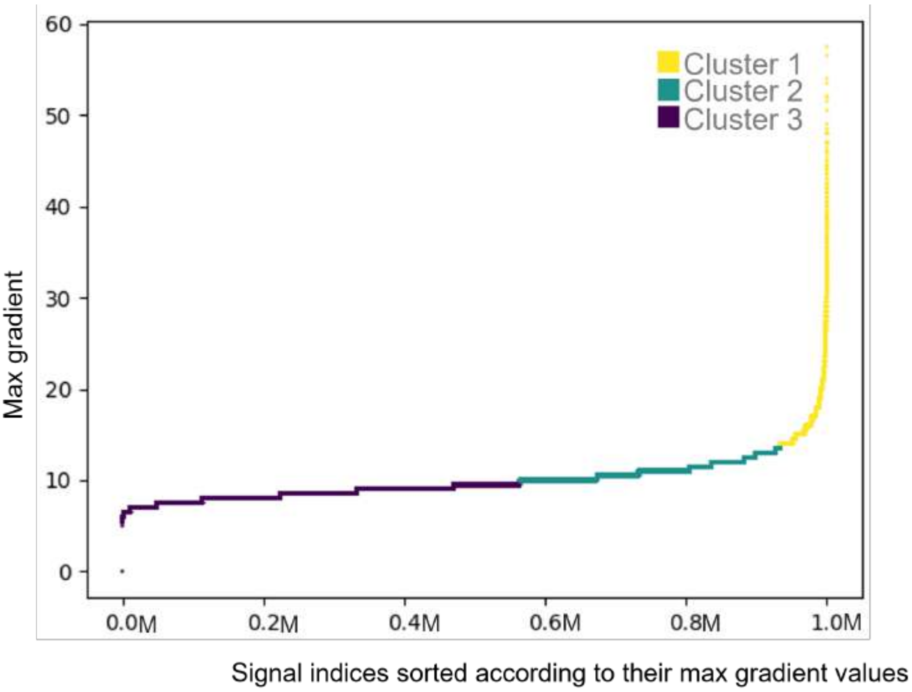


Figure 4.12: Using the K-means temporal signal clustering technique, the signals are grouped into specular, semi-specular, and diffuse based on their maximum gradient value. This example corresponds to the measurement of surface 14 by a ring acquisition at an elevation of 30°.

For calculating the threshold, we use two different ways depending on the nature of reflectance. If the reflectance is specular in nature, we use half the highest gradient observed in the dense reflectance signal. If the reflectance is diffuse, we use twice the highest gradient value observed in the dense reflectance signal. This is done to maintain uniform gradient in the decimated reflectance signal. This strategy is illustrated in Figure 4.13

Mathematically, we can represent the decimation process as follows:

Let S be the set of all points in the dense acquisition, and $s_i \in S$ be an individual point. Let $d(s_i, s_j)$ be the distance between points s_i and s_j, $g(s_i)$ be the gradient of the signal at point s_i. Let T_d be the threshold distance for sampling, T_v be the threshold gradient or difference in signal value for sampling.

The set of sampled points S' can be defined as:

$$S' = \{s_i \in S \mid d(s_i, s_{i-1}) > T_d \lor g(s_i) > T_v \lor |v(s_i) - v(s_{i-1})| > T_v\}, \tag{4.4}$$

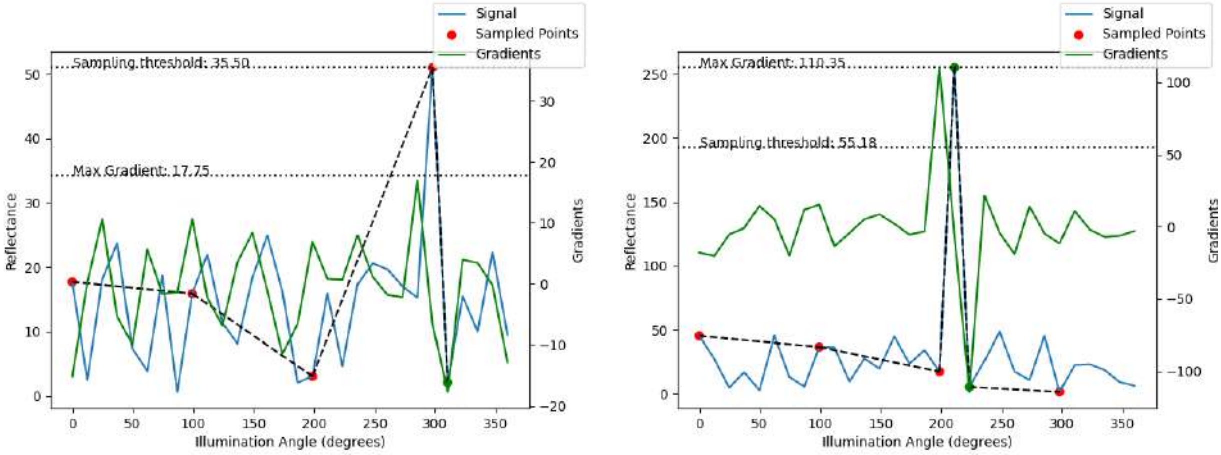
where $T_v = \frac{1}{2} \max_{s_i \in S} g(s_i)$ if the reflectance is specular and $T_g = 2 \times \max_{s_i \in S} g(s_i)$ if the reflectance is semi-specular. Note that the gradient criteria is never satisfied for the semi specular points.

It is computationally challenging to sample all the signals (surface points) simultaneously. To make the process efficient, we First gather all the dense acquisition of surface reflectance signals and put them into a single matrix called the acquisition data matrix M.

To perform the proposed sampling method efficiently, we use matrix column operations instead of looping through each signal and acquisition point individually. This allows us to apply the sampling method to all signals simultaneously, which is much faster and more efficient than processing each signal separately.

Azimuth-Elevation space

Adding the elevation dimension to the problem greatly increases its complexity. While gradient-based sampling is an effective technique for analyzing 1D signal decimation, it becomes much more



(a) Semi specular reflectance signal sampling. (b) Specular reflectance signal sampling.

Figure 4.13: Examples of reflectance signals measured of two different points belonging to the same surface. For the semi specular point (a), the reflectance signal has low gradient magnitudes, hence it is sufficient to sample points at locations of significant gradient changes compared to the previously sampled point (twice the max of the original signal gradients). For the specular point (b), since the gradient variation is huge, we choose a threshold sampling value half the maximum gradient observed in the original signal.

challenging when applied to azimuth-elevation data. This is because the non-continuous dense signal contains data that are not arranged in a regular grid. For performing gradient descent, computing partial derivatives along the x and y directions by regularizing the grid using any surface fitting methods for millions of signals would be computationally prohibitive.

Decimating multiple signals simultaneously is challenging to preserve the characteristics of each signal while reducing their complexity. Our approach to solving this problem is to represent each signal as a 3D pointcloud and decimate these pointclouds. This approach has the advantage of allowing for more control over the decimation process, and it can result in more accurate and faithful representations of the original signals.

Similar to our approach in the azimuth-only space, the surface points are first classified as diffuse, and specular. However, in azimuth-elevation space, computing the signal gradients by comparing each signal point to its neighbours is computationally very expensive. Instead we identify the specular and semi specular signals by simply comparing the minimum and maximum values observed in a signal. If the difference between the max and min is lower than 20, we classify the signal as diffuse, if it is higher than 20 we classify it semi specular. This is an approximate classification a for the preliminary step of the decimation process. To optimize the light positions, we focus on specular surface points. These light positions optimized for the specular surface points are also expected to work well for diffuse points, as the reflectance of diffuse points is uniform.

In our approach we create a pointcloud for each signal. To create a pointcloud for a signal, we create points in 3D space by combining the x and y coordinates of the light positions with the normalized signal values as z coordinates. We then decimate the pointclouds using the Quadric Error Metrics (QEM) decimation approach Garland and Heckbert [1997], Yu et al. [2014]. This method works by iteratively removing points from the point cloud, starting with the point that has the smallest quadric error when removed till we reach a desired size of the pointcloud. The desired size of the pointcloud corresponds to the desired number of light positions for the acquisition and it can be set to any value depending on the requirement. For PTM, DMD, any number of points $>$ 50 is generally considered reasonable size of an acquisition.

To compute the quadric error for a given point in the point cloud, a quadric matrix is first constructed to represent the local geometry of the points in the neighborhood of that point. The quadric matrix is constructed using a least squares fit to the points in the neighborhood of the point being considered for removal. This fit is used to estimate the local geometry of the points, and the resulting quadric matrix is used to evaluate the error introduced by removing the point.

The quadric error for a point is given by the following equation:

$$E = (x - \hat{x})^T Q (x - \hat{x}), \tag{4.5}$$

where x is the original position of the point, $\hat{x}$ is the position of the point after it has been removed, and Q is the quadric matrix as shown in Figure 4.14.

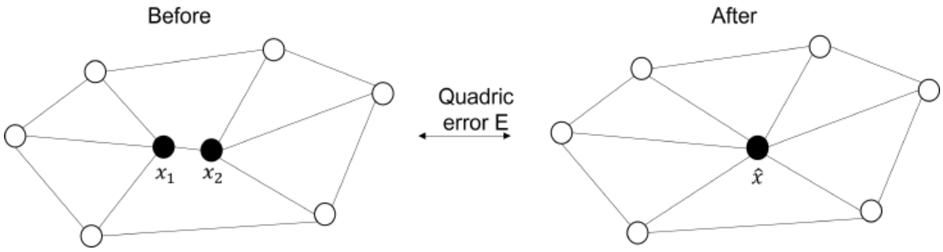


Figure 4.14: Quadric error of a point being removed.

The quadric matrix is given by the following equation:

$$Q = \sum_{i=1}^{n} (x_i - \bar{x})(x_i - \bar{x})^T, \tag{4.6}$$

where x_i are the positions of the points in the neighborhood of the point being considered for removal, and $\bar{x}$ is the centroid of these points.

The QEM decimation algorithm proceeds by iteratively removing the point with the smallest quadric error from the point cloud until the desired number of points is reached. This ensures that the points removed are the ones that introduce the least amount of error, and thus the shape of the point cloud is preserved as much as possible.

Calculating the quadric error for each point in a signal can be computationally expensive, especially when dealing with a large number of such signals. The quadric error minimization (QEM) decimation algorithm has a time complexity of $O(n^2)$, where n is the number of points in the point cloud. This is because the quadric error for each point must be computed individually.

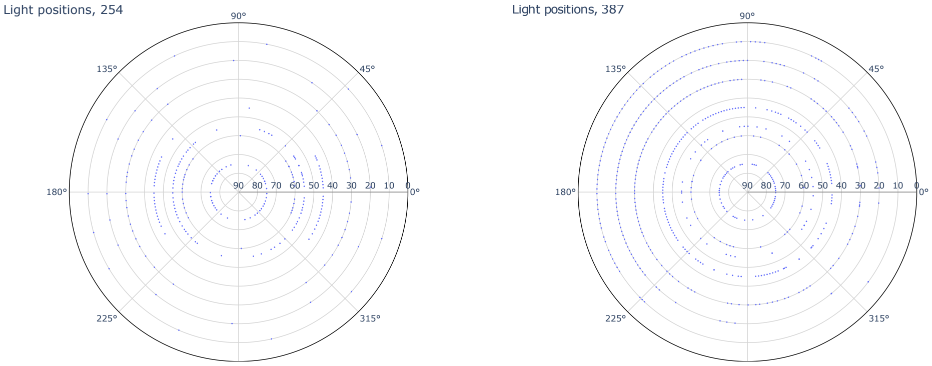
However, we can use a matrix kernel to compute the quadric error for all points simultaneously, reducing the time complexity to $O(n)$. To do this, we first compute the quadric matrix for each point using the points in its neighborhood. Then, we subtract the centroid of each neighborhood from each point in that neighborhood to obtain a set of centered points.

Finally, we compute the quadric error for all points simultaneously by taking the dot product of the centered points with the quadric matrix and the transpose of the centered points. This operation is available in standard programming libraries.

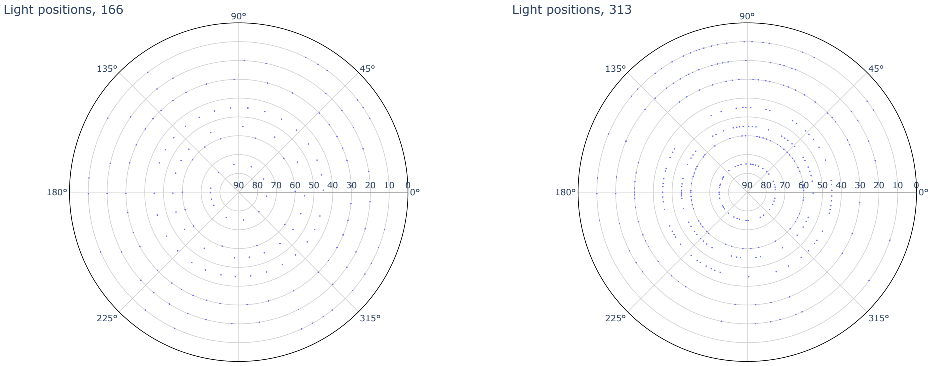
Using a matrix kernel to compute the quadric error allows us to significantly improve the performance of the QEM decimation algorithm. This makes it possible to decimate large number of sigmals more efficiently, and thus can be applied to derive reference best light positions in the azimuth-elevation space.

4.4 Results

Results for selected surfaces representing the general behavior of each surface category are presented here. More results for are provided in the appendix.



(a) Surface 1 (Simple brushed metal) (b) Surface 3 (Brushed metal with a dent)



(c) Surface 13 (Canvas painting) (d) Surface 14 (Ancient coin1)

Figure 4.15: The reference good points on each ring are shown together, showing the overall distribution of surface light positions.

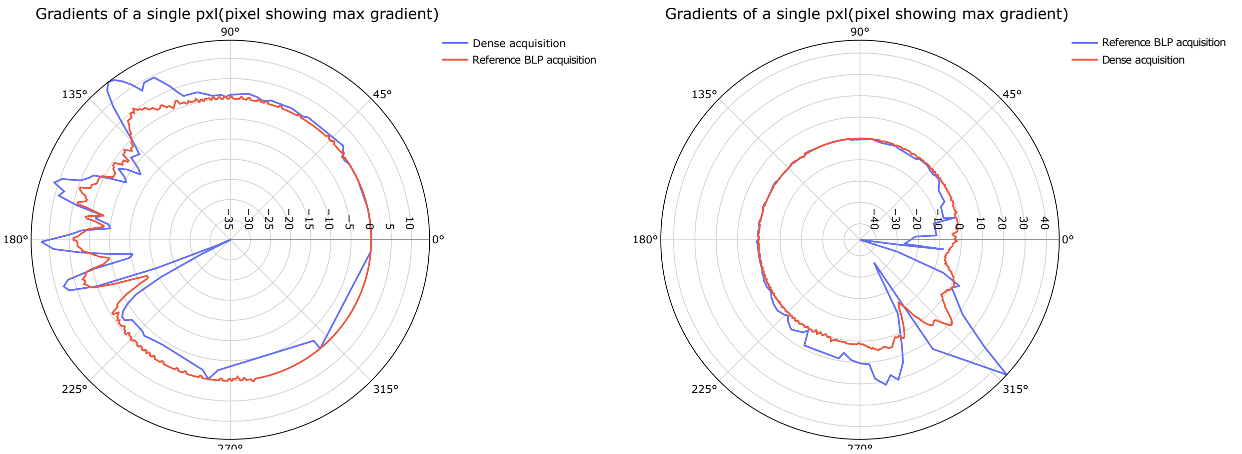
4.4.1 Azimuth only space

Light positions: Fig 4.15 depicts the reference good light positions obtained using the proposed method 4.3.2 for the brushed metal plate (Surface 1), the brushed metal plate with a dent defect (Surface 2), the canvas painting (Surface 13), and the antique coin 1 (surface 14). Surfaces 1, 2, 13 are virtual surfaces and surface 14 is a real surface. It can be observed that the proposed method is able to adaptively derive the reference good light positions to the surface. Since surface 13 is a perfectly diffuse surface, the number of reference best light positions is small and nearly evenly distributed along the rings. For anisotropic surface 1, the method correctly identifies the direction exhibiting anisotropy and samples more points along that direction than other directions to preserve gradient uniformity. Surface 3, Surface 14 contain details showing specular reflections of many random surface points, so the distribution of reference light positions is numerically larger.

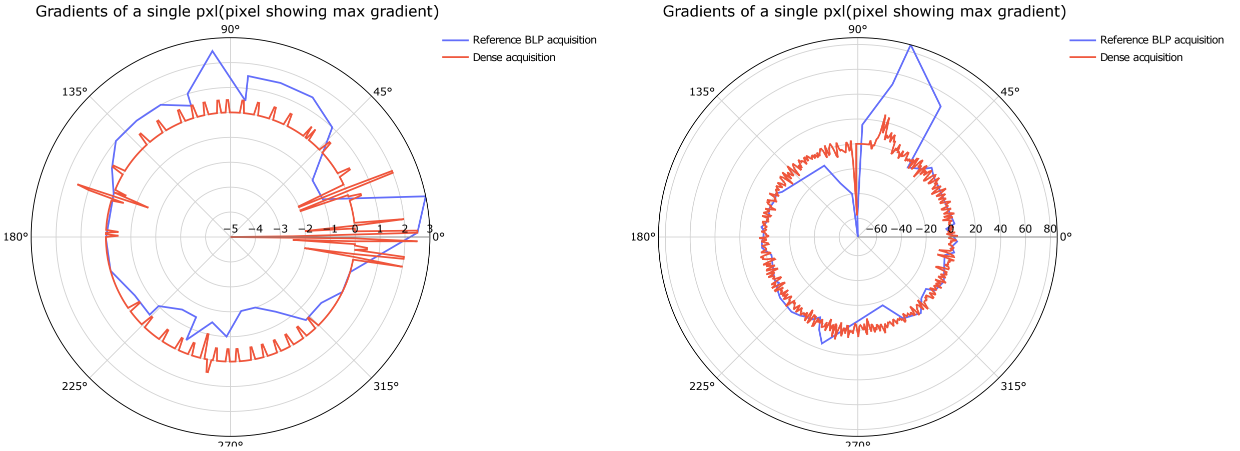
Reflectance signal gradients: The proposed method 4.3.2 can achieve the goal of maintaining the gradient uniformity of the reflected signal. Figure. 4.16 compares the gradient at a single surface point (here we have chosen the pixel showing the largest gradient) in the reference best light positions acquisition versus the dense acquisition.

To illustrate the overall dispersion of signal gradients considering all surface points, we use the coefficient of variation as a measure of dispersion. For example, a dense ring acquisition contains 1000 images (orientations), each containing measurements of 1 million surface points. We compute the gradient of the surface points individually along the 1000 measurements and compute the coefficient of variations for each surface point from the computed gradients.

Figure. 4.17, compares the distribution of these coefficient of variations in reference good light positions acquisition vs dense acquisition. It can be seen that the reference good light positions acquisition shows smaller level of gradient spread compared to the dense acquisitions.



(a) Surface 1 (Simple brushed metal) (b) Surface 3 (Brushed metal with a dent)



(c) Surface 13 (Canvas painting) (d) Surface 14 (Ancient coin 1)

Figure 4.16: Comparison of reflectance signal gradients from dense and reference best light positions acquisitions of surfaces.

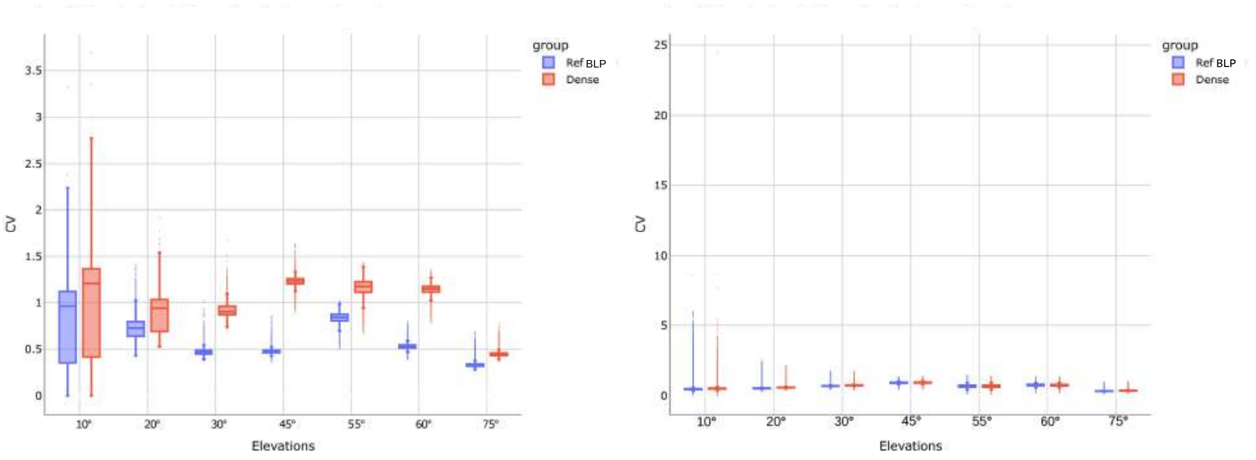
4.4.2 Azimuth-Elevation space

Light positions: Figure 4.18 shows the reference good light positions obtained using the proposed method 4.4.2 for four surfaces. In this example, the minimum number of light positions is set to 50 for all surfaces. The method adaptively generates a set of light positions that are most effective at capturing the reflectance of the surface. For the perfectly diffuse surface 13, the number of reference light positions is equal to the minimum number set and is evenly distributed along rings. For the anisotropic surface 1, the specular reflections are regular and symmetric, so most specular points have similar directions. Surface 3, which has details around a dent and a brush feature, has many more specular directions, and the method successfully identifies and samples them. Surface 14 has details with specular reflections at many random surface points, so the distribution of reference light positions is concentrated in random regions and evenly spread in most other regions.

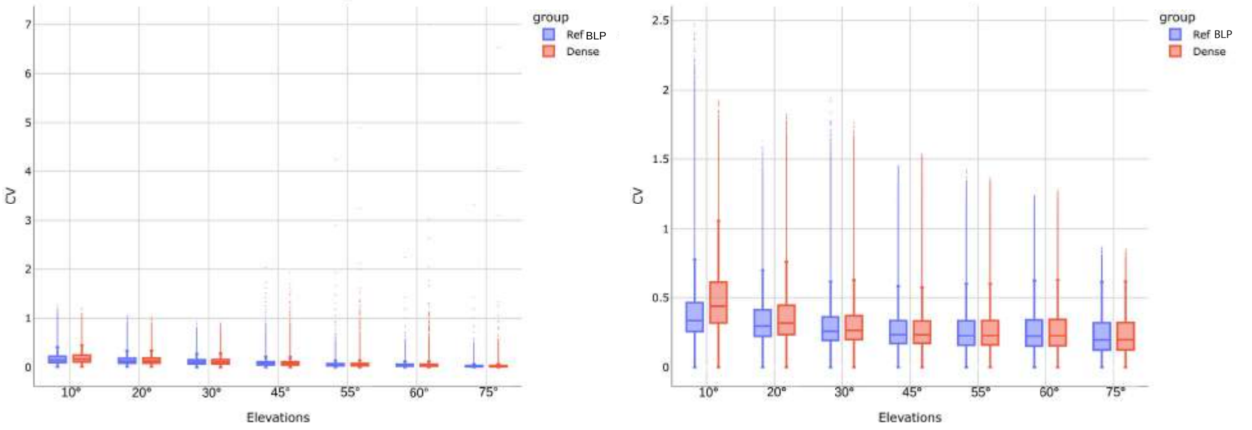
In contrast to ring acquisitions (which only have azimuth space), azimuth-elevation space acquisitions have both strong linear (elevation) and strong non-linear (azimuth) behavior. Therefore, using dispersion of gradients as a measure of performance is not appropriate for this case. Instead, we compare the reconstruction errors between the best light positions derived using our method, dense light positions, and sparsely homogeneous light positions.

To illustrate the impact of the choice of light position on RTI modeling, we perform a statistical analysis of the reconstruction errors. For that, we consider the following four acquisitions.

1. Uniformly distributed dense acquisition (1000 unique directions)
2. Uniformly distributed sparse acquisition (100 unique directions)



(a) Surface 1 (Simple brushed metal) (b) Surface 3 (Brushed metal with a dent)



(c) Surface 13 (Canvas painting) (d) Surface 14 (Ancient coin1)

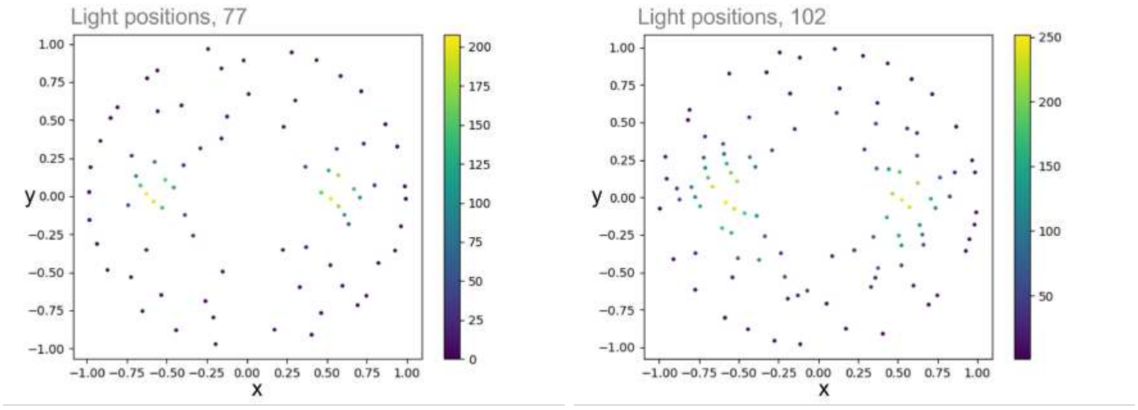
Figure 4.17: Comparison of the distribution of the coefficient of variation of the reflected signal gradients between dense vs reference best light positions acquisitions

3. Reference good light positions acquisition (number of unique directions adaptive to the surfaces as in Figure. 4.15)

4. Test acquisition containing 500 random light positions.

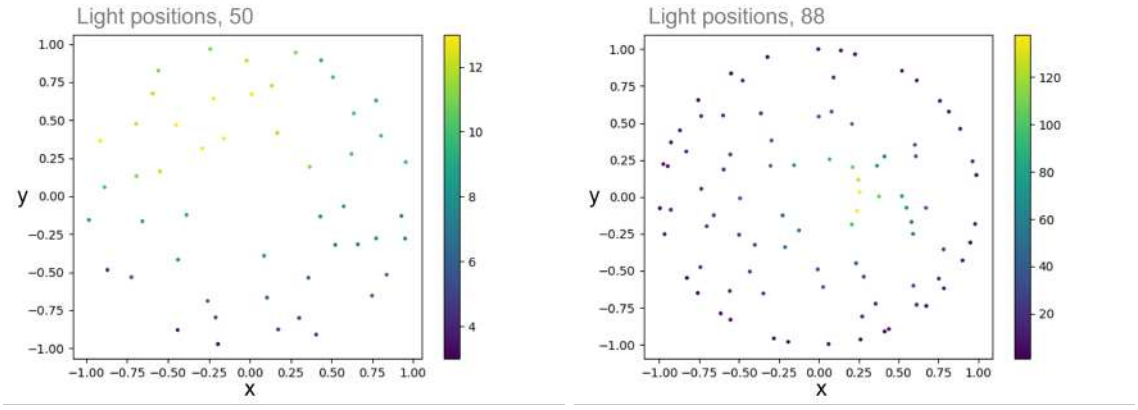
Our approach to the statistical analysis of reconstruction errors is illustrated in Figure. 4.19. We perform DMD and PTM RTI modeling on acquisitions with uniform distribution of dense light positions, uniform distribution of sparse light positions, and reference good light positions acquisitions. We then use the fitted model to relight the surface from the test acquisition light positions. The relighted images are compared to the corresponding actual captured images in the test acquisition by calculating the absolute point to point differences. The comparison gives the errors in the surface points reflection reconstruction. We estimate the probability density function and cumulative distribution function [Schervish and DeGroot, 2012] of the measured errors in both DMD and PTM relighted images.

Figure. 4.20 to 4.23 shows the PDF and CDF plots of the reconstruction errors of the surfaces. The x-axis ranges from 0 to 255, corresponding to the smallest and largest possible differences between the reconstructed image and the actual image. The height of the pdf plot at any point indicates how likely the reconstruction error is equal to that value. The slope of the cdf plot at any point indicates how likely that value is for the reconstruction error. The total area under a pdf plot is always equal to 1. The value of a cdf plot always ranges from 0 to 1. As can be seen from the figure, in the case of reference best light positions acquisition, the overall reconstruction error is



(a) Surface 1 (Simple brushed metal)

(b) Surface 3 (Brushed metal with a dent)



(c) Surface 13 (Canvas painting)

(d) Surface 14 (Ancient coin1)

Figure 4.18: Best light positions obtained for the surfaces from the respective dense acquisition

lower than in the case of uniform sparse acquisition. Reference good light positions acquisition errors are almost as small as dense acquisitions. In the case of non-Lambertian and semi-non-Lambertian surfaces (Surfaces 1, 3, 14), the difference is larger and more prominent.

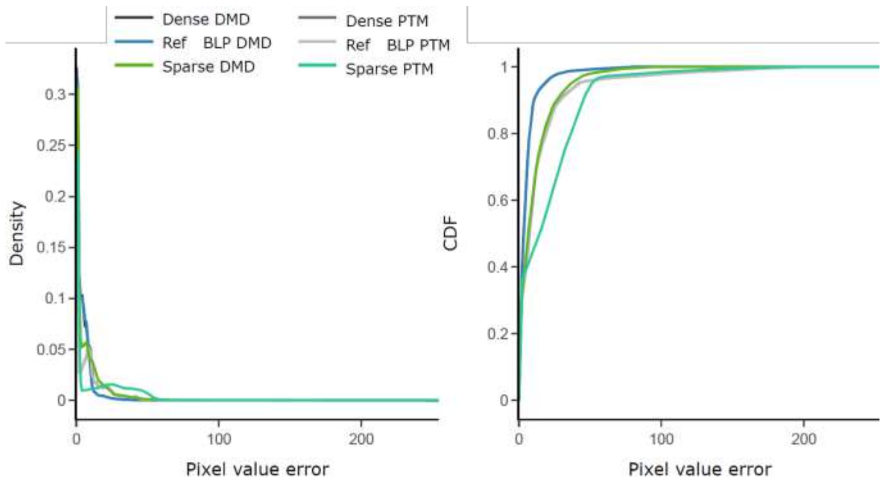


Figure 4.20: PDF and CDF of the reconstruction errors of surface 1 (simple brushed metal).

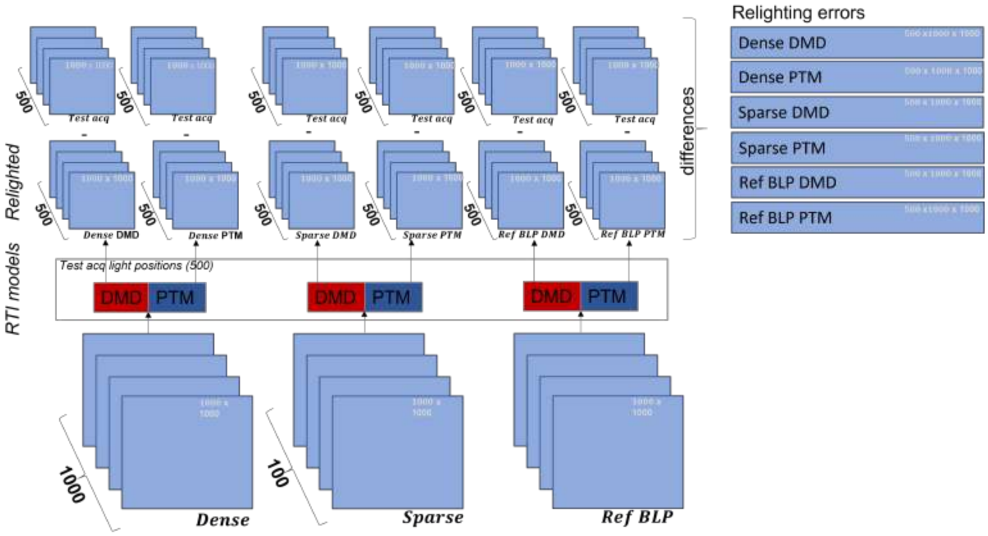


Figure 4.19

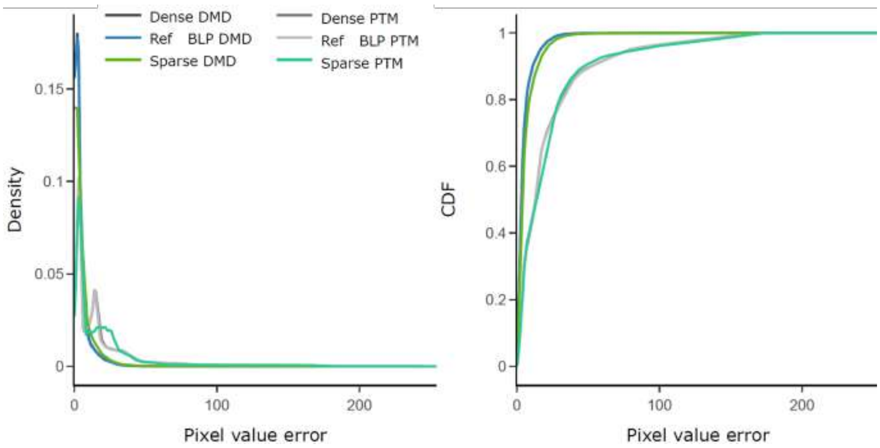


Figure 4.21: PDF and CDF of the reconstruction errors of surface 3 (brushed metal with dent).

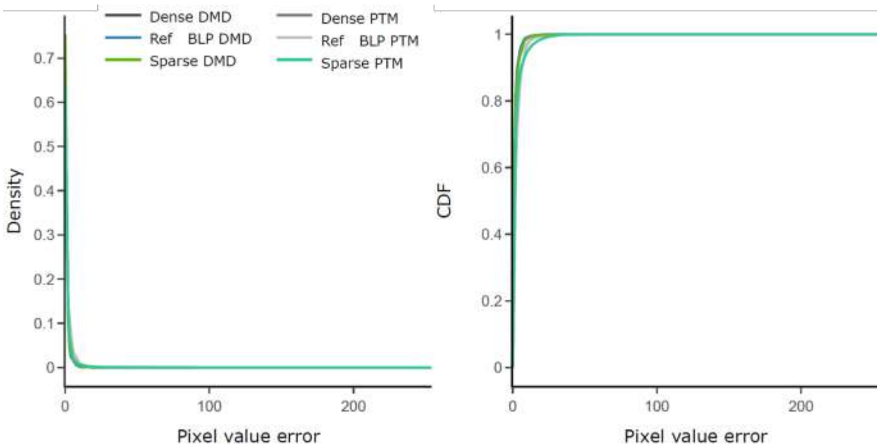


Figure 4.22: PDF and CDF of the reconstruction errors of surface 13 (canvas painting).

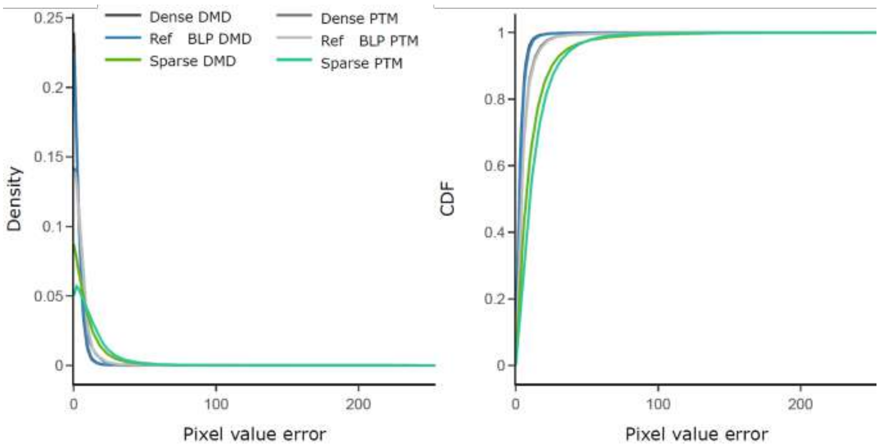


Figure 4.23: PDF and CDF of the reconstruction errors of ancient surface 14 (Ancient coin 1).

4.5 Conclusion

In this chapter, we presented a benchmark dataset containing dense acquisitions and reference sparse acquisitions with good light positions derived using a proposed novel method for both azimuth-only space acquisitions and azimuth-elevation space acquisitions. Our results showed that the proposed method was able to effectively generate a set of good light positions that improved the reconstruction of reflectance maps in both acquisition spaces. The contributions of our work include the development of a novel method for generating reference good light positions from dense acquisition and the creation of a benchmark dataset that can be used to evaluate the performance of different strategies used for estimating the light positions for a given surface.

The implications of our results are significant for the field of RTI, as they demonstrate the effectiveness of the proposed method for improving the quality of reflectance map reconstructions with fewer acquisitions.

Future work could explore the various strategies of RTI acquisition and benchmark them with these reference good light positions for the surfaces. Overall, our study provides a promising approach for improving the quality of reflectance map reconstructions through the use of optimal light positions. We believe that the proposed method and benchmark dataset will be useful resources for researchers and practitioners working in this field.

In the next chapter, we will explore strategies for *Next Best Light Position* problem, which is autonomously computing the good light positions for performing RTI acquisition of an unknown apriori object.

CHAPTER 5

Next Best Light Position for adaptive RTI acquisition

> ***Chapter overview***
>
> ***In the last chapter, we discussed the significance of selecting optimal light positions for RTI acquisition. We also presented techniques to determine the good light positions from dense acquisitions to assess the quality of an RTI acquisition. This chapter introduces new strategies to perform RTI acquisition that estimate good light positions as the acquisition progresses, adapting to the surface being captured. We present our approach to solve the problem of autonomously finding the Next Best Light Position in RTI acquisitions. These methods aim to address the challenges associated with capturing high-quality RTI data for objects with complex shapes and uneven surfaces. The Next Best Light Position (NBLP) problem refers to the challenge of determining the optimal light directions adaptive to a given surface for performing high quality RTI acquisition. The goal is to determine the minimum set of light positions during RTI acquisition to achieve the best fit to the surface properties of an apriori unknown object.***

5.1 Introduction

In RTI, when the light positions are chosen intuitively or set uniformly around the object, capturing the surface from all the pertinent light directions are not guaranteed and conversely all the captured directions often do not contribute significantly in the angular reflectance information. For example, at a grazing angle, the reflection of a relatively flat surface is often close to zero, and at orthogonal positions the pixel intensities are often saturated. Having too many images with zero or saturated pixel intensities, are not desirable, and can even be detrimental for the RTI model quality. We pioneered the identification of this problem and have developed innovative methods to address it. We coined the term "Next Best Light Position" (NBLP) to refer to this concept, analogous to the term "Next Best View" (NBV) used in 3D scanning [Pito, 1999, Karaszewski et al., 2016]. The quality of the resulting RTI model depends heavily on the lighting configuration under which the images are captured. By solving the NBLP problem, we can ensure that the surface appearance are best captured and modeled. In addition to that, autonomously detecting the best light positions adaptive to the surface being acquired will make the acquisition process itself efficient and saves time by avoiding repeating acquisitions by trial and error approach to achieve quality.

Furthermore, NBLP is an important step towards automation of RTI. The manual acquisition of RTI data is a laborious and time-consuming process. Automation of the RTI process can reduce the time and effort required for data acquisition and enable researchers to analyze large datasets efficiently. With the advancements in automation of RTI, the technology can become more accessible and can have a significant impact on the fields of archaeology, art conservation, and material inspection. Conventionally, RTI is performed with a set of 50 to more light directions that are evenly distributed in their spherical coordinates. These light directions are refined manually by trial and error [Webb and Wachowiak, 2011] for complex surfaces, which can be time-consuming and challenging for objects with intricate details or uneven surfaces.

Improving reproducibility in RTI acquisition is another reason to address this issue. Automating the acquisition process and using NBLP methods can help achieve this. For instance, in the cultural heritage field, monitoring changes in an object's surface appearance requires reproducing measurements. Manually reproducing RTI acquisition can be challenging due to factors such as positioning the surface and camera accurately. Using methods that autonomously determine the best light positions and adapt to new conditions can help capture the surface more accurately for comparison with previous measurements.

In summary, the motivation for developing NBLP methods include:

- *Acquisition quality:* To ensure quality acquisition by identifying all the light directions that reveals the details on the surface.
- *Acquisition size:* To reduce the size of the acquisition without compromising its quality, making it easier to post-process and analyze data such as HD-RTI, Focus variation RTI, and multi-spectral RTI that are usually computationally heavy methods.
- *Acquisition time:* To improve the efficiency of the acquisition process and enable RTI users to perform RTI acquisition more quickly.
- *Reproducibility:* To enhance reproducibility in applications such as monitoring changes in cultural heritage objects over time.

NBLP problem is still an open research problem in the RTI field. Methods exist for choosing optimal light directions in photometric stereo acquisitions Woodham [1980], which is a similar technique to RTI that involves capturing multiple images of an object under different lighting conditions. Photometric stereo uses the captured images to calculate the surface normals of the object, allowing for the reconstruction of the 3D surface. Unlike photometric stereo, RTI captures information about both the surface normals and the surface reflectance properties. This allows for a more detailed and accurate representation of the surface. In general photometric stereo, homogenously distributed light positions are enough since the objective is to measure the overall global shape of an abject and not the surface details.

[Gardi et al., 2022] proposed a method that computes the optimal positions of the light sources for the calibrated Photometric Stereo of a Lambertian surface. According to the paper, they estimate optimal light configurations by using a two-step approach: Firstly, they use Parameter Estimation (PE) to estimate the surface normal of an object from a set of images taken under different lighting positions. Secondly, they use Optimum Experimental Design (OED) to find the best lighting positions for capturing new images that can improve the accuracy of the surface normal estimation. They use a criterion called D-optimality that measures how well the lighting positions can distinguish different surface normals. The paper claims that their approach can produce more detailed and accurate surface normal estimates than conventional methods that use fixed or random lighting positions. However the paper states that their approach is limited and both uncalibrated photometric stereo methods and non-Lambertian property of the objects are beyond the scope of their approach.

[Drbohlav and Chantler, 2005] presents a similar theory for the optimal placement of photometric stereo lighting in the presence of camera noise. Their theory is based on the idea of Fisher information [Fisher et al., 1920], which is a measure of how much information an observable random variable (such as image intensity) carries about an unknown parameter (such as surface normal). The Fisher information matrix is used to calculate the covariance matrices associated with maximum-likelihood estimates. The authors use this matrix to quantify the uncertainty in estimating the surface normal from a set of images under different lighting conditions. They show that for three lights, the optimal light directions are orthogonal to each other. This minimises the uncertainty in scaled normal computation. A scaled normal is a vector that represents both the direction and the brightness of a surface point. The direction of the scaled normal is given by the surface normal. The magnitude of the scaled normal is equal to the albedo. For more than three lights, they use a numerical optimization technique to find the optimal light directions that minimize the uncertainty function. They start with an initial guess of light directions and iteratively update them until they reach a local minimum of the uncertainty function. The limitations of their approach are: it assumes Lambertian reflectance. It does not consider occlusions or shadows, which may significantly affect the observed intensities. It requires a prior knowledge of camera noise level, which may not be easy to obtain or vary across images.

[Argyriou et al., 2013] approaches the problem by creating a sparse representation of the illumination arrangement and estimates the light directions using L1 optimisation. Again, they assume that the surface is Lambertian. Their approach is to sample a set of possible light directions on a hemisphere around the object and then select a subset of them that minimizes the reconstruction error and the presence of shadows. They formulate this as an L1 optimization problem, where they try to find a sparse vector of weights for each sampled light direction that best explains the observed image intensities. The authors state that their method can work with any kind of light source, but they need to have some idea of what the object looks like in general. For example, if they want to reconstruct faces, they use a synthetic model face to find the best light positions, and then use those positions for all real faces. They also say that the number of samples they need depends on the object type, and they can use a prototype object to get a generic shape.

The methods discussed cannot be applied directly to RTI acquisition of an unknown apriori surface because they are limited to Lambertian surfaces and focuses on 3D information than reflectance properties of the object surfaces. There is clearly a need for novel methods to find best adaptive light configurations to perform RTI acquisitions. However, we draw inspiration from some of the error minimization approaches utilized in these works.

5.2 Methods

In the previous chapter we discussed methods to remove less impacting light directions from dense acquisition to sample the signals. Contrast to that, in this chapter, we aim to find the pertinent light directions to add as an acquisition evolves. We use the list of symbols, abbreviations, and their meanings as given in the table 5.1 in this chapter.

Symbol	Description
θ	Azimuth angle
ϕ	Elevation angle
m	Number of light positions in an acquisition.
n	Resolution of the camera used to acquire the surface (number of pixels used to capture the surface. Each pixel represents a point on the surface).
i	Index of a pixel in a vectorized image.
j	Index of a light direction in an acquisition.
p	A point on the surface. p_i represents the i^{th} pixel (surface point) in an acquired image
d	Light direction. d_j represents the j^{th} light direction in an acquisition
s	Luminance signal corresponding to a point p. s_{ij} represents the signal value of the point p_i the direction j. The signal value is just the pixel value which is a direct measure of luminance.
g	Gradient of a signal. The gradient of a point p_i at direction d_j is represented as g_{ij}
w	Weight of a direction. w_{ij} represents the weight of p_{ij} representing the relative effect of point to the acquisition quality
$\hat{s}$	Normalized signal. It represents the gradients of a signal normalized across all the light directions between 0 to 1. The normalized value of signal $\hat{s}_i$ at light direction j, $\hat{s}_{ij} = \frac{s_{ij} - \min_j (s_{ij})}{\max_j (s_{ij}) - \min_j (s_{ij})}$
G	Weighted average gradient score. G_j represents the weighted average gradient score at direction j and is calculated as $G_j = \frac{\sum_{i=0}^{n} w g_{ij}}{n}$

Table 5.1: Notations used in this chapter

5.2.1 Optimal light directions in azimuthal space for a single surface point

To address the NBLP problem for RTI, we initially simplified it by focusing on a single surface point (represented by one pixel) and attempting to determine the optimal light positions for that point. Our intention was to later extend this estimation method to all surface points to find set of light directions which is globally optimal. To further simplify the problem, we first limited our analysis to the azimuth space. This was motivated by the fact that the non-linearity of surface reflectance in the azimuth space is generally higher than in the elevation space. This observation is illustrated in the intensities plot Figure 5.1 where it is observed that at different elevations, the shape remains same and just the magnitude varies. It can be seen that the signal is almost linear along the elevation, but non-linear along the azimuth. Therefore, we hypothesize that it is adequate to optimize the positions only in the azimuth space and then interpolate the optimized azimuthal points onto evenly spaced elevations.

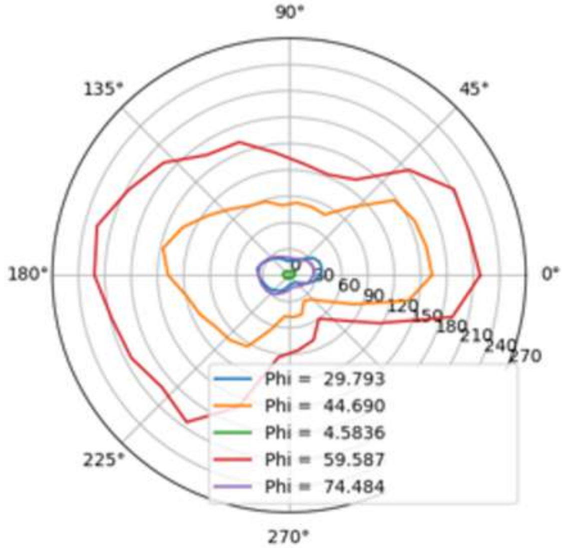


Figure 5.1: Plot of variation of intensity observed for a single pixel with changes in elevation and azimuth

To find the good light directions in azimuthal space for a single pixel, we use an adaptive approach where we start with a few evenly spaced directions. Then we capture additional directions where the gradient g is high, and avoid new acquisition points in areas where the gradient is low.

The process consists of the following steps:

1. Acquire Initial Measurements: Take luminance measurements of the surface point illuminated from a set of evenly spaced directions in the azimuthal space. A more detailed analysis on the choice of initial number of light directions will be given later in the results and analysis section.

2. Append the first luminance value and azimuth angle to the end of the respective arrays to make them circular, i.e., $s_{m+1} = s_1$ and $\theta_{m+1} = \theta_1$, where m is the number of initial light directions.

3. Compute the gradients of the point between adjacent directions by taking the difference in luminance values divided by the azimuthal difference between the points, given by:

   $$g_{ij} = abs(\frac{s_{i(j+1)} - s_{ij}}{\theta_{(j+1)} - \theta_j}), \tag{5.1}$$

   where s_{ij} and s_{ij+1} are the luminance values at directions j and $j+1$ respectively, and θ_j and θ_{j+1} are the corresponding azimuth angles.

4. Identify the points where the gradient is higher than a threshold. We used the standard deviation of the signal, σ to compute the threshold as $\max(\sigma, 10)$. However, we observed that any statistical thresholding strategy like interquartile range, z-scores gives similar results as the acquisition evolves.

5. Interpolation of points: Insert additional points to the acquisition between the previously identified points using an adaptive spacing scheme based on the gradient values. Higher gradient values correspond to smaller spacing and more sample points, while lower gradient values correspond to larger spacing and fewer sample points.

6. Acquire the additional measurements with the interpolated points and append the newly acquired measurements to the previous measurements.
7. *Termination:* Repeat steps 2-6 until, firstly, there are no additional points where the graident is higher than an absolute minimum that depends on the object being acquired and the desired size of the acquisition. Secondly, all the additional points are very close to already acquired points.

Results

In order to test the effectiveness of the adaptive sampling approach, we conducted a series of experiments on different surfaces in our dataset. We choose different points, that are diffuse, semi-specular, and specular. For each surface point, we acquired data using a varying number of initial light directions (1, 5, 10, 15, 50, and 75) and applied the adaptive sampling approach. The goal of the experiment is to determine if the strategy can accurately identify the light positions that produce unique features in the surface's reflectance characterization. Additionally, we aim to analyze and suggest an optimal number of initial light positions that will ensure accurate reflectance characterization as data acquisition progresses. We present the results of this adaptive sampling approach to acquire a surface point by plotting the pixel values observed in radial polar form, where the radius represents the pixel value and the angles represents the azimuth angles of the light directions at each iteration. The progress of the acquisition is shown when the initial number of light directions is changed.

Diffuse point Figure.5.3 illustrates the progress of acquisitions optimized for a diffuse point with different initial numbers of uniformly distributed light directions. This point corresponds to the pixel (1298, 925) of the surface painting surface (surface 13 in the rti-dataset presented in the previous chapter) as shown in Figure.5.4

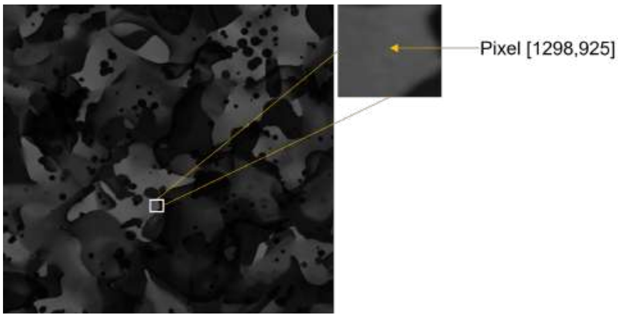
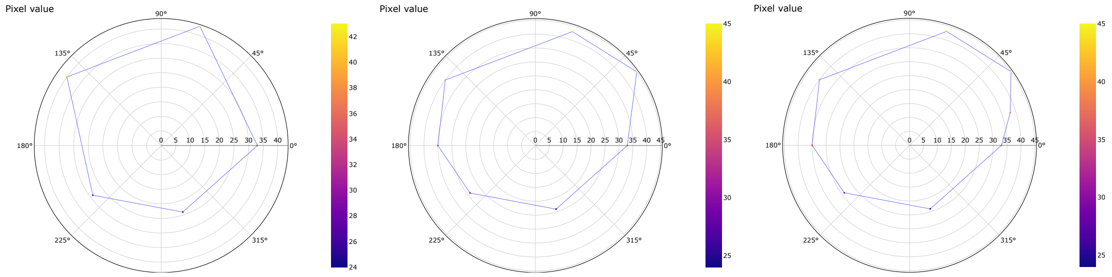
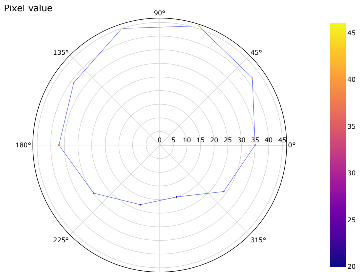


Figure 5.2: A diffuse point from the canvas painting surface.



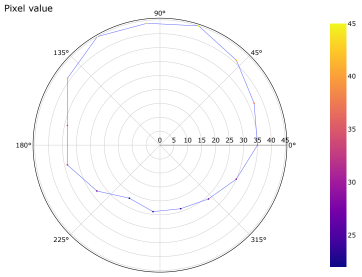
Initial 1^{st} iteration 2^{nd} iteration

(a) Initial number of directions: 5. The acquisition stopped after 2^{nd} iteration where all the gradients are than 10, the threshold.



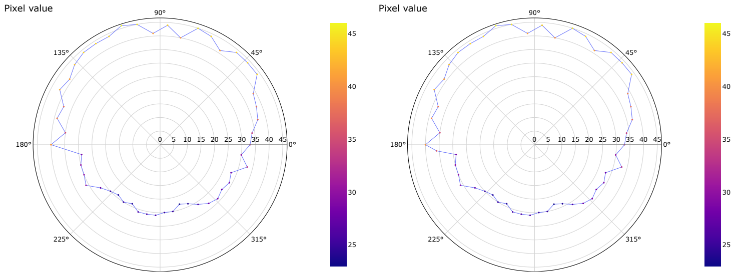
Initial

(b) Initial number of directions: 10. The acquisition did not continue beyond initialization since the max gradient is already less than 10, the threshold



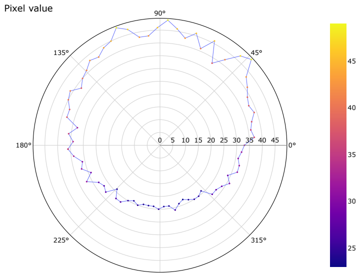
Initial

(c) Initial number of directions: 15. The acquisition did not continue beyond initialization since the max gradient is already less than 10, the threshold



Initial 1^{st} iteration

(d) Initial number of directions: 50. The acquisition stopped after 1^{st} iteration where all the gradients are than 10, the threshold..



Initial

(e) Initial number of directions: 75. The acquisition did not continue beyond initialization since the max gradient is already less than 10, the threshold

Figure 5.3: Optimizing the acquisition of a diffuse point on the ancient coin1 (Surface 13 in the RTI-dataset) in the azimuthal space with different initial number of light directions. The These acquisitions were carried out by positioning the light source at various angles in a ring at 45˙ elevation.

A diffuse point reflects light uniformly in all directions, regardless of the incident light direction. The intensity of the pixel corresponding to the diffuse point remains almost constant for any azimuth angle unless there is masking or shadow casting on it due to the 3D shape of the object. The change in reflectance due to masking and shadowing is generally observed over a wide range of azimuth angles, unlike the narrow angle observed in the case of specular reflection. Our results show that any initial number of light directions greater than 5 is sufficient to achieve quality capture of diffuse reflectance.

Semi specular point In Figure.5.5, we have shown the evolving of adaptive acquisition for a sample semi specular point. This point corresponds to the pixel (1000, 1000) of the ancient coin surface (surface 14 in the rti-dataset presented in the previous chapter) as shown in Figure.5.4

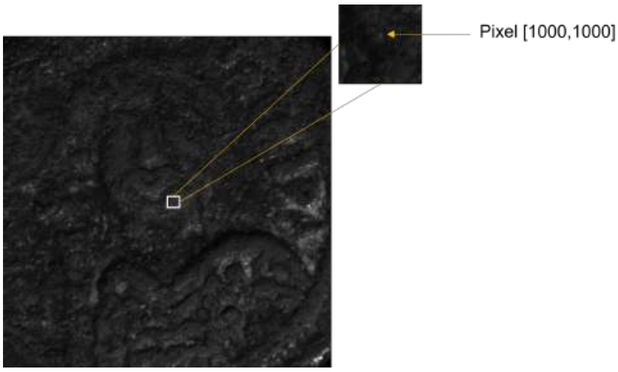
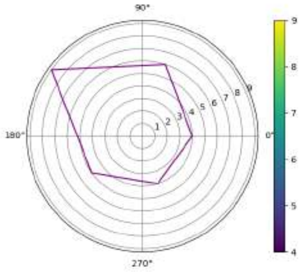
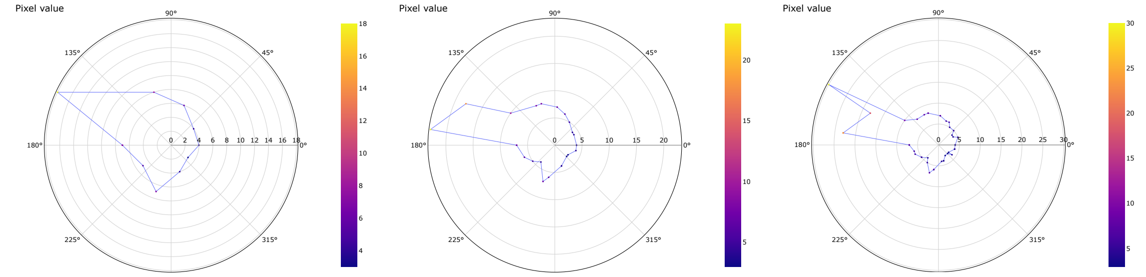


Figure 5.4: A semi specular point from the coin surface.



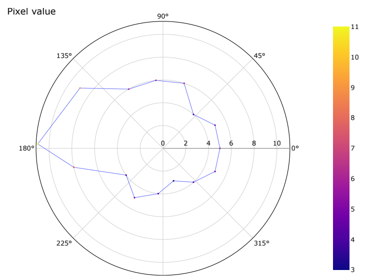
Initial

(a) Initial number of directions: 5. The acquisition did not proceed after initialization since all the gradients are than 10, the threshold.



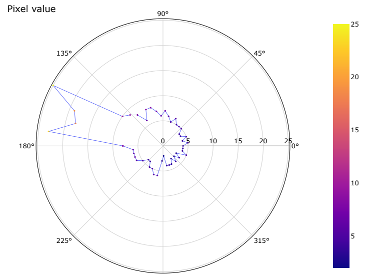
Initial $1^{st}iteration.$ $2^{nd}iteration.$

(b) Initial number of directions: 10. The acquisition stopped after 2^{nd} iteration since the max gradient is less than 10, the threshold



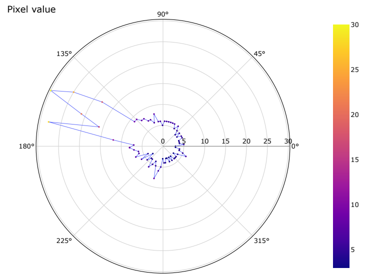
Initial

(c) Initial number of directions: 15. The acquisition did not proceed after initialization since all the gradients are than 10, the threshold.



Initial

(d) Initial number of directions: 50. The acquisition did not proceed after initialization since all the gradients are than 10, the threshold.



Initial

(e) Initial number of directions: 75. The acquisition did not proceed after initialization since all the gradients are than 10, the threshold.

Figure 5.5: Optimizing the acquisition of a semi specular point on the ancient coin1 (Surface 14 in the RTI-dataset) in the azimuthal space with different initial number of light directions. These ring acquisitions by positioning the light source in 45˙ elevation ring.

Among these results, it can be observed that the semi-specular lobe is captured when the initial number of light positions is 10, 50, and 75. In case of 5 number of initial light directions, the acquisition terminates at the first iteration itself, since none of the light directions satisfy the threshold criteria described in step 4. In case of 10 number of initial light directions, the acquisition progresses since here the locations of the initial light directions are advantageous in detecting semi specular lobe. Interestingly, in case of 15 number of initial light directions which is more than the previous case, the initial light positions doesn't meet the threshold criteria and hence the acquisition terminates at the first iteration itself, thereby misses detecting the semi-specular lobe. In case of 50 and 75 number of initial light directions, the number of sampling points are already more than enough to capture the semi-specular lobe. Number of initial light directions higher than this just improves the resolution of the reflection lobe and does not contribute any significantly important information than the previous acquisitions.

Specular point Figure.5.7 illustrates the progress of acquisitions optimized for a specular point with different initial numbers of uniformly distributed light directions. This point corresponds to the pixel (1000, 1000) of the brushed metal surface (surface 1 in the rti-dataset presented in the previous chapter) as shown in Figure.5.6

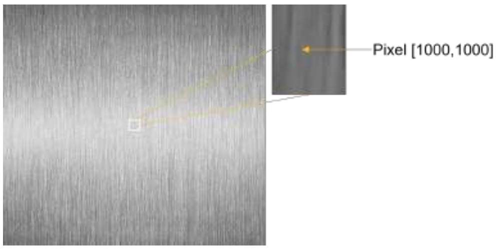
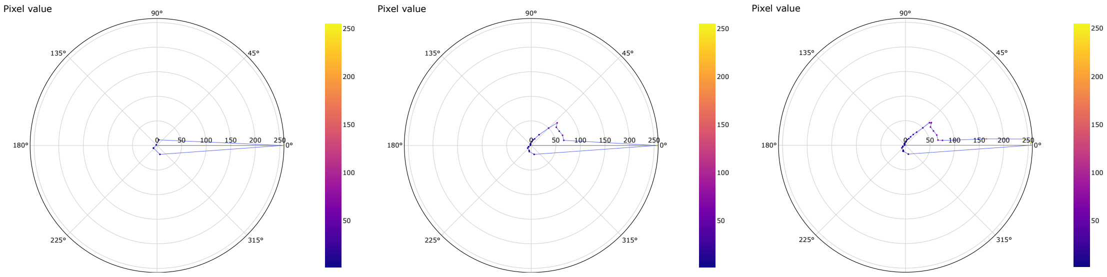


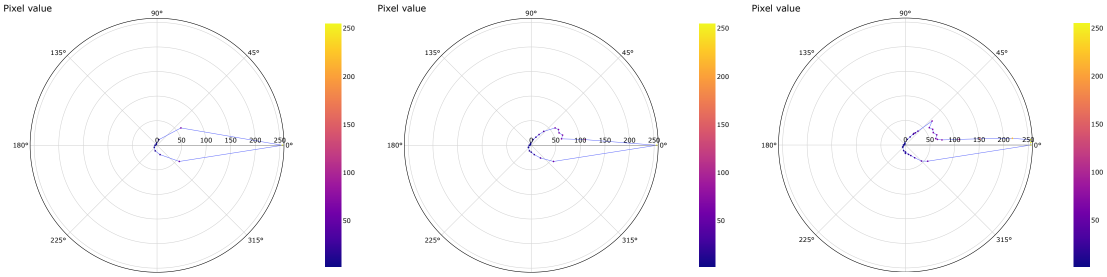
Figure 5.6: A specular point from the brushed metal surface.

A specular point reflects light in a specific direction. Unlike, the semi-specular case, the specular points exhibits huge gradients towards the specular direction. The rise is very steep that all the points in the region closer to the inflection point of the plot have higher gradients. It can be obsered that, in case of 5 number of initial acquisition, the acquisition progresses till 3 iterations and succeeded in detecting the specular direction. However, it identifies only one of the two inflection points. In case of 10 initial number light directions also, the acquisition progresses till three iterations and still misses the second inflection point. We observe that for specular points, 15 number of intial light directions are generally sufficient to achieve quality acquisition.



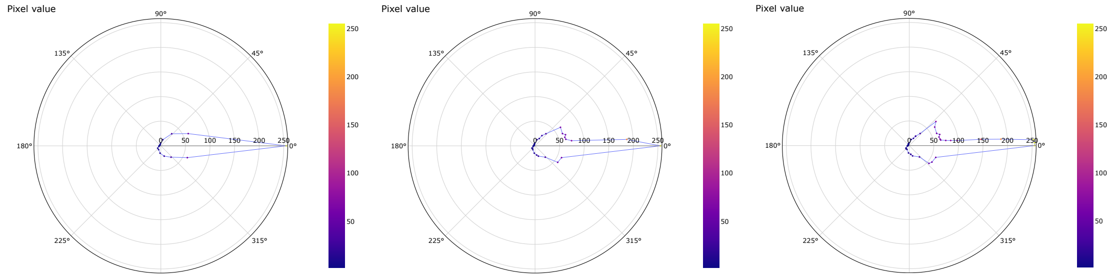
Initial 1st iteration 2nd iteration

(a) Number of initial directions: 5. The acquisition stopped after 2^{nd} iteration where all the gradients are than 10, the threshold.

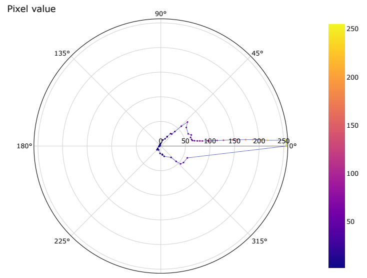


Initial 1st iteration 2nd iteration

(b) Number of initial directions: 10. The acquisition stopped after 2^{nd} iteration where all the gradients are than 10, the threshold.

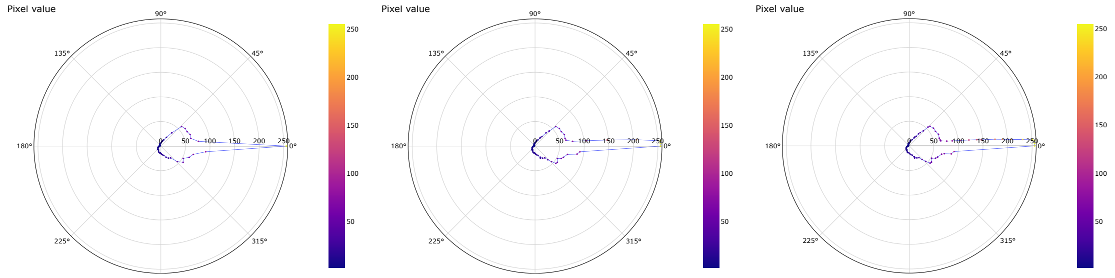


Initial 1^{st} iteration 2^{nd} iteration



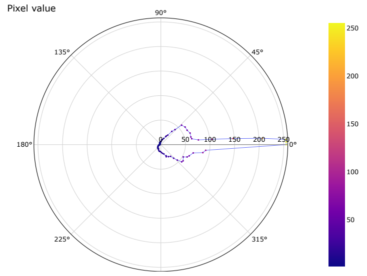
3^{rd} iteration

(c) Number of initial directions: 15. The acquisition stopped after 3^{rd} iteration where all the gradients are than 10, the threshold.



Initial 1^{st} iteration 2^{nd} iteration

(d) Number of initial directions: 50. The acquisition stopped after 2^{nd} iteration where all the gradients are than 10, the threshold.



Initial

(e) Number of initial directions: 75. The acquisition did not proceed after initialization since there were no gradient value higher than the threshold 10.

Figure 5.7: Optimizing the acquisition of a specular point on the simple burshed metal surface (Surface 1 in the RTI-dataset) in the azimuthal space with different initial number of light directions. These acquisitions we carried out by positioning the light source in a ring at 45˙ elevation.

In summary, adaptive acquisition strategy is able to capture the overall reflectance of a point. We observe that 15-20 initial light positions is generally good to achieve quality acquisition of any surface point using this strategy. The choice of initial light directions is more critical in case of a semi specular point than diffuse or specular point. We can overcome the problem of missing the critical light directions due to bad choice of initial set of light directions by adding a few random directions in addition to the uniformly spaced set of light directions. By critical light direction we mean those direction where the surface reflection has strong changes.

5.2.2 Optimizing light directions globally in Azimuthal space

As discussed earlier, the adaptive sampling approach using gradients is effective in optimizing light directions for acquiring a single point in the azimuthal space. However, extending this approach to optimize light directions for all surface points globally is not that straightforward. Firstly, this approach is limited by computational cost, and secondly, optimizing light directions for each surface point individually will lead to oversampling, resulting in excessive redundancy.

One possible solution is to average the signals and optimize the light directions based on the averaged signal. This approach may work for homogenous 2D surfaces where the reflectance of the points is similar. However, in case of non-homogenous 3D surfaces, especially when active light is used for measurement, averaging the signals can adversely affect the measurement and cannot be used for optimization.

We propose a method to optimize light directions globally using weighted averages approach. The process consists of the following steps:

1. Acquire initial acquisition: We make an initial acquisition of the surface by illuminating it from set of evenly spaced directions in the azimuthal space. Based on our study in the last section, we acquire 15 number of evenly spaced angles as our initial acquisition.

2. Since our goal is the global optimization of light directions, it is required to normalize each signal to a range (0,1) in order to treat every signal equally. A signal s_i is normalized as

$$\hat{s_i} = \frac{s_i - \min s_i}{\max s_i - \min s_i}. \tag{5.2}$$

3. Compute the gradients of the normalized signal values between adjacent directions by taking the difference in normalized signal values divided by the azimuthal difference between the points, given by:

$$g_{ij} = \frac{\hat{s}_{i(j+1)} - \hat{s}_{ij}}{\theta_{(j+1)} - \theta_j}. \tag{5.3}$$

4. Let's say we have n surface points and m light directions. For each light direction d_j, we assign weights to each point p_i based on its gradient at that light direction. The weight w_{ij} for point p_i at light direction d_j is calculated as

$$w_{ij} = \frac{n \times g_{ij}}{\sum\limits_{i=0}^{n} g_{ij}}, \tag{5.4}$$

5. Once we have calculated the weights for all points at all light directions, we use these weights to calculate the weighted average of the gradients at each light direction. The weighted average gradient G_j at light direction j is calculated as:

$$G_j = \sum_{i=1}^{n} w_{ij} g_{ij}, \tag{5.5}$$

6. We use these weighted gradient scores to estimate the next best light positions. We apply the same standard deviation thresholding used in the single pixel method in this case too to interpolate new light directions between light directions with higher weighted gradient scores.

7. Termination: We repeat steps 2-6 until all the gradient values of the actual signals (not the normalized signal) fall below a certain threshold value. This threshold is dependent on the desired acquisition size and accuracy. Setting a very low threshold results in a dense acquisition. In our experiments, we choose a threshold gradient gray level of 10.

	d_1	d_2	$\cdots$	d_m
p_1	g_{11}, w_{11}	g_{12}, w_{12}	$\cdots$	g_{1m}, w_{1m}
p_2	g_{21}, w_{21}	g_{22}, w_{22}	$\cdots$	g_{2m}, w_{2m}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
p_n	g_{n1}, w_{n1}	g_{n2}, w_{n2}	$\cdots$	g_{nm}, w_{nm}
Weighted gradient scores, G	$\sum_{i=1}^{n} w_{i1}g_{i1}$	$\sum_{i=1}^{n} w_{i2}g_{i2}$	$\cdots$	$\sum_{i=1}^{n} w_{im}g_{im}$

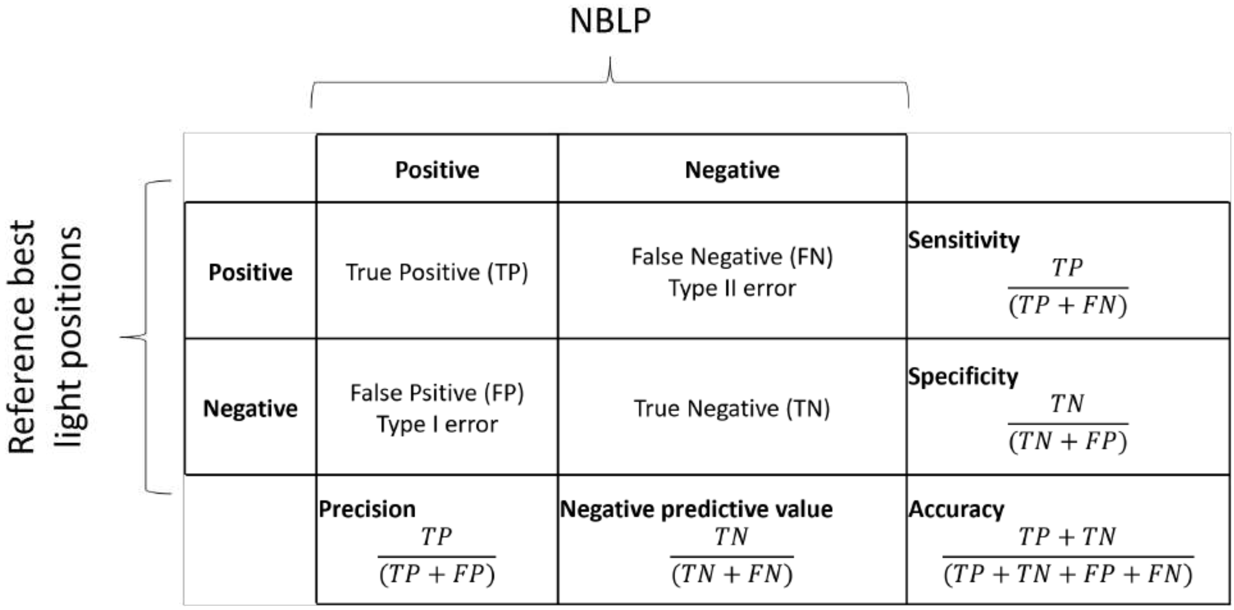
Table 5.2: Weighted average gradients approach to optimize light directions globally in the azimuth space.

The method of weighted gradient scoring is simple yet powerful approach for optimizing global azimuth. The weighted gradients score computation is illustrated in Table.5.2, where the gradients of each point are calculated for every light direction. Based on these gradients, weights are assigned to each point. The weighted average gradients are then computed and used as a score to estimate the optimal position for the next illumination. This process involves only a few matrix operations, making it highly efficient to compute. By utilizing this approach, it is possible to quickly determine the best light position for global azimuths optimization.

Evaluation metric In order to evaluate the performance of our approach, we compare our results with the reference good light positions for the surfaces derived from dense acquisitions in the previous chapter. To compare the light positions with the reference good light positions, we construct a confusion matrix. A confusion matrix as shown in Figure.5.8 is a table that is used to describe the performance of a classification model. In our case, we are classifying the points in the azimuthal ring as good or not good light positions.

First, we divide the azimuthal ring into 360 points, 1 degree apart each. This allows us to have a fine-grained representation of the azimuthal space. Then, we classify each point as an good light position or not an good light position based on whether or not it is present in the reference good light position set. To account for small variations in the light positions, we use a tolerance of +/- 1.0 degrees when classifying the points.

We do the same with NBLP test light positions and compare their classification with that of the reference good light positions. This comparison allows us to see how well the test light positions match the reference good light positions. From this comparison, we construct a confusion matrix, which shows the number of true positives, false positives, true negatives, and false negatives.



		NBLP		
		Positive	**Negative**	
Reference best light positions	**Positive**	True Positive (TP)	False Negative (FN) Type II error	**Sensitivity** $\frac{TP}{(TP+FN)}$
	Negative	False Psitive (FP) Type I error	True Negative (TN)	**Specificity** $\frac{TN}{(TN+FP)}$
		Precision $\frac{TP}{(TP+FP)}$	**Negative predictive value** $\frac{TN}{(TN+FN)}$	**Accuracy** $\frac{TP+TN}{(TP+TN+FP+FN)}$

Figure 5.8: confusion matrix

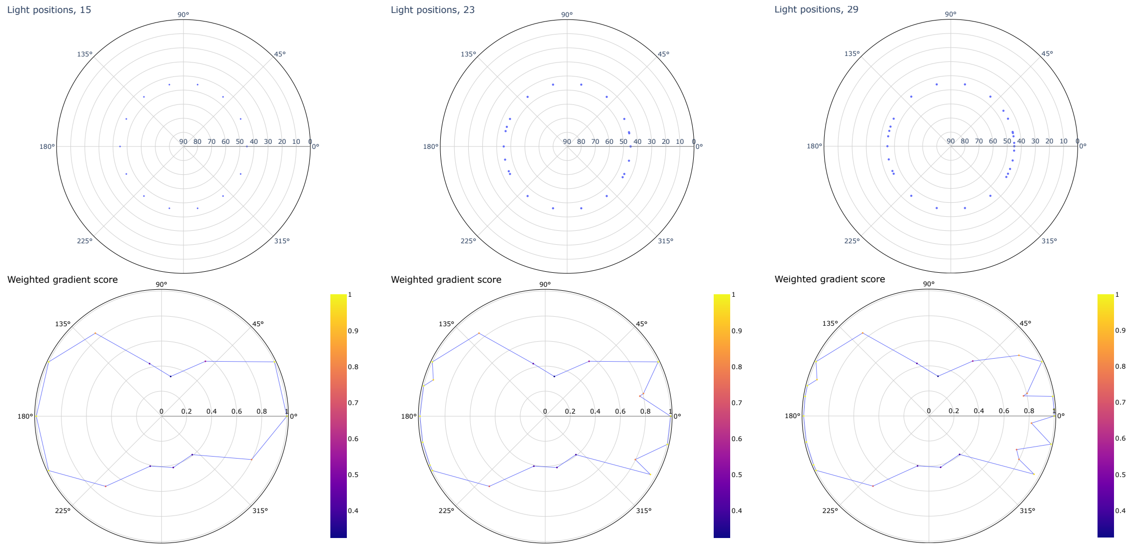
The confusion matrix provides a summary of how well the NBLP test light positions match the reference good light positions. By looking at the number of true positives and true negatives, we can see how many NBLP light positions were actually good light positions. Also, there are several metrics that can be calculated from the confusion matrix to evaluate the performance of our method. Some of these metrics include precision, negative predictive value, sensitivity (also known as recall), specificity, and accuracy. Precision measures how many of the NBLP test light positions that were identified as good light positions were actually good light positions. Negative predictive measures how many of the NBLP test light positions that were classified as not good light positions were actually not good light positions. Sensitivity (or recall) measures how many of the good light positions were correctly identified by the method. Specificity measures how many of the not good light positions were correctly identified by the method. Accuracy measures how many of the NBLP test light positions were correctly identified overall. These metrics provide different perspectives on the performance of the classification and can help us evaluate how well the NBLP light positions match the reference good light positions.

Results In order to test our method, we implemented and tested it on the same four surfaces in the RTI dataset presented in the previous chapter. We performed the NBLP acquisition in rings at elevations of 10˙, 20˙, 30˙, 45˙, 55˙, 60˙, 75˙ similar to the previous chapter.

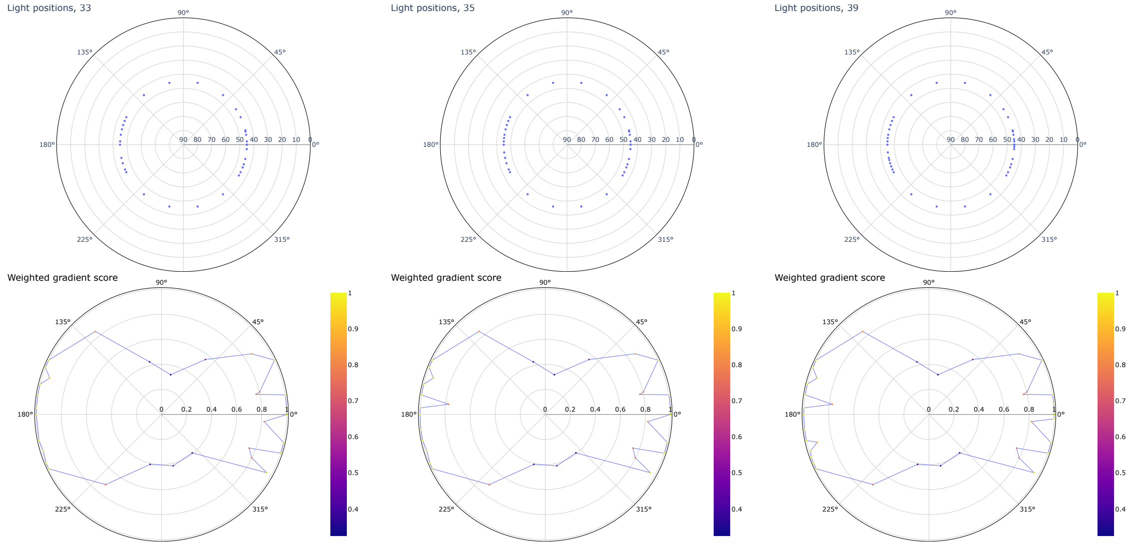
To present the information in a clear and concise manner, in this section we demonstrate only a selected few results. Additional results are available in the appendix for further reference. We present the evolution of acquisition of simple brushed metal at elevation 30˙. We present this by plotting the weighted gradient scores in a radial polar plot and showing the respective light positions in each iteration. This allows us to visually see how the good light positions are estimated as the acquisition progresses. Then we present its final overall light positions compared with that of the reference light positions. For the other 3 surfaces - brushed metal with a dent, ancient coin 1 and the canvas painting, we present the over all NBLP light positions compared with that of the reference light positions. To ease comparison, we present the confusion matrix for each acquisition immediately following the respective acquisition results.

Simple brushed metal The NBLP results for a simple brushed metal surface at the elevation 45˙ are shown in Figure.5.9 . The top figure displays the iteration steps during the NBLP acquisition, while the bottom figure compares the overall distribution of light positions between the reference good acquisition and the final NBLP acquisition. The illustration is accompanied by its corresponding confusion matrix.

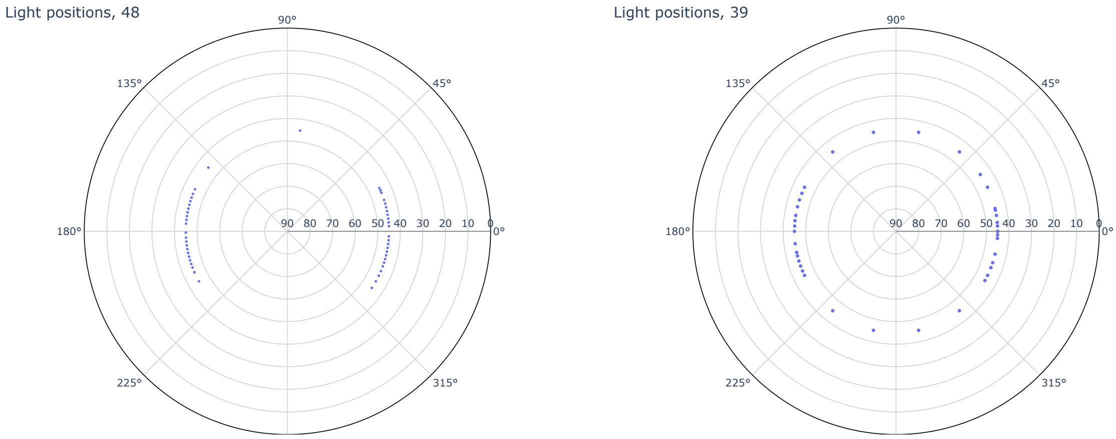
Elevation 45˙



(a) Initial acquisition (b) 1^{st}iteration (c) 2^{nd}iteration



(d) 3^{rd}iteration (e) 4^{th}iteration (f) 5^{th}iteration



(g) Overall distribution of the light positions in reference good acquisition (left) and NBLP acquisition (right)
Figure 5.9: NBLP acquisition on a simple brushed metal surface in the azimuthal space ring at an elevation of 45˙.

	Positive	Negative	
Positive	43	15	Sensitivity 0.74
Negative	41	261	Specificity 0.84
	Precision 0.51	Negative predictive value 0.94	Accuracy 0.86

Table 5.3: Confusion matrix of the results for simple brushed metal surface at elevation 45˙

In the plots, we observe that our method iteratively adapts to the surface and identifies optimal light positions. Convergence is typically achieved within 4-6 iterations. Due to the anisotropic nature of brushed metal, we observe that the optimal light directions tend to be more towards the orthogonal direction to the brush direction. The results show higher performance in identifying and avoiding the non-optimal light directions which is also an important objective of NBLP that reduces the acquisition size without compromising the quality.

Brushed metal with a dent In Figure.5.10 we present the comparison of overall NBLP estimated light positions and the reference light positions for the brushed metal surface with a dent (surface 2 in our dataset) at elevation 45˙.

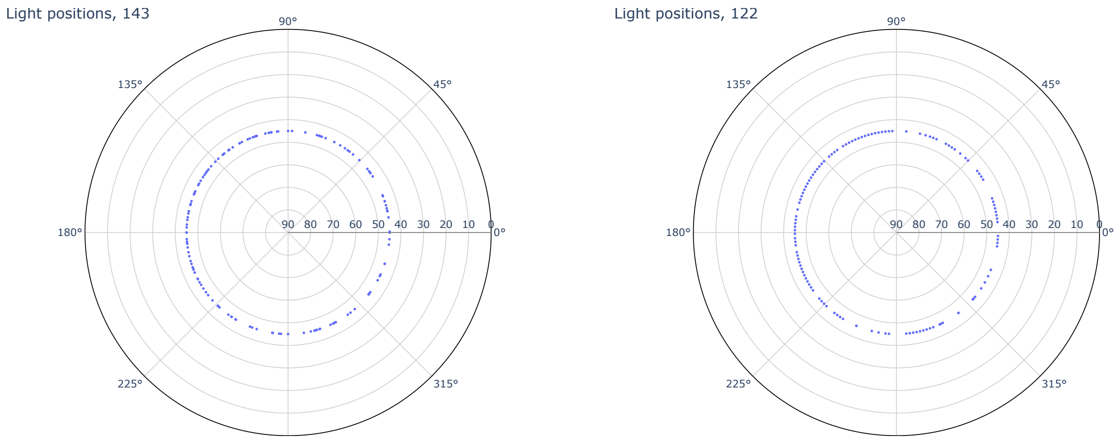


Figure 5.10: Overall distribution of the light positions in reference good acquisition (left) and NBLP acquisition (right) of brushed metal surface with a dent at elevation 45˙

	Positive	Negative	
Positive	139	17	Sensitivity 0.89
Negative	22	182	Specificity 0.89
	Precision 0.86	Negative predictive value 0.91	Accuracy 0.89

Table 5.4: Confusion matrix of the results for brushed metal surface with a dent at elevation 45˙

For brushed metal with a dent, all the metric scores are high. Here we see that number of true positives are higher than the light positions - this is because, some of the light positions in the reference good acquisition are close (within +/- 1 degrees apart) to more than one light position in the NBLP acquisition.

Ancient coin 1 In Figure.5.11 we present the comparison of overall NBLP estimated light positions and the reference light positions for the ancient coin 1 (surface 14 in our dataset) at elevation 45˙.

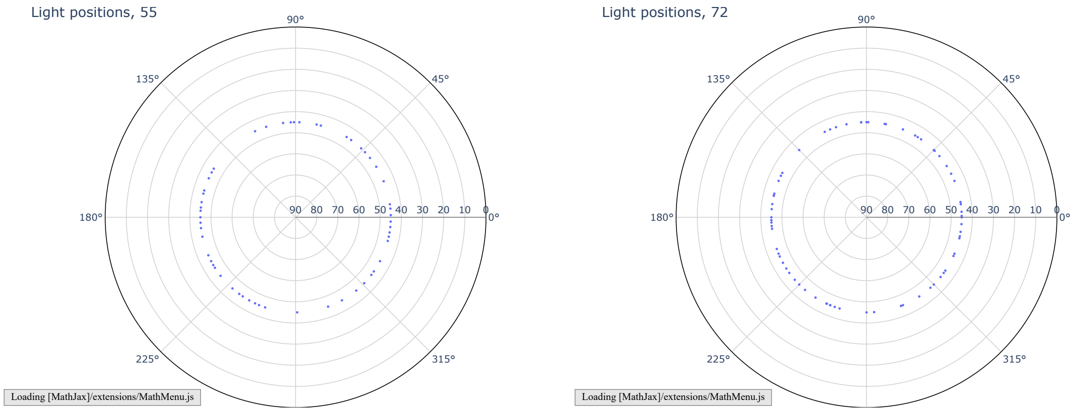


Figure 5.11: Overall distribution of the light positions in reference good acquisition (left) and NBLP acquisition (right) of ancient coin 1 at elevation 45˙

	Positive	Negative	
Positive	70	20	Sensitivity 0.77
Negative	53	217	Specificity 0.80
	Precision 0.56	Negative predictive value 0.91	Accuracy 0.79

Table 5.5: Confusion matrix of the results for ancient coin 1 at elevation 45˙

This surface is relatively complex surface phenomenon consisting of varying degree of specularity among different points. It contains details that result in random surface normals. It is interesting to note that our method is able to adaptively find the directions that effectively characterizes all the points simultaneouly. As seen the NBLP estimated light positions are close to the distribution of light directions in the reference good acquisition.

canvas painting Here, we present the comparison of overall NBLP estimated light positions and the reference light positions for the canvas painting surface(surface 14 in our dataset).

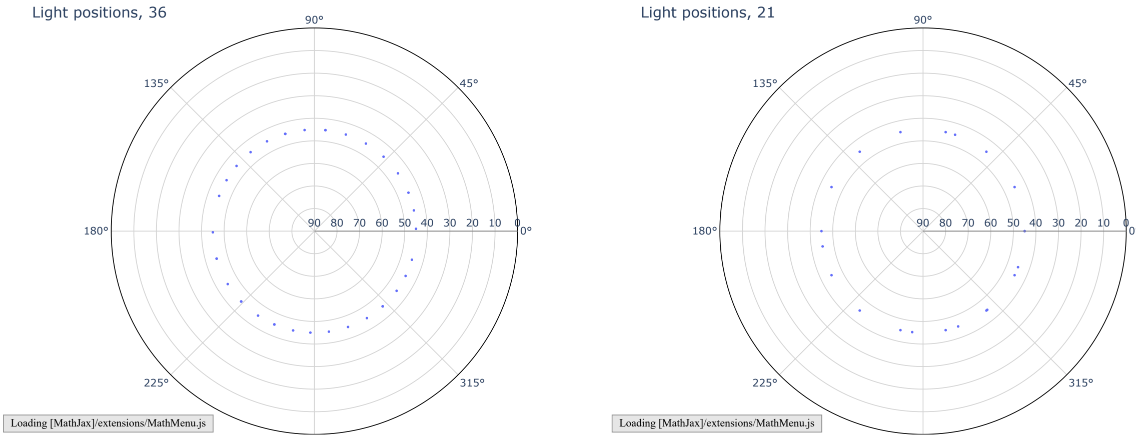


Figure 5.12: Overall distribution of the light positions in reference good acquisition (left) and NBLP acquisition (right) of canvas painting at elevation 45˙

	Positive	Negative	
Positive	12	49	Sensitivity 0.19
Negative	36	263	Specificity 0.87
	Precision 0.25	Negative predictive value 0.84	Accuracy 0.76

Table 5.6: Confusion matrix of the results for canvas painting at elevation 45˙

For a diffuse point, every direction carries approximately the same weights. In this example, we see that the metrics scores are lower. The lower score is due to the choice of initial LPs. An initial LP that might be closer to the reference LPs, will clearly have a better metric score. However, the metrics are not that relevant to the diffuse surface as much as for specular and semi specular surfaces. Since any set of uniformly distributed light positions are good to characterize the diffuse points and are comparably valid.

Summary We tested our method on different surfaces and obtained results for ring acquisitions at different elevations. Figure.5.13 are the plots of the measured metrics of the implemented method. Our observations show that the method is adaptive relatively to the surface inspected, with an overall accuracy of 70% to 85% in classifying azimuth space into important and non-critical directions. However, the precision in identifying good light directions is relatively lower, particularly for diffuse surfaces such as canvas paintings where reflectance is nearly uniform. The accuracy in finding the good light direction is also relatively low at raking angles and higher elevations due to the effect of zero-intensity pixels and pixel saturation, respectively. This is a limitation of our method.

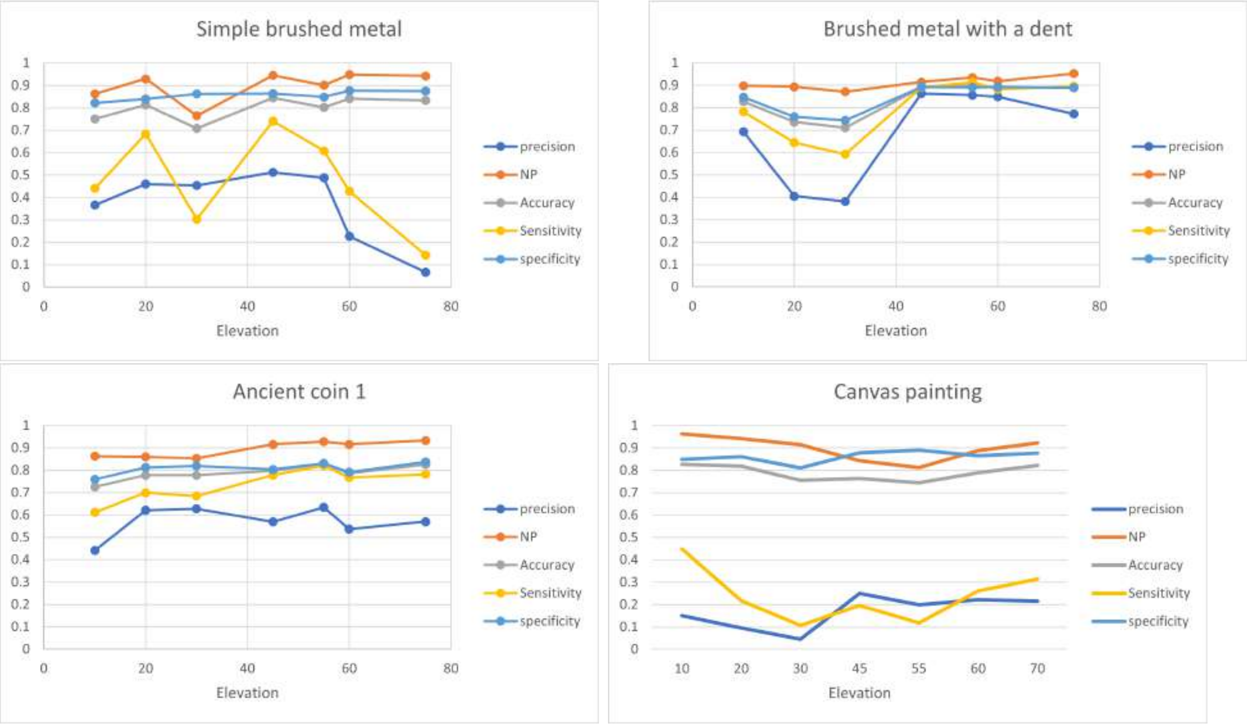


Figure 5.13: Overall perfomance of the implemented NBLP method on the surfaces at different elevations

5.2.3 Optimizing light directions in Azimuthal-Elevation space

To extend our method from azimuthal-only space to azimuthal-elevation space, two critical challenges must be addressed.

1. The first challenge is the higher non-linearity in change in luminance due to change in azimuth compared to the nearly linear change due to change in elevation. The luminance of a point changes more dramatically in elevation space due to the placement of the camera directly above the surface, increasing the probability of light rays hitting the surface and bouncing to the camera at higher elevations (For example, as seen in Figure.5.15, where the signal peaks rapidly towards $(0, 0)$ - the apex of the hemisphere where the elevation is $90^{\cdot}$). Therefore, it is necessary to account for this difference in the influence of azimuth and elevation on luminance gradients.

   To account for the difference in the influence of azimuth and elevation on luminance gradients we apply the following approach:

   In the pre-initialization step, we illuminate the surface at different elevation but keeping the azimuth same. Let L_i be the luminance value of pixel i in image P_j, where j represents the index of the elevations. Let ϕ_j be the elevation corresponding to image P_j. We fit a linear curve to L_i as a function of ϕ_j for each pixel i as illustrated in Figure.5.14 and Eq.5.6

$$L_i = (a_i \times \phi_j) + b_i \tag{5.6}$$

where a_i is the slope of the linear curve for pixel i, representing the rate of change of luminance with respect to elevation. This means that for each pixel, we can calculate how much its luminance changes as the elevation. This pre-initialization is highly limited by the computational cost. Although its a simple linear fitting of signals at less than 10 light positions, simultaneously doing it for the every pixel is computationally very expensive. Hence we need to be careful in choosing the number of elevations in the pre initialization step. For subsequently acquired light position with elevation ϕ_k, we can approximate the change in luminance due to a change in elevation as:

$$\delta L_i = a_i \times (\phi_k - \phi_j) \tag{5.7}$$

This equation allows us to approximate how much the luminance of each pixel will change when the elevation changes from ϕ_j to ϕ_k. We then subtract this value from the actual difference in luminance values to account for the difference in the influence of azimuth and elevation on luminance gradients. In other words, this method allows us to correct for the effect of changes in elevation on the luminance values of each pixel, so that azimuths and elevation are treated equally in gradients calculations. We may consider fitting the initial signal to cosineous instead of linear fitting following the Lambert's cosine law. However, there are always a significant number of outliers due to various problems like saturations, shadowing, masking etc. Hence fitting line is simpler and more closer to the signal.

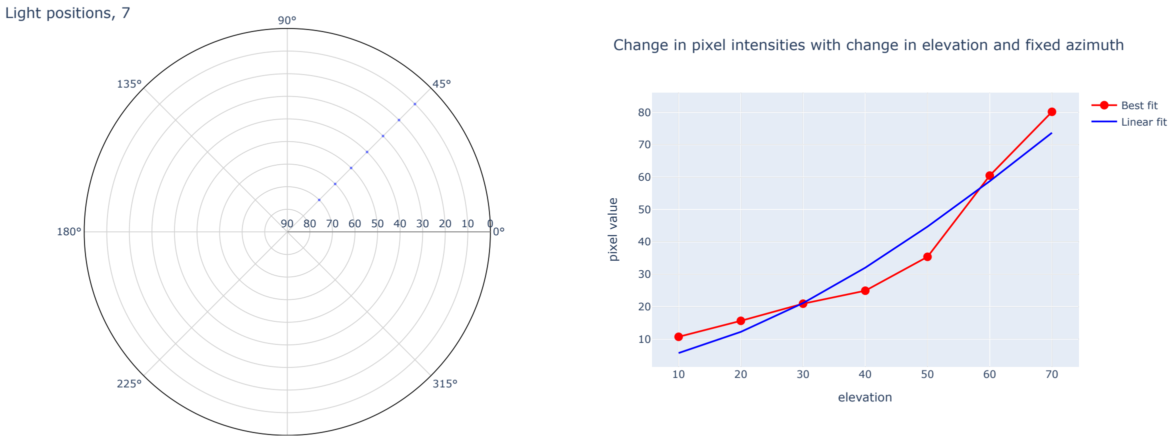
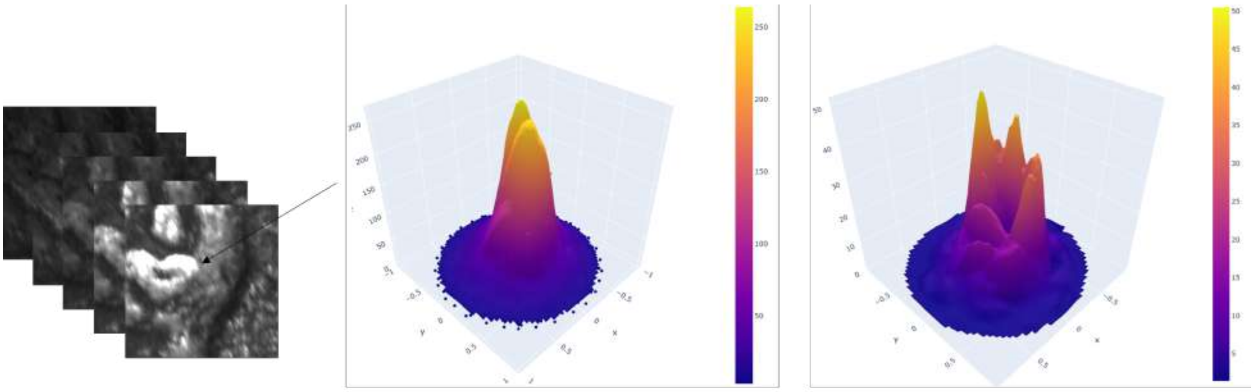


Figure 5.14: (Left) Light source positioned at different elevation keeping the azimuth same at 45°

2. The second challenge is the computational complexity of calculating the gradients of a large number of signals simultaneously in azimuthal-elevation space. In azimuthal-only space, where the signal is sorted according to azimuths and already on a regularized grid, calculating gradients is relatively straightforward. However, in azimuthal-elevation space, it is necessary to extrapolate or interpolate signal values using polynomial fitting Turley [2018] in order to form a regularized grid for computing gradients as illustrated in Figure.5.15. Doing this for a very large number of signals is not feasible.



Acquisition stack Acquired pixel signal Gradient of the signal

Figure 5.15: Example of an acquisition in the azimuthal-elevation space. The signal plot shows the value of the indicated pixel when the surface is illuminated from various directions. The gradient plot illustrates the gradients of the pixel computed from the fitted curve.

To overcome this challenge, we estimate the gradient of a pixel p_i at a light position d_j by comparing the pixel value with that at light positions closest to d_j. We construct a KD tree Ram and Sinha [2019] of all the light directions (x,y,z from azimuths and elevations). Using the KD tree, we find k nearest neighbors for each light position. Then we find the gradient of a surface point p_i at direction d_j as the average of the gradients between the luminance of p_i when it is illuminated by light at the direction d_j and its value when it is illuminated by a light at the neighboring directions.

Let's say for a surface point p_i illuminated by a light at direction d_j we have k signal values s_{il} for light directions $l = 1$ to k neighbours, corresponding to (θ_l, ϕ_l) directions. The gradient between s^l_{ij} and the signal s_{ij} is

$$\nabla s^l_{ij} = \frac{|s_{il} - s_{ij}|}{\sqrt{(\theta_l - \theta_j)^2 + (\phi_l - \phi_j)^2}}, \text{ for } l = 1 \text{ to } k. \tag{5.8}$$

The magnitude of the average gradient of p_i at light direction d_j is then calculated as,

$$g_{ij} = \frac{\sum_{l=1}^{k} \nabla s^l_{ij}}{k}, \tag{5.9}$$

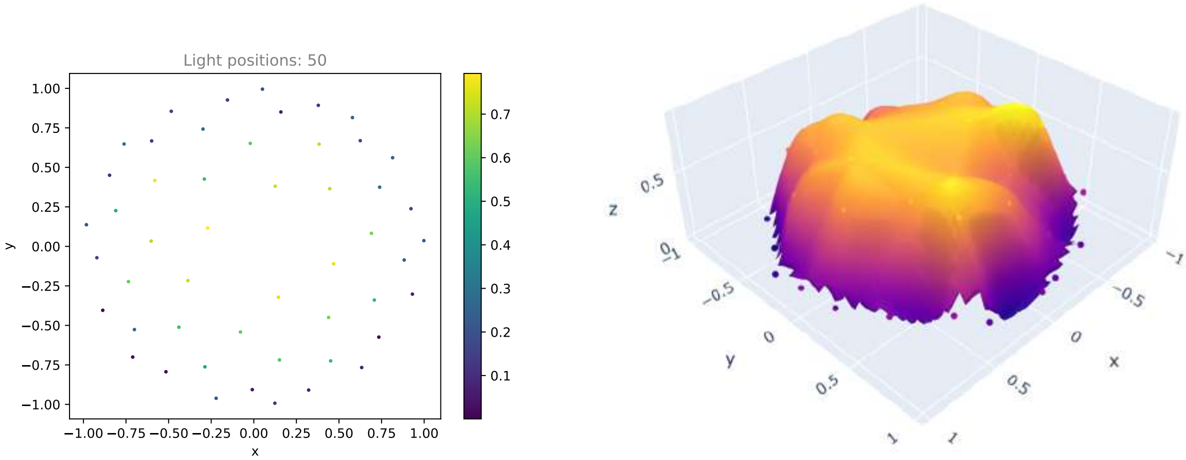
The resultant direction is calculated as the weighted average of the (θ, ϕ) values, where the weights are gradients ∇s^l_{il}. The weighted average would be,

$$\text{Resultant direction, } \hat{g_{ij}} = \left(\frac{\sum_{l=1}^{k} \nabla s^l_{ij} \theta_l}{\sum_{l=1}^{k} \nabla s^l_{ij}}, \frac{\sum_{l=1}^{k} \nabla s^l_{ij} \phi_l}{\sum_{l=1}^{k} \nabla s^l_{ij}} \right) \tag{5.10}$$

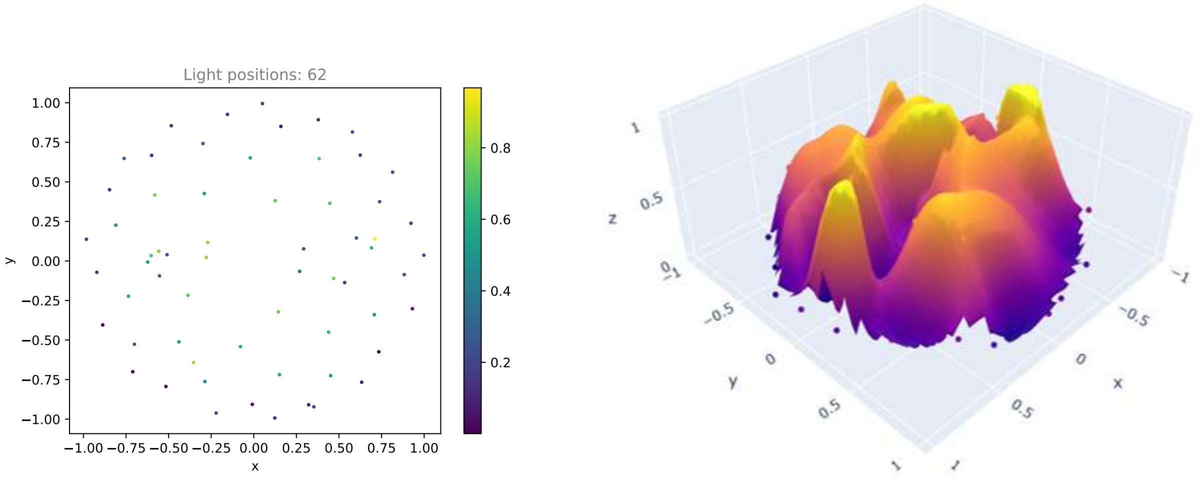
We use the gradient based identification of next best light positions approach for the azimuthal-elevation space similar to the azimuth case discussed in the previous section. We use the same steps as before while the correction of gradients to account for the difference in luminance change with azimuth and elevation is carried out before the computation of gradient score.

Results We implemented our method and tested it on the 4 surfaces used as before. We present the results by showing the light positions through the iterations and the respective gradient scores in the form of a 3D surface plot. The images on the left represents the distribution of the light positions projected on a 2D plane and the images on the right represents the normalized gradient scores at the current iteration. Although it is possible to quantify the results using a confusion matrix metric similar to the one used before, analyzing the azimuth-elevation space is more complex due to its two-dimensional grid and the computational cost of identifying the filling of the grid. Therefore, visual analysis of the distribution of the NBLP (Next-Best Light Position) light positions compared to the reference best light positions is performed by plotting the light positions on a 2D plane. To present the results concisely, in this section, we present the evolution of NBLP azimuth-elevation acquisition for the simple brushed metal surface and for the other three surfaces we present just the initial and final acquisition and the comparison of NBLP acquistion points with that of the reference good acquisition points. We have shown more acquisitions in the appendix for reference.

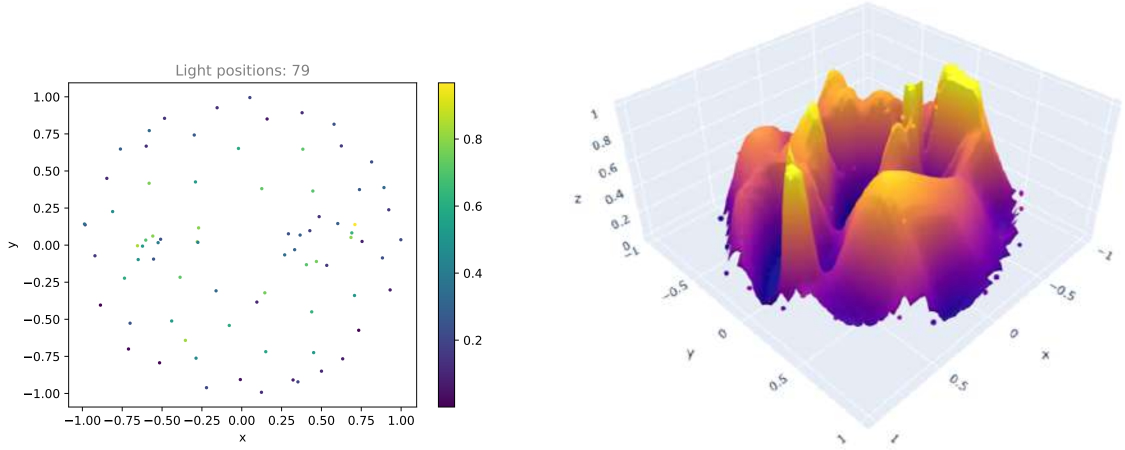
Simple brushed metal Figure.5.16 shows the NBLP acquisition of the simple brushed metal surface. The acquisition terminated after 3 iterations because all gradients fell below the threshold value of 10. The method identifies more critical light positions close to the central x direction than along the y direction, which is the same as the brush direction.



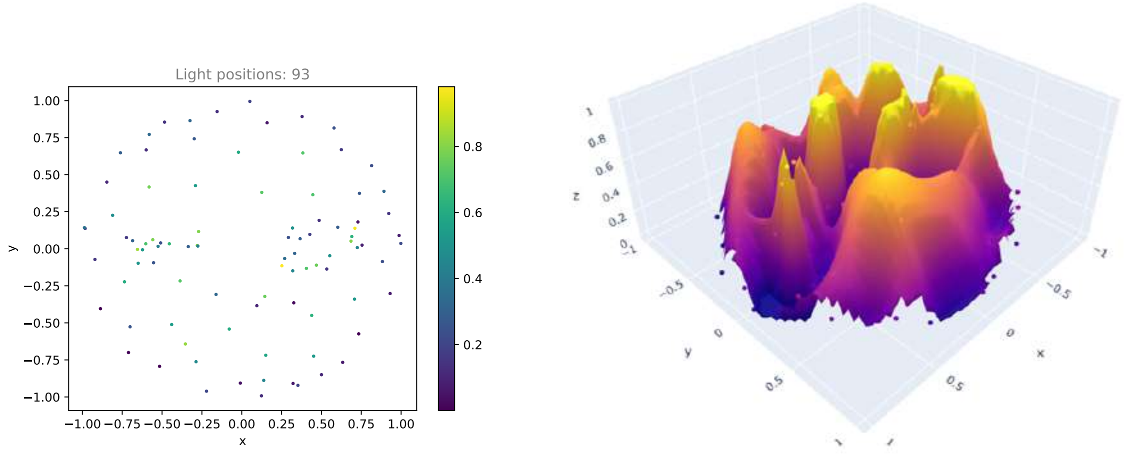
(a) Initial acquisition. (Left) Light positions, (Right) Normalized gradient scores



(b) 1^{st} iteration. (Left) Light positions, (Right) Normalized gradient scores



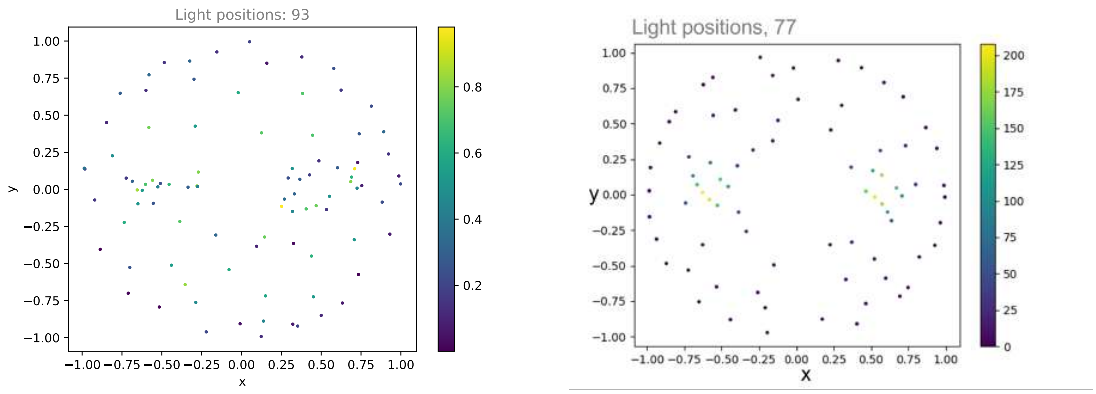
(c) 2^{nd} iteration. (Left) Light positions, (Right) Normalized gradient scores



(d) 3^{rd} iteration. (Left) Light positions, (Right) Normalized gradient scores

Figure 5.16: Azimuthal-elevation space NBLP acquisition of the simple brushed metal surface

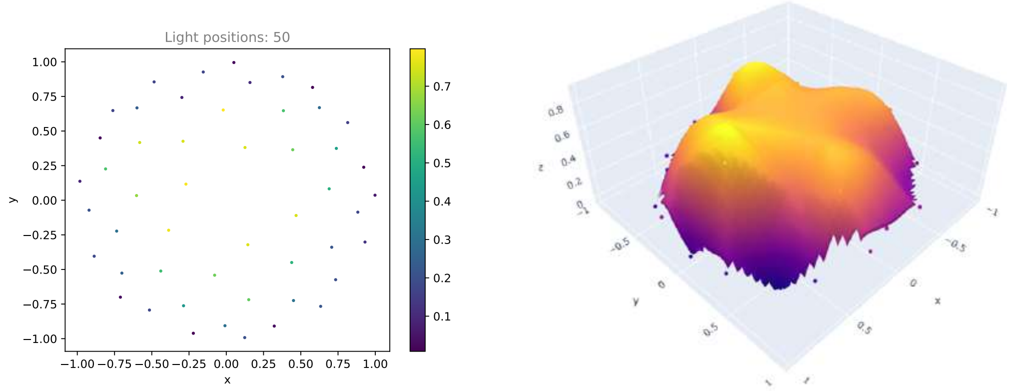
Figure.5.17 compares the light positions estimated by the implemented NBLP method with the best light positions estimated from dense acquisition in Chapter. 4, Figure.4.18. The redundant light positions estimated by the NBLP method as observed in the figure are not desired and require further optimization and improvement in the approach.



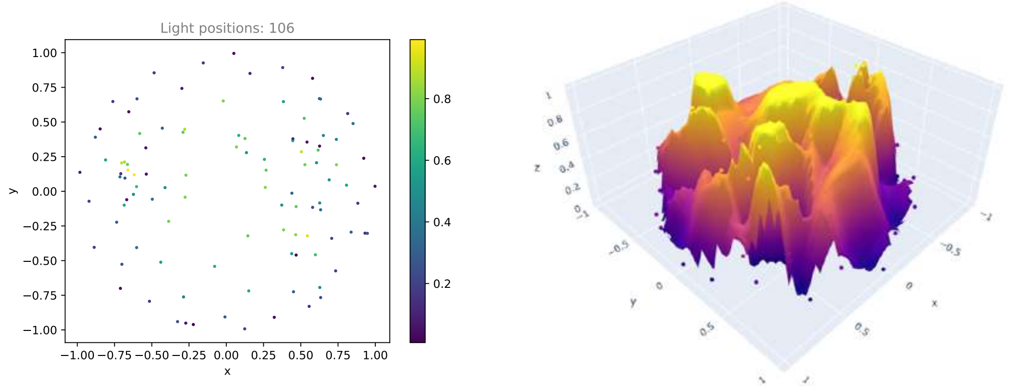
(a) NBLP estimated light position (b) Best light positions derived in chapter. 4

Figure 5.17: Comparison of NBLP estimated light positions with that of the best light positions estimated from dense acquisition in chapter. 4 for the simple brushed metal surface.

Brushed metal with dent Figure.5.18 shows the initial and terminal steps in the NBLP acquisition of the brushed metal surface with a dent. The acquisition terminated after 3 iterations as well. The method adapts the light positions to capture the details around the defect (dent) on the brushed metal surface.



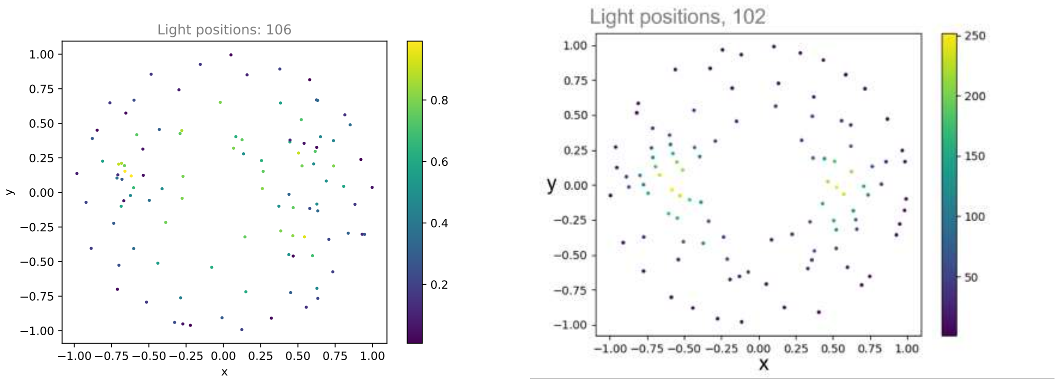
(a) Initial acquisition



(b) 3^{rd} iteration. (Left) Light positions, (Right) Normalized gradient scores

Figure 5.18: Azimuthal-elevation space NBLP acquisition of the brushed metal with a dent surface

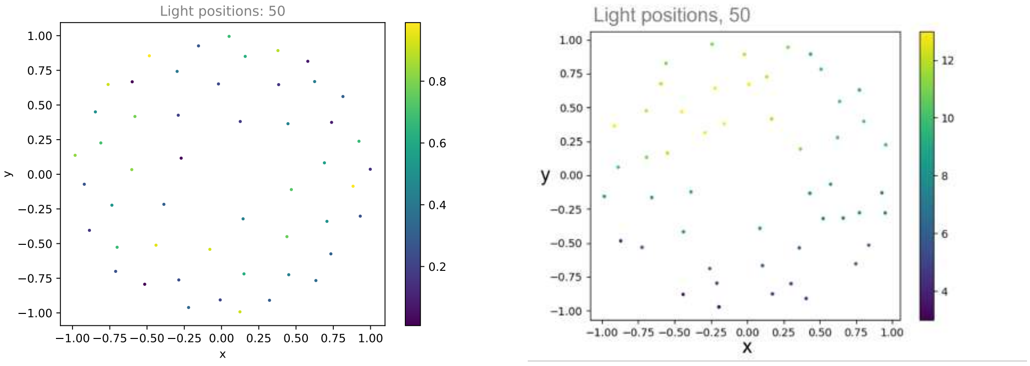
Figure.5.19 compares the light positions estimated by the implemented NBLP method with the best light positions estimated from dense acquisition in Chapter. 4, Figure.4.18. As seen, the light positions are spread such that it is in general towards the good light directions but still far from achieving the same result as the reference good light directions.



(a) NBLP estimated light position (b) Best light positions derived in chapter. 4

Figure 5.19: Comparison of NBLP estimated light positions with that of the best light positions estimated from dense acquisition in chapter. 4 for the brushed metal surface with a dent.

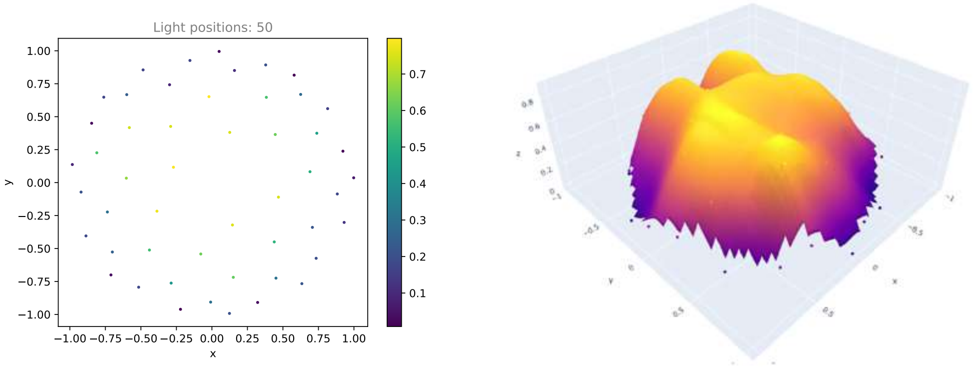
Canvas painting Figure.5.20 shows the NBLP acquisition of the canvas painting surface in the rti-dataset. The acquisition did not proceed after the initial acquisition since none of the gradients were above the threshold value of 10. The surface is diffuse and hence the light rays gets reflected uniformly at all directions. This results in very low luminance gradient values.



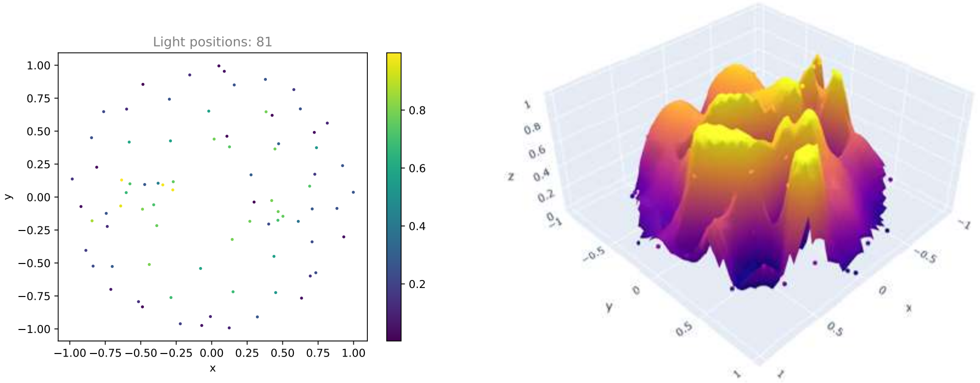
(a) NBLP estimated light position (b) Reference good light positions

Figure 5.20: NBLP estimated light positions and good reference light positions for the canvas painting surface.

Ancient coin 1 Figure.5.21 shows the NBLP acquisition of the ancient coin1 surface in the rti-dataset. The acquisition terminated after 6 iterations. The method progresses slowly in this case, because there are points with a range of reflectance properties unlike brushed metal (anisotropic) or canvas painting (diffuse).



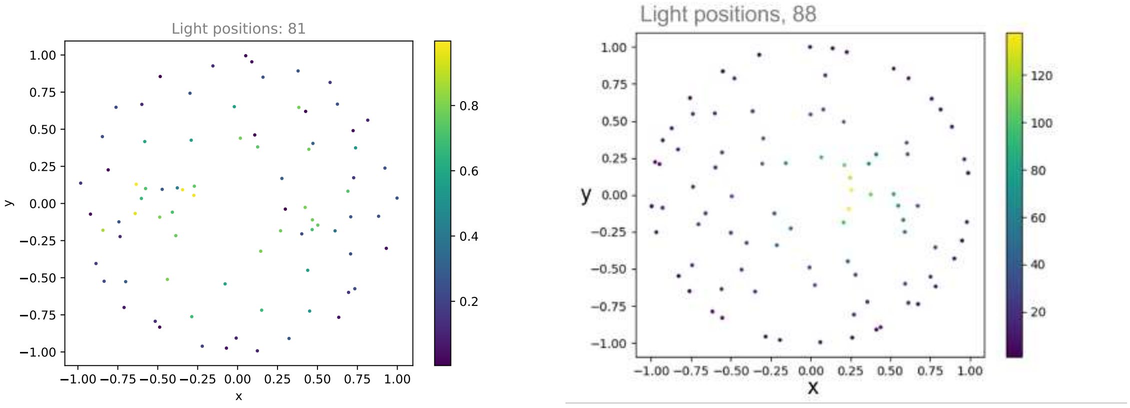
(a) Initial acquisition



(b) 6^{th} iteration (termination). (Left) Light positions, (Right) Normalized gradient scores

Figure 5.21: NBLP iterations in acquisition of the ancient coin1 surface in the rti-dataset

Figure.5.22 compares the NBLP acquisition result with the reference good acquisition light positions for the ancient coin1 surface in the rti-dataset. The NBLP resulted in undersampling for this surface, unlike the redundant sampling for the brushed metal surface. The approximation of linearity in elevation space for correction significantly affects the NBLP estimation for this surface. This approximation suppresses the masking and shadowing of points from it 3D geometry, which can hinder the identification of certain critical light directions.



(a) NBLP estimated light position

(b) Reference good light positions

Figure 5.22: NBLP estimated light positions and good reference light positions for the ancient coin1 surface.

5.3 Conclusion

In conclusion, this chapter presented new strategies for RTI acquisition that estimate good light positions as the acquisition progresses, adapting to the surface being captured. Our approach to solve the Next Best Light Position (NBLP) problem in RTI acquisitions uses gradients as a measure of the influence of a light direction in revealing surface details. We introduced an evaluation metric to quantitatively assess the performance of our proposed NBLP methods. For ring acquisition (azimuth-only space), the average accuracy of NBLP compared to the reference good acquisition derived in the rti-dataset is around 80%, which is reasonable. In the azimuthal-elevation space, our method showed good adaptability to surfaces - for example in case of canvas paintings (diffuse) surface, it terminated after the initial acquisition of 50 light positions and for complex surfaces like the brushed metal and ancient coin surfaces, the acquisition progressed to more number of iterations and acquired more light directions. For the ancient coin surface, it identifies most of the light directions that reveal more details on the surface. However, there is still room for improvement and this problem is to be studied and improved in future works. Overall, this chapter aimed to address the challenges associated with capturing high-quality RTI data for objects with complex shapes and uneven surfaces.

Conclusion and future work

Reflectance imaging is a powerful technique that can be used to capture and reconstruct the surface properties of objects. It has a wide range of applications in cultural heritage and art conservation, including: 3D documentation of objects, surface analysis, damage assessment, restoration planning, virtual reconstruction.

In this thesis, we have presented a number of contributions to the field of reflectance imaging. These contributions include:

- The development of a robotic system for automated and adaptive RTI acquisitions
- The creation of a benchmark dataset for evaluating RTI acquisition strategies
- The development of new methods for improving the quality of RTI data

The robotic system that we developed, called the LightBot, is a state-of-the-art system that can be used to automate and improve the efficiency of RTI acquisitions. The LightBot is equipped with a camera, a light source, a robotic arm and a XY platform. The robotic arm is used to move the light source around the object and the XY platform to translate the surface. The LightBot can be used to acquire RTI data of objects of range of size or shape. We created an API of the robotic acquisition system that enable other researchers to easily use to test and develop strategies related to acquisition of reflectance transformation imaging. We used LightBot to develop stitching of RTI data of medium sized objects, thereby demonstrating the advantage of the system we built. We strongly believe that LightBot will open up new perspective of research in the field of RTI. We believe that LightBot will further the extent of applying RTI technique in the cultural heritage field.

Next, we created a benchmark dataset for evaluating RTI acquisition strategies. We recognized the need for a dataset and corresponding reference best light positions to strategize and evaluate the quality of RTI acquisitions. We created such a benchmark dataset consisting of RTI acquisition data of 20 different (13 virtual and 7 real). We developed methods to identify the best light positions to perform RTI acquisition adaptive to a given surface. This reference best light positions are derived from a very dense acquisition of the surfaces. They are considered as the benchmark acquisitions for evaluating the quality of a RTI acquisition performed on the surfaces. We presented methods to derive best light positions for azimuth only space and azimuth-elevation space. The benchmark dataset can be used to quantitatively evaluate the performance of different RTI acquisition strategies in terms of accuracy, efficiency. We are publishing this benchmark dataset as opensource for researchers to use.

Finally, we developed new methods for improving the quality of RTI data. These methods include:

☞ *Methods for correcting the illumination in photometric stereo data and perform stitching of RTI data.* To reveal the details of surface and to gather as much information as possible from the surface effectively, RTI acquisitions are usually performed at high resolutions. As the surfaces being acquired are larger, the working distance of the camera from the surface also gets larger to cover the whole surface. This results in the reduction in the number of pixels per unit area of the surface, thereby affecting the resolution of the capture. To overcome this issue, the acquisition of the surface is broken down to multiple acquisitions and then stitched together. We presented methods to do this that involves illumination correction, alignment and stitching of images. We implemented the methods and presented the results. This is a significant progress towards improving the quality of the RTI acquisition.

☞ *Identifying optimal light directions in performing RTI acquisition.* Our work is the first extensive study and attempt to find the optimal light directions in performing RTI acquisitions. We defined this problem of identifying the best light positions as Next Best Light Position problem analogous to the Next Best View problem in 3D imaging. The objective of addressing this problem is to achieve and ensure higher acquisition quality, reduce the number of light positions in an acquisition as much as possible, optimize the acquisition time and to achieve reproducibility of RTI of surfaces. We presented novel methods to identify optimal light positions during the acquisition process and also presented metrics to quantitatively evaluate their performance comparing with the benchmark dataset.

The methods that we developed have been shown to improve the quality of RTI data by reducing noise and artifacts. These methods can be used to improve the accuracy of RTI data, which can be used for a variety of applications in cultural heritage and art conservation. We believe that the contributions that we have made in this thesis will advance the field of reflectance imaging and make it a more powerful tool for cultural heritage and art conservation.

Publications

1. Luxman, R., Nurit, M., Goïc, G. L., Marzani, F., & Mansouri, A. (2021). Next Best Light Position: A self configuring approach for the Reflectance Transformation Imaging acquisition process. Electronic Imaging, 2021(5), 132-1.

2. Luxman, R., Castro, Y. E., Chatoux, H., Nurit, M., Siatou, A., Le Goïc, G., Brambilla, L., Degrigny, C., Marzani, F., & Mansouri, A. (2022). LightBot: A Multi-Light Position Robotic Acquisition System for Adaptive Capturing of Cultural Heritage Surfaces. Journal of Imaging, 8(5), 134.

3. Luxman, R., Chatoux, H. , Goïc, G. L., Hardeberg, Jon Y, Marzani, F., & Mansouri, A. (2023). A benchmark dataset and evaluation for best light configuration in Reflectance Transformation Imaging. Archiving Conference. Vol. 20. Society for Imaging Science and Technology, 2023.

4. Castro, Y., Siatou, A., Rossé, M., Chatoux, H., Luxman, R., Le Goïc, G., & Mansouri, A. (2022). Extended framework for Multispectral RTI. In Archiving Conference (Vol. 19, pp. 56-61).

Future works

We believe that the optimizations in reflectance transformation imaging technology is promising. We believe that RTI will continue to advance and become a more powerful tool for cultural heritage and art conservation.

There are a number of directions for future work in the field of reflectance transformation imaging. These directions include:

☞ LightBot is a significant step towards the automation of RTI acquisition. However, there are still some challenges that need to be addressed in the future, such as the automatic calibration of the system.

☞ We initiated the investigation on the search for optimal light directions for performing surface adaptive RTI acquisition. This is a topic that requires further study in the future.

☞ Artificial intelligence (AI), particularly in the field of image processing, is rapidly advancing. It would be very interesting to exploit AI to various RTI related optimizations. For example, using AI to optimize the light directions to perform RTI adaptive to surface being inspected is a topic that needs to be explored. Aslo AI could also be effectively utilized for RTI data anamoly detection in surfaces with RTI imaging.

☞ We have demonstrated the results of our methods for carrying out light corrections. However, there is still room for improvement, and this is a topic that will be studied in the future.

☞ Another area of research in RTI is the use of Multispectral lighting that would enable the acquisition of complex surfaces (multi-material) and a more detailed analysis of surfaces by adapting the wavelength to the desired surface state characteristics. For example, the study of how a surface's color can be used to indicate changes or as an indicator of heterogeneity in the materials that make up the object's surface.

☞ There are still many areas of research that need to be explored in RTI such as High Dynamic Range RTI, integrating focus variation methods in RTI.

Bibliography

Automata tech. URL `https://automata.tech/`.

Broncolor rti. URL `https://broncolor.swiss/fr/produits/scope-d50?variant=6147`.

Kinova. URL `https://assistive.kinovarobotics.com/product/jaco-robotic-arm`.

Henrik Aanæs, Anders Lindbjerg Dahl, and Kim Steenstrup Pedersen. Interesting interest points. *International Journal of Computer Vision*, 97(1):18–35, 2012.

Jens Ackermann, Michael Goesele, et al. A survey of photometric stereo techniques. *Foundations and Trends® in Computer Graphics and Vision*, 9(3-4):149–254, 2015.

Ebtsam Adel, Mohammed Elmogy, and Hazem Elbakry. Image stitching based on feature extraction techniques: a survey. *International Journal of Computer Applications*, 99(6):1–8, 2014.

Edward H Adelson, Charles H Anderson, James R Bergen, Peter J Burt, and Joan M Ogden. Pyramid methods in image processing. *RCA engineer*, 29(6):33–41, 1984.

John Amanatides. Realism in computer graphics: A survey. *IEEE Computer Graphics and Applications*, 7(1):44–56, 1987.

Vasileios Argyriou, Stefanos Zafeiriou, Barbara Villarini, and Maria Petrou. A sparse representation method for determining the optimal illumination directions in photometric stereo. *Signal Processing*, 93(11):3027–3038, 2013.

Frederick O Bartell, Eustace L Dereniak, and William L Wolfe. The theory and measurement of bidirectional reflectance distribution function (brdf) and bidirectional transmittance distribution function (btdf). In *Radiation scattering in optical systems*, volume 257, pages 154–160. SPIE, 1981.

Herbert Bay, Tinne Tuytelaars, and Luc Van Gool. Surf: Speeded up robust features. In *Computer Vision–ECCV 2006: 9th European Conference on Computer Vision, Graz, Austria, May 7-13, 2006. Proceedings, Part I 9*, pages 404–417. Springer, 2006.

Khrystyna Berladir, T Hovorun, I Pererva, G Rudenko, and I Martynov. Modern materials for automotive industry. *Journal of engineering sciences*, 4:F8–F18, 12 2017. doi: 10.21272/jes.2017.4(2).f8.

Michael S Bessell et al. Standard photometric systems. *Annual Review of Astronomy and Astrophysics*, 43(1):293–336, 2005.

Alex Byrne and David R Hilbert. Color realism and color science. *Behavioral and brain sciences*, 26(1):3–21, 2003.

Kun Chen, Jingjing Wen, Jing Wu, and Ze Ji. Automated robot-based large-scale 3d surface imaging. *Procedia Computer Science*, 176:2949–2958, 2020.

Yi-Cheng Chung, Pi-Ju Cheng, Yu-Hsun Chou, Bo-Tsun Chou, Kuo-Bin Hong, Jheng-Hong Shih, Sheng-Di Lin, Tien-Chang Lu, and Tzy-Rong Lin. Surface roughness effects on aluminium-based ultraviolet plasmonic nanolasers. *Scientific reports*, 7(1):1–9, 2017.

FJJ Clarke and DJ Parry. Helmholtz reciprocity: Its validity and application to reflectometry. *Lighting Research & Technology*, 17(1):1–11, 1985.

D Coleman. Integrating ikfast with moveit!: A tutorial, 2014.

Robert L Cook and Kenneth E. Torrance. A reflectance model for computer graphics. *ACM Transactions on Graphics (ToG)*, 1(1):7–24, 1982.

Copa-Hypotheses. Rti, 2019. URL `https://copa.hypotheses.org/category/3d-surface-imaging/rti`. Accessed January 04, 2023.

HE Coules, PJ Orrock, and C Er Seow. Reflectance transformation imaging as a tool for engineering failure analysis. *Engineering Failure Analysis*, 105:1006–1017, 2019.

Kristin J Dana, Bram Van Ginneken, Shree K Nayar, and Jan J Koenderink. Reflectance and texture of real-world surfaces. *ACM Transactions On Graphics (TOG)*, 18(1):1–34, 1999.

Rae David F, Earnshaw. State of the art in computer graphics: visualization and modeling. *Springer Science*, 1991.

Paul E Debevec and Jitendra Malik. Recovering high dynamic range radiance maps from photographs. In *ACM SIGGRAPH 2008 classes*, pages 1–10. 2008.

Christian Degrigny, V Argyropoulos, P Pouli, M Grech, K Kreislova, M Harith, F Mirambet, N Haddad, E Angelini, E Cano, et al. The methodology for the promet project to develop/test new non-toxic corrosion inhibitors and coatings for iron and copper alloy objects housed in mediterranean museums. In *Metal 07. Proceedings of the Interim meeting of the ICOM-CC Metal WG*, volume 5, pages 31–37, 2007.

Matteo Dellepiane, Massimiliano Corsini, Marco Callieri, and Roberto Scopigno. High quality ptm acquisition: Reflection transformation imaging for large objects. In *VAST*, pages 179–186, 2006.

Rosen Diankov. Ikfast: The robot kinematics compiler. *URL here*, 2010.

Rosen Diankov and James Kuffner. Openrave: A planning architecture for autonomous robotics. *Robotics Institute, Pittsburgh, PA, Tech. Rep. CMU-RI-TR-08-34*, 79, 2008.

Ondrej Drbohlav and Mike Chantler. On optimal light configurations in photometric stereo. In *Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1*, volume 2, pages 1707–1712. IEEE, 2005.

Aaron D'Souza, Sethu Vijayakumar, and Stefan Schaal. Learning inverse kinematics. In *Proceedings 2001 IEEE/RSJ International Conference on Intelligent Robots and Systems. Expanding the Societal Role of Robotics in the the Next Millennium (Cat. No. 01CH37180)*, volume 1, pages 298–303. IEEE, 2001.

Tinsae G Dulecha, Filippo A Fanni, Federico Ponchio, Fabio Pellacini, and Andrea Giachetti. Neural reflectance transformation imaging. *The Visual Computer*, 36(10):2161–2174, 2020.

Tinsae Gebrechristos Dulecha. *Surface Analysis and Visualization from Multi-Light Image Collections*. PhD thesis, University of Verona, 2021.

Graeme Earl, Philip Basford, Alexander Bischoff, Alan Bowman, Charles Crowther, Jacob Dahl, Michael Hodgson, Leif Isaksen, Eleni Kotoula, Kirk Martinez, et al. Reflectance transformation imaging systems for ancient documentary artefacts. *Electronic Visualisation and the Arts (EVA 2011)*, pages 147–154, 2011.

Aarthipoornima Elangovan and T Jeyaseelan. Medical imaging modalities: a survey. In *2016 International Conference on emerging trends in engineering, technology and science (ICETETS)*, pages 1–4. ieee, 2016.

Georgios D Evangelidis and Emmanouil Z Psarakis. Parametric image alignment using enhanced correlation coefficient maximization. *IEEE transactions on pattern analysis and machine intelligence*, 30(10):1858–1865, 2008.

Mark Fiala. Artag revision 1, a fiducial marker system using digital techniques. *National Research Council Publication*, 47419:1–47, 2004.

Olga Filipova. *Learning Vue. js 2*. Packt Publishing Ltd, 2016.

Martin A Fischler and Robert C Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, 1981.

Ronald Aylmer Fisher et al. 012: A mathematical examination of the methods of determining the accuracy of an observation by the mean error, and by the mean square error. 1920.

Roland W Fleming, Ron O Dror, and Edward H Adelson. Real-world illumination and the perception of surface reflectance properties. *Journal of vision*, 3(5):3–3, 2003.

Silvia Florindi, Anna Revedin, Biancamaria Aranguren, and Vincenzo Palleschi. Application of reflectance transformation imaging to experimental archaeology studies. *Heritage*, 3(4):1279–1286, 2020.

Museum für Kunst und Gewerbe Hamburg. An egyptian tomb stone. `https://skfb.ly/oBXSH`, 2021. Accessed: October 10, 2022.

Hamza Gardi, Sebastian F Walter, and Christoph S Garbe. An optimal experimental design approach for light configurations in photometric stereo. *arXiv preprint arXiv:2204.05218*, 2022.

Michael Garland and Paul S Heckbert. Surface simplification using quadric error metrics. In *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, pages 209–216, 1997.

Pascal Gautron, Jaroslav Krivánek, Sumanta N Pattanaik, and Kadi Bouatouch. A novel hemispherical basis for accurate and efficient rendering. *Rendering Techniques*, 2004:321–330, 2004.

Jason Geng. Structured-light 3d surface imaging: a tutorial. *Advances in Optics and Photonics*, 3 (2):128–160, 2011.

Andrea Giachetti, Irina Ciortan, Claudia Daffara, Ruggero Pintus, Enrico Gobbetti, et al. Multispectral rti analysis of heterogeneous artworks. 2017.

Andrew S Glassner. *An introduction to ray tracing*. Morgan Kaufmann, 1989.

A Ardeshir Goshtasby. *Image registration: Principles, tools and methods*. Springer Science & Business Media, 2012.

Megandhren Govender, Kershani Chetty, and Hartley Bulcock. A review of hyperspectral remote sensing and its application in vegetation and water resource studies. *Water Sa*, 33(2):145–151, 2007.

Allan Grønlund, Kasper Green Larsen, Alexander Mathiasen, Jesper Sindahl Nielsen, Stefan Schneider, and Mingzhou Song. Fast exact k-means, k-medians and bregman divergence clustering in 1d. *arXiv preprint arXiv:1701.07204*, 2017.

Grant Haskins, Uwe Kruger, and Pingkun Yan. Deep learning in medical image registration: a survey. *Machine Vision and Applications*, 31(1):1–18, 2020.

Mathieu Hébert, Roger D Hersch, and Patrick Emmel. Fundamentals of optics and radiometry for color reproduction. *Handbook of Digital Imaging*, pages 1–57, 2015.

David Holt, Olugbenga Okusanya, Ryan Judy, Ollin Venegas, Jack Jiang, Elizabeth DeJesus, Evgeniy Eruslanov, Jon Quatromoni, Pratik Bhojnagarwala, Charuhas Deshpande, et al. Intraoperative near-infrared imaging can distinguish cancer from normal tissue but not inflammation. *PLoS One*, 9(7):e103342, 2014.

Radu Horaud and Fadi Dornaika. Hand-eye calibration. *The international journal of robotics research*, 14(3):195–210, 1995.

Xiang Huang, Marc Walton, Greg Bearman, and Oliver Cossairt. Near light correction for image relighting and 3d shape recovery. In *2015 Digital Heritage*, volume 1, pages 215–222. IEEE, 2015.

Xiang Huang, Erich Uffelman, Oliver Cossairt, Marc Walton, and Aggelos K Katsaggelos. Computational imaging for cultural heritage: Recent developments in spectral imaging, 3-d surface measurement, image relighting, and x-ray mapping. *IEEE Signal Processing Magazine*, 33(5): 130–138, 2016.

Ivo Ihrke, John Restrepo, and Lois Mignard-Debise. Principles of light field imaging: Briefly revisiting 25 years of research. *IEEE Signal Processing Magazine*, 33(5):59–69, 2016.

Cultural Heritage Imaging. Chi. URL `https://culturalheritageimaging.org/Technologies/RTI/index.html`. Accessed January 04, 2023.

Rasmus Jensen, Anders Dahl, George Vogiatzis, Engin Tola, and Henrik Aanæs. Large scale multi-view stereopsis evaluation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 406–413, 2014.

Deane B Judd. Fresnel reflection of diffusely incident light. *J. Res. Natl. Bur. Stand*, 29(5):329, 1942.

Michail Kalaitzakis, Brennan Cain, Sabrina Carroll, Anand Ambrosi, Camden Whitehead, and Nikolaos Vitzilaios. Fiducial markers for pose estimation. *Journal of Intelligent & Robotic Systems*, 101(4):1–26, 2021.

Sertac Karaman and Emilio Frazzoli. Sampling-based algorithms for optimal motion planning. *The international journal of robotics research*, 30(7):846–894, 2011.

Sertac Karaman, Matthew R Walter, Alejandro Perez, Emilio Frazzoli, and Seth Teller. Anytime motion planning using the rrt. In *2011 IEEE International Conference on Robotics and Automation*, pages 1478–1483. IEEE, 2011.

Maciej Karaszewski, Marcin Adamczyk, and Robert Sitnik. Assessment of next-best-view algorithms performance with various 3d scanners and manipulator. *ISPRS Journal of Photogrammetry and Remote Sensing*, 119:320–333, 2016.

Ted Kinsman. An easy to build reflectance transformation imaging (rti) system. *Journal of Biocommunication*, 40(1), 2016.

Sean J Kirkpatrick. A primer on radiometry. *Dental Materials*, 21(1):21–26, 2005.

Vlado Kitanovski and Jon Yngve Hardeberg. Objective evaluation of relighting models on translucent materials from multispectral rti images. 2021.

WJ Knox. Shading models for realistic image synthesis. In *Proceedings of the IEEE National Aerospace and Electronics Conference*, pages 596–603. IEEE, 1989.

Sanjeev J Koppal. Lambertian reflectance. *Computer vision: a reference guide*, pages 1–3, 2020.

Anis Koubâa et al. *Robot Operating System (ROS).*, volume 1. Springer, 2017.

Vít Krátkỳ, Pavel Petráček, Vojtěch Spurnỳ, and Martin Saska. Autonomous reflectance transformation imaging by a team of unmanned aerial vehicles. *IEEE Robotics and Automation Letters*, 5(2):2302–2309, 2020.

Serdar Kucuk and Zafer Bingul. *Robot kinematics: Forward and inverse kinematics.* INTECH Open Access Publisher London, UK, 2006.

Murat Kurt and Dave Edwards. A survey of brdf models for computer graphics. *ACM SIGGRAPH Computer Graphics*, 43(2):1–7, 2009.

Alex Kyriakidis, Kostas Maniatis, and Evan You. *The majesty of Vue. js.* Packt Publishing, 2016.

Gaëtan Le Goïc, Hermine Chatoux, Marvin Nurit, Ramamoorthy Luxman, Vincent Brost, Anthony Pamart, Livio De Luca, Alamin Mansouri, et al. A new pixel-wise illumination correction method for reflectance transformation imaging acquisition and modeling.

David Lewis, Marvin Nurit, Hermine Chatoux, Fabrice Meriaudeau, and Alamin Mansouri. An automated adaptive focus pipeline for reflectance transformation imaging. *Electronic Imaging*, 2021(18):63–1, 2021.

Hongsong Li, Sing-Choong Foo, Kenneth E Torrance, and Stephen H Westin. Automated three-axis gonioreflectometer for computer graphics applications. *Optical Engineering*, 45(4):043605, 2006.

G Lowe. Sift-the scale invariant feature transform. *Int. J*, 2(91-110):2, 2004.

David Lu. Urdf and you, 2016.

Boxin Shi Yasuyuki Matsushita Yongtian Wang Lun Wu, Arvind Ganesh and Yi Ma. Robust photometric stereo via low-rank matrix completion and recovery. In *Proceedings of Asian Conference on Computer Vision (ACCV)*, 2010.

Ramamoorthy Luxman, Yuly Emilia Castro, Hermine Chatoux, Marvin Nurit, Amalia Siatou, Gaëtan Le Goïc, Laura Brambilla, Christian Degrigny, Franck Marzani, and Alamin Mansouri. Lightbot: A multi-light position robotic acquisition system for adaptive capturing of cultural heritage surfaces. *Journal of Imaging*, 8(5):134, 2022.

Tom Malzbender, Dan Gelb, and Hans Wolters. Polynomial texture maps. *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pages 519–528, 2001a.

Tom Malzbender, Dan Gelb, and Hans Wolters. Polynomial texture maps. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pages 519–528, 2001b.

Marcello Manfredi, Greg Williamson, Dale Kronkright, Eric Doehne, Megan Jacobs, Emilio Marengo, and Gregory Bearman. Measuring changes in cultural heritage objects with reflectance transformation imaging. In *2013 Digital Heritage International Congress (DigitalHeritage)*, volume 1, pages 189–192. IEEE, 2013.

Ricardo Marques, Christian Bouville, Kadi Bouatouch, and Josep Blat. Extensible spherical fibonacci grids. *IEEE Transactions on Visualization and Computer Graphics*, 27(4):2341–2354, 2019.

Stephen R Marschner, Stephen H Westin, Eric PF Lafortune, and Kenneth E Torrance. Image-based bidirectional reflectance distribution function measurement. *Applied optics*, 39(16):2592–2600, 2000.

William Ross McCluney. *Introduction to radiometry and photometry.* Artech House, 2014.

Matthew McGuigan and Jacqueline Christmas. Automating rti: automatic light direction detection and correcting non-uniform lighting for more accurate surface normals. *Computer Vision and Image Understanding*, 192:102880, 2020.

S MONICA. Ros based 3d modeling using urdf tools. *Intelligent Systems And Communication (NCISC-2016)*, page 75, 2016.

Rosana Montes and Carlos Ureña. An overview of brdf models. *University of Grenada, Technical Report LSI-2012-001*, 2012.

Harold Mytum and JR Peterson. The application of reflectance transformation imaging (rti) in historical archaeology. *Historical Archaeology*, 52:489–503, 2018.

Christian Nansen and Norman Elliott. Remote sensing and reflectance profiling in entomology. *Annual Review of Entomology*, 61(1):139–158, 2016.

JAUWAIRIA Nasir, FAHAD Islam, and YASAR Ayaz. Adaptive rapidly-exploring-random-tree-star (rrt*)-smart: algorithm characteristics and behavior analysis in complex environments. *Asia-Pacific Journal of Information Technology and Multimedia*, 2(2):39–51, 2013.

Shree K Nayar. Lecture notes. URL `https://slideplayer.com/slide/17779333/`. Accessed January 04, 2023.

NCPTT. H-rti. URL `https://ncptt.nps.gov/blog/conservation-and-h%E2%80%90rti-at-the-metropolitan-museum-of-art/`. Accessed January 04, 2023.

Addy Ngan, Frédo Durand, and Wojciech Matusik. Experimental analysis of brdf models. *Rendering Techniques*, 2005(16th):2, 2005.

Iram Noreen, Amna Khan, and Zulfiqar Habib. Optimal path planning using rrt* based approaches: a survey and future directions. *International Journal of Advanced Computer Science and Applications*, 7(11), 2016.

Marvin Nurit, Y Castro, Abir Zendagui, Gaëtan Le Goïc, Hugues Favreliere, and Alamin Mansouri. High dynamic range reflectance transformation imaging: an adaptive multi-light approach for visual surface quality assessment. In *Fourteenth International Conference on Quality Control by Artificial Vision*, volume 11172, pages 276–284. SPIE, 2019.

Marvin Nurit, Gaëtan Le Goïc, David Lewis, Yuly Castro, Abir Zendagui, Hermine Chatoux, Hugues Favrelière, Stéphane Maniglier, Pierre Jochum, and Alamin Mansouri. Hd-rti: An adaptive multi-light imaging approach for the quality assessment of manufactured surfaces. *Computers in Industry*, 132:103500, 2021.

Cleveland Museum of Art. 1998.121 sutra container. `https://skfb.ly/oxA6o`, 2022. Accessed: October 10, 2022.

Michael Oren and Shree K Nayar. Generalization of lambert's reflectance model. In *Proceedings of the 21st annual conference on Computer graphics and interactive techniques*, pages 239–246, 1994.

Anthony Pamart, Federico Ponchio, Violette Abergel, A Alaoui M'Darhri, Massimiliano Corsini, Matteo Dellepiane, François Morlet, Roberto Scopigno, and Livio De Luca. A complete framework operating spatially-oriented rti in a 3d/2d cultural heritage documentation and analysis tool. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, 2019.

Patrick Pérez, Michel Gangnet, and Andrew Blake. Poisson image editing. In *ACM SIGGRAPH 2003 Papers*, pages 313–318. 2003.

Matt Pharr, Wenzel Jakob, and Greg Humphreys. *Physically based rendering: From theory to implementation*. Morgan Kaufmann, 2016.

Bernadette Philips-Invernizzi, Daniel Dupont, and Claude Caze. Bibliographical review for reflectance of diffusing media. *Optical Engineering*, 40(6):1082–1092, 2001.

Ruggero Pintus, Irina Ciortan, Andrea Giachetti, Enrico Gobbetti, et al. Practical free-form rti acquisition with local spot lights. 2016.

Ruggero Pintus, Tinsae Gebrechristos Dulecha, Irina Ciortan, Enrico Gobbetti, and Andrea Giachetti. State-of-the-art in multi-light image collections for surface visualization and analysis. In *Computer Graphics Forum*, volume 38, pages 909–934. Wiley Online Library, 2019.

Gilles Pitard, Gaëtan Le Goïc, Hugues Favrelière, Serge Samper, Simon-Frédéric Desage, and Maurice Pillet. Discrete modal decomposition for surface appearance modelling and rendering. In *Optical Measurement Systems for Industrial Inspection IX*, volume 9525, pages 489–498. SPIE, 2015.

Richard Pito. A solution to the next best view problem for automated surface acquisition. *IEEE Transactions on pattern analysis and machine intelligence*, 21(10):1016–1030, 1999.

Federico Ponchio, Marion Lamé, Roberto Scopigno, and Bruce Robertson. Visualizing and transcribing complex writings through rti. In *2018 IEEE 5th International Congress on Information Science and Technology (CiSt)*, pages 227–231. IEEE, 2018.

Thomas Porter and Tom Duff. Compositing digital images. In *Proceedings of the 11th annual conference on Computer graphics and interactive techniques*, pages 253–259, 1984.

Richard J Potton. Reciprocity in optics. *Reports on Progress in Physics*, 67(5):717, 2004.

Morgan Quigley, Ken Conley, Brian Gerkey, Josh Faust, Tully Foote, Jeremy Leibs, Rob Wheeler, Andrew Y Ng, et al. Ros: an open-source robot operating system. In *ICRA workshop on open source software*, volume 3, page 5. Kobe, Japan, 2009.

Morgan Quigley, Brian Gerkey, and William D Smart. *Programming Robots with ROS: a practical introduction to the Robot Operating System*. " O'Reilly Media, Inc.", 2015.

Parikshit Ram and Kaushik Sinha. Revisiting kd-tree for nearest neighbor search. In *Proceedings of the 25th acm sigkdd international conference on knowledge discovery & data mining*, pages 1378–1388, 2019.

Sandrine Rémy, Michel Dhome, Jean-Marc Lavest, and Nadine Daucher. Hand-eye calibration. In *Proceedings of the 1997 IEEE/RSJ International Conference on Intelligent Robot and Systems. Innovative Robotics for Real-World Applications. IROS'97*, volume 2, pages 1057–1065. IEEE, 1997.

Miguel Ribo and Markus Brandner. State of the art on vision-based structured light systems for 3d measurements. In *International Workshop on Robotic Sensors: Robotic and Sensor Environments, 2005.*, pages 2–6. IEEE, 2005.

Alexis Roche, Grégoire Malandain, Xavier Pennec, and Nicholas Ayache. The correlation ratio as a new similarity measure for multimodal image registration. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 1115–1124. Springer, 1998.

RTP-Company. ulprospector. URL `https://knowledge.ulprospector.com/4467/pe-color-science-101-coloring-plastics/`. Accessed January 04, 2023.

Ethan Rublee, Vincent Rabaud, Kurt Konolige, and Gary Bradski. Orb: An efficient alternative to sift or surf. In *2011 International conference on computer vision*, pages 2564–2571. Ieee, 2011.

Pedro Santos, Martin Ritz, Constanze Fuhrmann, and Dieter Fellner. 3d mass digitization: a milestone for archeological documentation. *Virtual Archaeology Review*, 8(16):1–11, 2017.

Yasuyuki Matsushita Satoshi Ikehata, David P. Wipf and Kiyoharu Aizawa. Photometric stereo using sparse bayesian regression for general diffuse surfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 36(9):1078–1091, 2014.

Yasuyuki Matsushita Satoshi Ikehata, David Wipf and Kiyoharu Aizawa. Robust photometric stereo using sparse regression. In *Proceedings of Computer Vision and Pattern Recognition (CVPR)*, 2012.

Mark J Schervish and Morris H DeGroot. *Probability and statistics.* Pearson Education, 2012.

Christophe Schlick. A survey of shading and reflectance models. In *computer graphics forum*, volume 13, pages 121–131. Wiley Online Library, 1994.

Boxin Shi, Zhe Wu, Zhipeng Mo, Dinglong Duan, Sai-Kit Yeung, and Ping Tan. A benchmark dataset and evaluation for non-lambertian and uncalibrated photometric stereo. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3707–3716, 2016.

William M Silver. *Determining shape and reflectance using multiple images.* PhD thesis, Massachusetts Institute of Technology, 1980.

Klaus H Strobl and Gerd Hirzinger. Optimal hand-eye calibration. In *2006 IEEE/RSJ international conference on intelligent robots and systems*, pages 4647–4653. IEEE, 2006.

Ping Tan. Phong reflectance model. *Computer Vision: A Reference Guide*, pages 1–3, 2020.

Kenneth E Torrance and Ephraim M Sparrow. Theory for off-specular reflection from roughened surfaces. *Josa*, 57(9):1105–1114, 1967.

Chih-Fong Tsai. A review of image retrieval methods for digital cultural heritage resources. *Online Information Review*, 2007.

R Steven Turley. Polynomial fitting. 2018.

Kenneth J Voss, Albert Chapin, Marco Monti, and Hao Zhang. Instrument to measure the bidirectional reflectance distribution function of surfaces. *Applied Optics*, 39(33):6197–6206, 2000.

Nikolaos Voudoukis and Sarantos Oikonomidis. Inverse square law for light and radiation: A unifying educational approach. *European Journal of Engineering and Technology Research*, 2(11):23–27, 2017.

Bruce Walter, Stephen R Marschner, Hongsong Li, and Kenneth E Torrance. Microfacet models for refraction through rough surfaces. In *Proceedings of the 18th Eurographics conference on Rendering Techniques*, pages 195–206, 2007.

Oliver Wang, Prabath Gunawardane, Steve Scher, and James Davis. Material classification using brdf slices. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pages 2805–2811. IEEE, 2009.

EK Webb and M Wachowiak. Imaging studio technical note: Flexible solutions for reflectance transformation imaging (rti). *Smithsonian Museum*, 2011.

Martin H. Weik. *Lambert's cosine law*, pages 868–868. Springer US, Boston, MA, 2001. ISBN 978-1-4020-0613-5. doi: 10.1007/1-4020-0613-6_9901. URL `https://doi.org/10.1007/1-4020-0613-6_9901`.

Robert J Woodham. Photometric method for determining surface orientation from multiple images. *Optical engineering*, 19(1):191139, 1980.

Gaochang Wu, Belen Masia, Adrian Jarabo, Yuchen Zhang, Liangyong Wang, Qionghai Dai, Tianyou Chai, and Yebin Liu. Light field image processing: An overview. *IEEE Journal of Selected Topics in Signal Processing*, 11(7):926–954, 2017.

Qingxiong Yang and Narendra Ahuja. Surface reflectance and normal estimation from photometric stereo. *Computer Vision and Image Understanding*, 116(7):793–802, 2012.

Yuki Mikamoto Yoshiki Kaminaka and Kazufumi Kaneda. Eurograhics, 2022. URL `http://vcg.isti.cnr.it/cgf/winner.php`. Accessed January 04, 2023.

Jinze Yu, Mingqiang Wei, Jing Qin, Jianhuang Wu, and Pheng-Ann Heng. Feature-preserving mesh denoising via normal guided quadric error metrics. *Optics and Lasers in Engineering*, 62:57–68, 2014.

AL Yuille. Lecture 10: the lambertian reflectance model, 2012.

Abir Zendagui, Jean-Baptiste Thomas, Gaëtan Le Goïc, Yuly Castro, Marvin Nurit, Alamin Mansouri, and Marius Pedersen. Quality assessment of reconstruction and relighting from rti images: application to manufactured surfaces. In *2019 15th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)*, pages 746–753. IEEE, 2019.

Abir Zendagui, Gaëtan Le Goïc, Hermine Chatoux, Jean-Baptiste Thomas, Yuly Castro, Marvin Nurit, and Alamin Mansouri. Quality assessment of dynamic virtual relighting from rti data: application to the inspection of engineering surfaces. In *Fifteenth International Conference on Quality Control by Artificial Vision*, volume 11794, pages 94–102. SPIE, 2021.

Ruo Zhang, Ping-Sing Tsai, James Edwin Cryer, and Mubarak Shah. Shape-from-shading: a survey. *IEEE transactions on pattern analysis and machine intelligence*, 21(8):690–706, 1999.

Xiaocheng Zhou, Xueping Liu, Xiaoqin Wang, Guojin He, Youshui Zhang, Guizhou Wang, and Zhaoming Zhang. Evaluation of surface reflectance products based on optimized 6s model using synchronous in situ measurements. *Remote Sensing*, 14(1):83, 2022.

Barbara Zitova and Jan Flusser. Image registration methods: a survey. *Image and vision computing*, 21(11):977–1000, 2003.

APPENDIX A

LightBot

A.1 XY platform control

The XY platform used in the LightBot system consists of two stepper motors one for each axis. We used TB6600 stepper motor driver and Arduino drive and control the motors. The Arduino controller is connected to the computer using the USB connector and programmed to control the X and Y position. The electrical connections for this motors are as depicted in Figure.A.1.

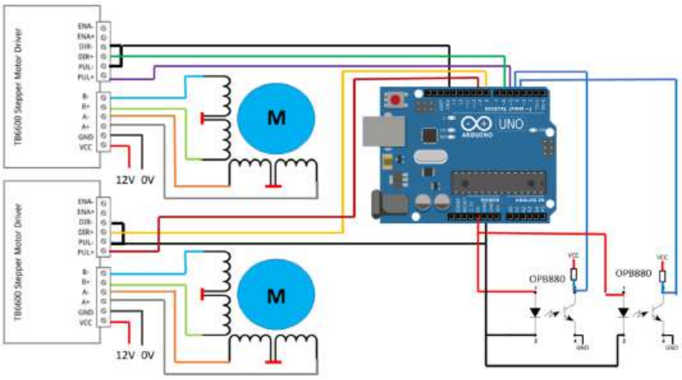
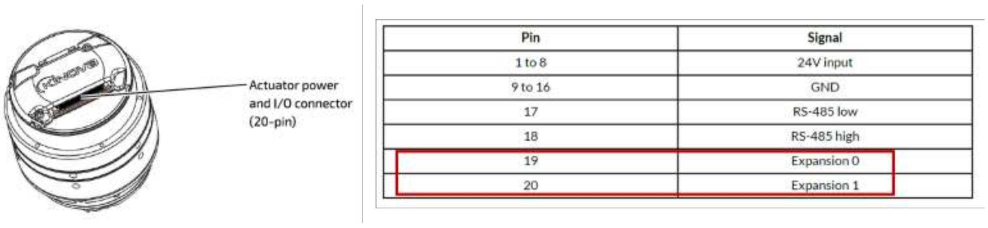


Figure A.1: Electrical connections of the XY platform motors.

A.2 Light source

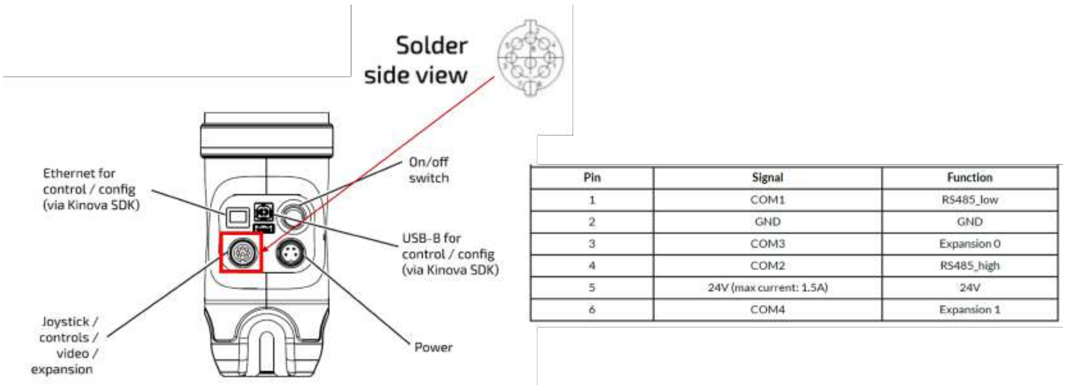
Kinova robotic arm offers two expansion pins at the end effector connected internally to the robot arm base. This is very convenient to wire the light source cleanly especially in a tight space like LightBot. Figure.A.2 shows the expansion pins available in the arm end effector that we use to power the light source.



Pin	Signal
1 to 8	24V input
9 to 16	GND
17	RS-485 low
18	RS-485 high
19	Expansion 0
20	Expansion 1

Figure A.2: Expansion pins in Kinova robotic arm end effector.

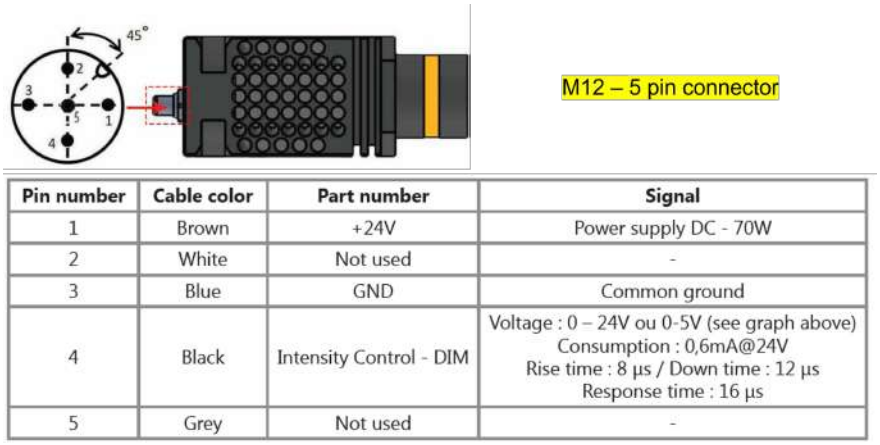
Figure.A.3 shows the expansion pins at the base of the robotic arm. These pins are connected to the corresponding expansion pins in the end effector through a diode.



Pin	Signal	Function
1	COM1	RS485_low
2	GND	GND
3	COM3	Expansion 0
4	COM2	RS485_high
5	24V (max current: 1.5A)	24V
6	COM4	Expansion 1

Figure A.3: Expansion pins which are connected from the end effector to the base of the Kinova robotic arm end effector.

Figure.A.4 shows the electrical connections of the EFFI spot light that we have used in LightBot 2 system. The power source of for the light is connected to the expansion pins at the base and then from the end effector, the power is tapped and connected to the light source.



Pin number	Cable color	Part number	Signal
1	Brown	+24V	Power supply DC - 70W
2	White	Not used	-
3	Blue	GND	Common ground
4	Black	Intensity Control - DIM	Voltage : 0 – 24V ou 0-5V (see graph above) Consumption : 0,6mA@24V Rise time : 8 µs / Down time : 12 µs Response time : 16 µs
5	Grey	Not used	-

Figure A.4: EFFI spot light source electrical connector

Figure.A.5 shows the PCB that we fabricated to connect the light source to the molex connector at the robotic arm end effector.

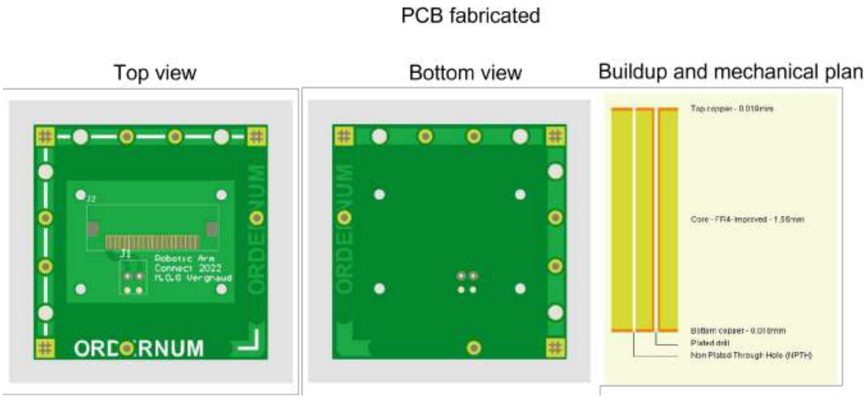


Figure A.5: PCB that we fabricated to connect the light source to the molex connector at the robotic arm end effector.

A.3 User interface

My Ph.D. is part of the multi disciplinary CHANGE project consisting of Ph.D students from cultural heritage background and our contribution involved strong collaboration with the colleagues from cultural heritage field. Hence we built an easy to use user interface to use our LightBot. Figure.A.6 gives an overview of the main page of the UI.

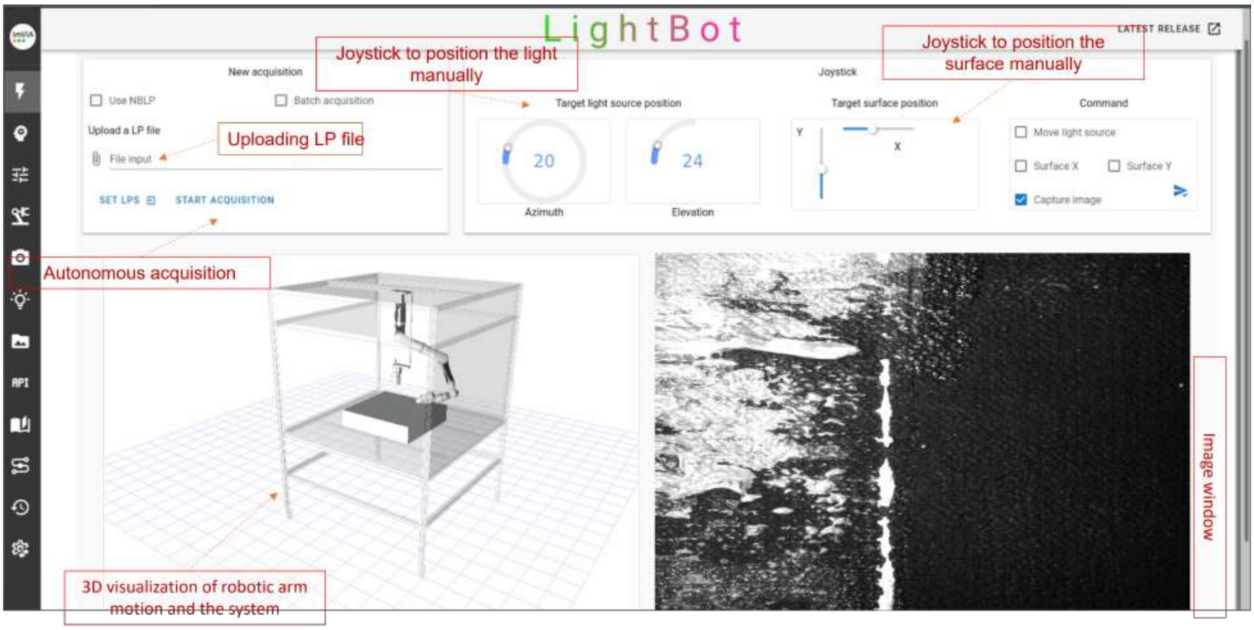


Figure A.6: Home page of the LightBot UI

Figure.A.7 illustrates the selection tabs for operating the system consisting of pages to control the robotic arm, camera, light source etc individually.

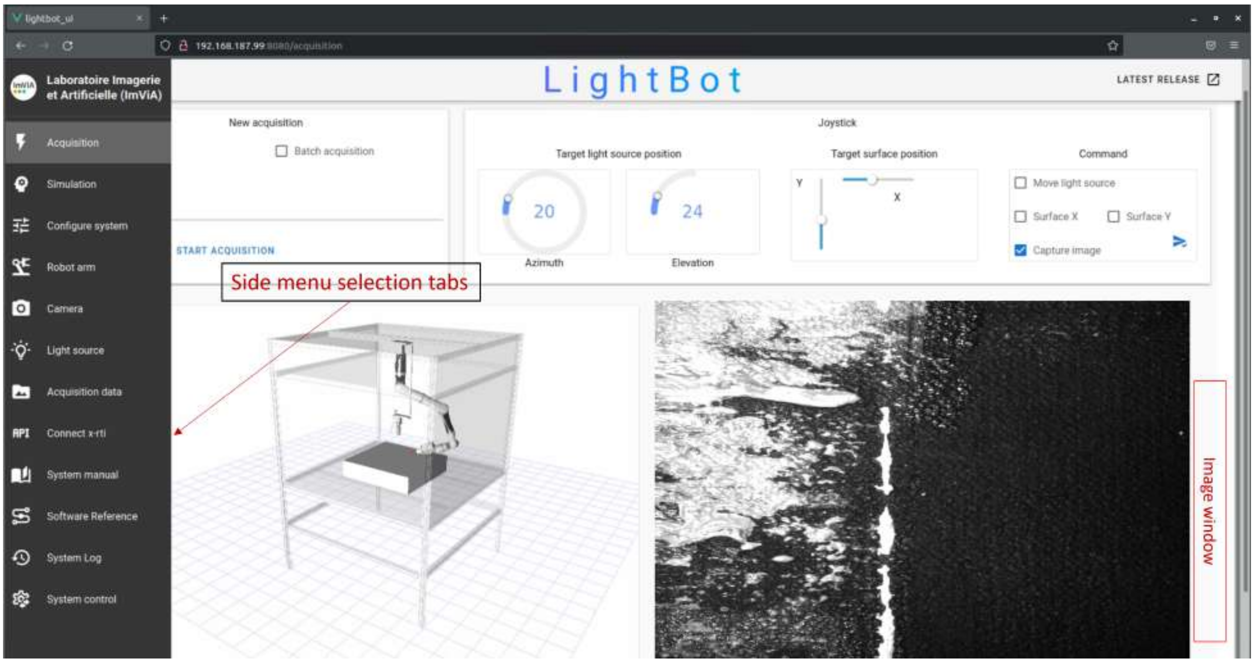


Figure A.7: Side menu selection tabs

We adopted modular architecture for the system which gives the ability to control and run the components independently. Figure.A.9 show the node status and control page. This page basically is the main control page from which we can turn on/off the system and startup the machine.

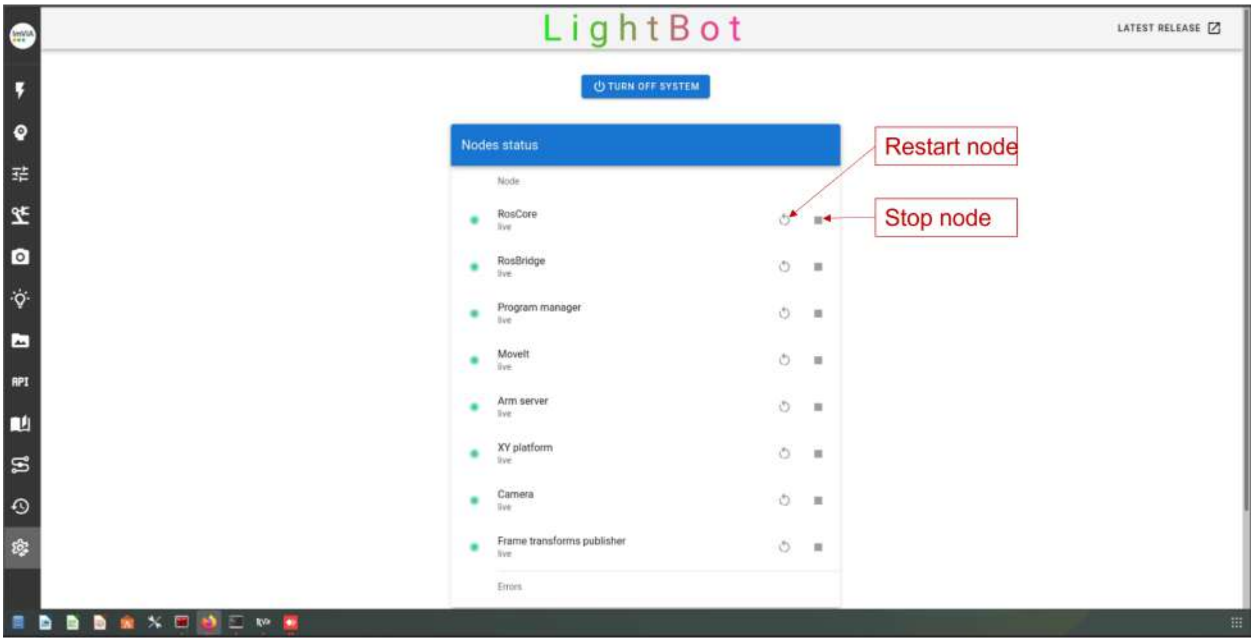
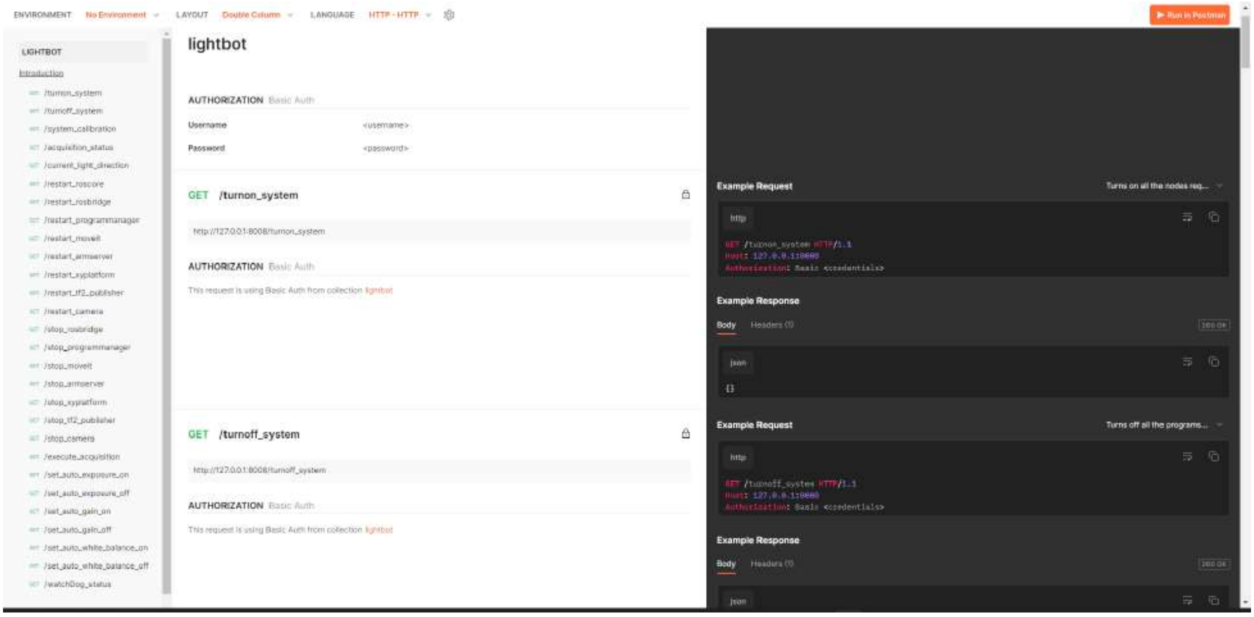


Figure A.8: Node status and control page

A.4 API

We developed the API documentation. Our LightBot API allows other researchers and developers to easily integrate LightBot to their systems to test their methods or control the system. We used Postman a popular tool for LightBot API development and testing. Postman is a convenient tool since it allows developers to create, test, share, document APIs and generate sample requests, code snippets in any programming languages to interface with the system. LightBot API documentation is available online at https://documenter.getpostman.com/view/8477825/UVXqECQf.



lightbot

AUTHORIZATION Basic Auth

Username <username>

Password <password>

GET /turnon_system

http://127.0.0.1:8008/turnon_system

AUTHORIZATION Basic Auth

This request is using Basic Auth from collection lightbot

GET /turnoff_system

http://127.0.0.1:8008/turnoff_system

AUTHORIZATION Basic Auth

This request is using Basic Auth from collection lightbot

Example Request — Turns on all the nodes req...

```
GET /turnon_system HTTP/1.1
Host: 127.0.0.1:8008
Authorization: Basic <credentials>
```

Example Response

```
{}
```

Example Request — Turns off all the programs...

```
GET /turnoff_system HTTP/1.1
Host: 127.0.0.1:8008
Authorization: Basic <credentials>
```

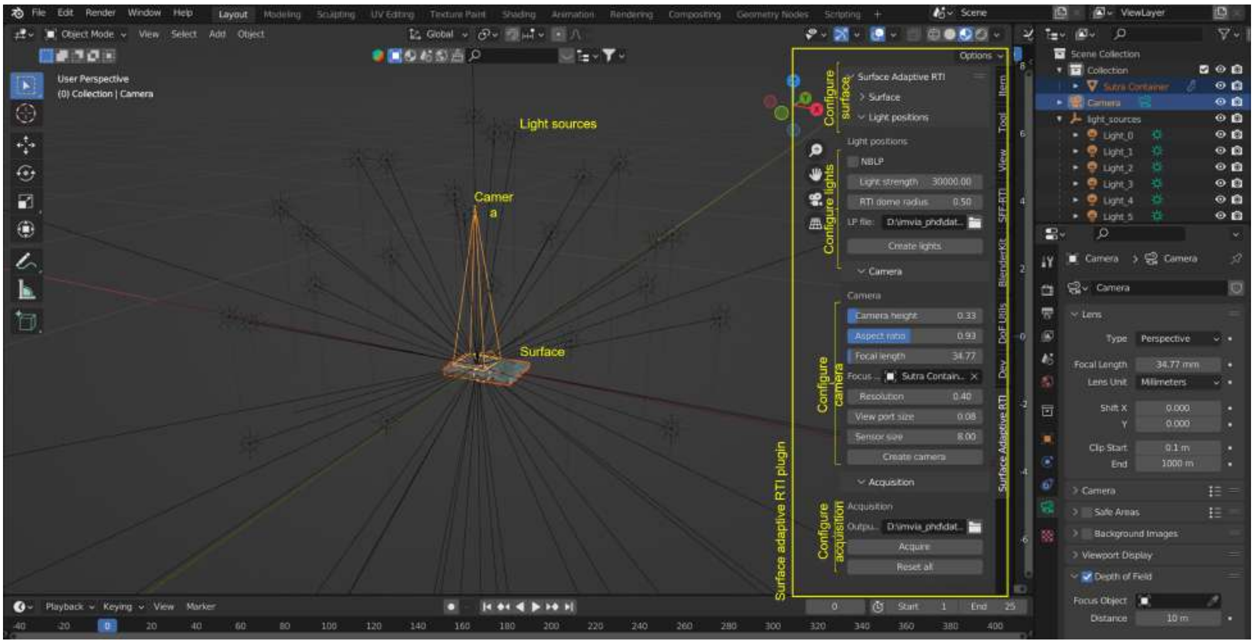
Figure A.9: Node status and control page

APPENDIX B

RTI-Dataset

B.1 Blender plugin for virtual NBLP RTI acquisitions

We developed a plugin for Blender software to perform NBLP RTI acquisition of virtual surfaces. This plugin can be used by researchers to test RTI methods in simulated environment. The plugin features easy to use GUI for performing virtual RTI acquisitions. This is created to perform acquisition of the virtual surfaces created in our RTI dataset. Figure.B.1 gives an overview of the plugin that we created in our work.



Light positions
☐ NBLP
Light strength 30000.00
RTI dome radius 0.50
LP file: D:\imvia_phd\dat...

Camera
Camera height 0.33
Aspect ratio 0.93
Focal length 34.77
Focus ... Sutra Contain...
Resolution 0.40
View port size 0.08
Sensor size 8.00

Acquisition
Outpu... D:\imvia_phd\dat...

Type Perspective
Focal Length 34.77 mm
Lens Unit Millimeters
Shift X 0.000
Y 0.000
Clip Start 0.1 m
End 1000 m
☑ Depth of Field
Focus Object
Distance 10 m

0 Start 1 End 25

Figure B.1: Overview of the NBLP-RTI plugin developed for the Blender software to perform virtual RTI acquisitions

B.2 RTI Stitching data

We acquired two medium sized surfaces using LightBot to develop and demonstrate our stitching methods. We have shared the acquisition data publicly through Zenodo as open source for any researcher to access and using it for developing new stitching methods. The data can be accessed form https://zenodo.org/record/7315762

APPENDIX C

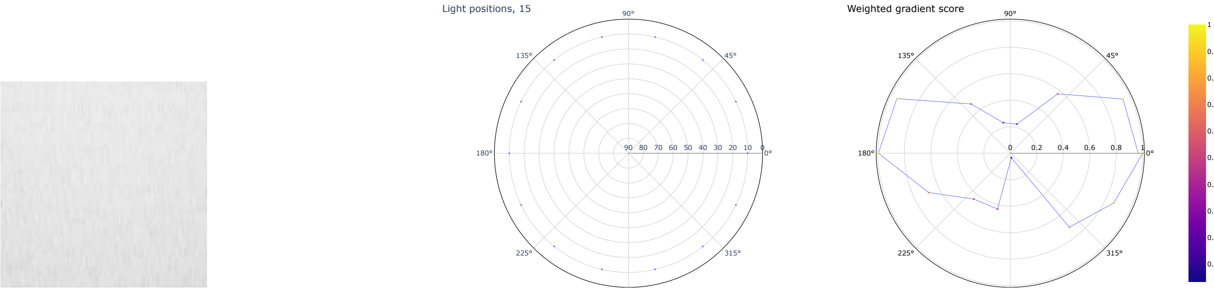
NBLP

C.1 Results of Azimuth only space NBLP

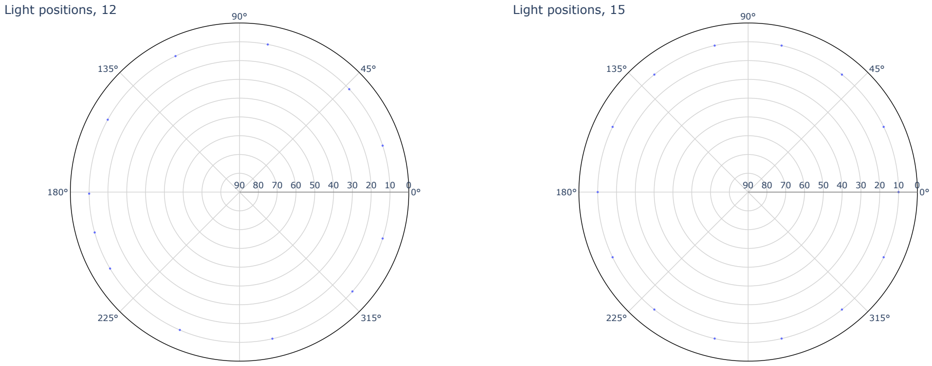
Here we show the Azimuth only space NBLP results for the surface are shown in Figures. The top figure displays the iteration steps during the NBLP acquisition, while the bottom figure compares the overall distribution of light positions between the ideal reference acquisition and the final NBLP acquisition. Each illustration is accompanied by its corresponding confusion matrix.

C.1.1 Simple brushed metal

Elevation 10˙



(a) Simple brushed metal surface (b) Initial acquisition



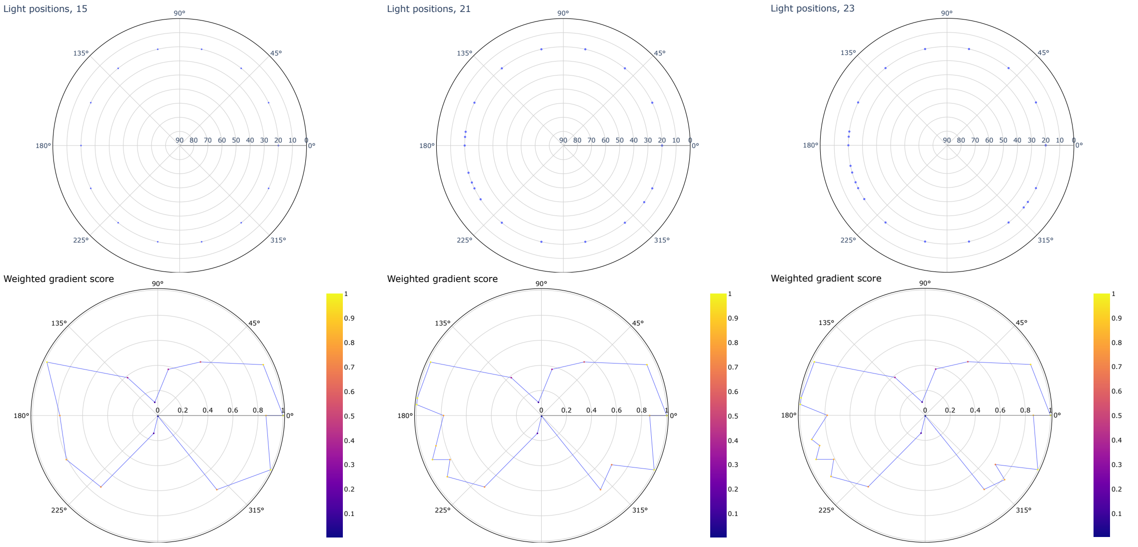
(c) Overall distribution of the light positions in ideal reference acquisition (left) and NBLP acquisition (right)

Figure C.1: NBLP acquisition on a simple brushed metal surface in the azimuthal space ring at an elevation of 10˙.

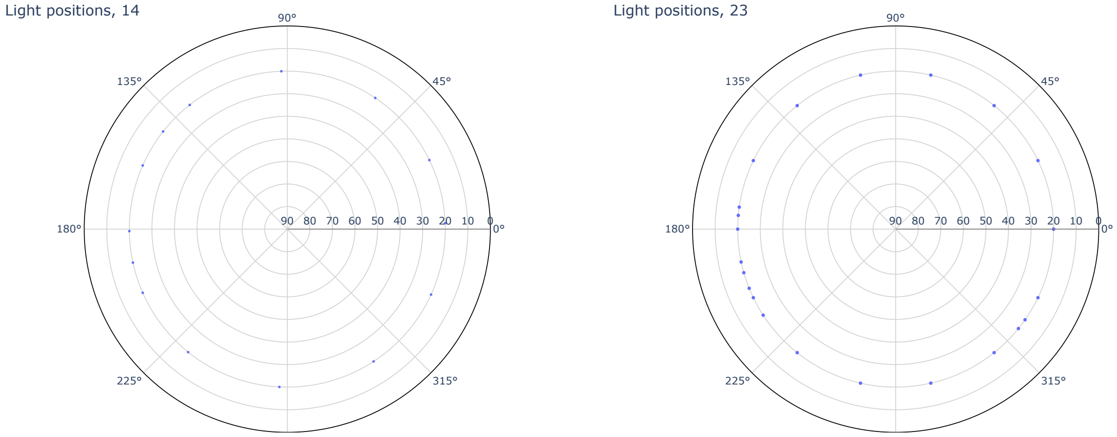
	Positive	Negative	
Positive	2	18	Sensitivity 0.14
Negative	42	297	Specificity 0.87
	Precision 0.06	Negative predictive value 0.94	Accuracy 0.83

Table C.1: Confusion matrix of the results for simple brushed metal surface at elevation 10˙

Elevation 20˙



(a) Initial acquisition (b) 1^{st}iteration (c) 2^{nd}iteration



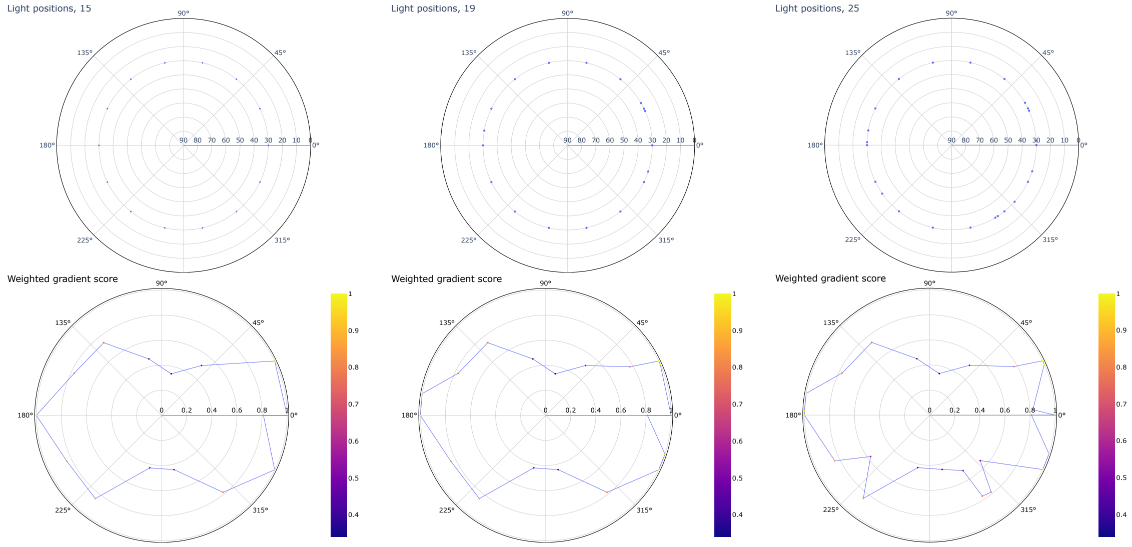
(d) Overall distribution of the light positions in ideal reference acquisition (left) and NBLP acquisition (right)

Figure C.2: NBLP acquisition on a simple brushed metal surface in the azimuthal space ring at an elevation of 20˙.

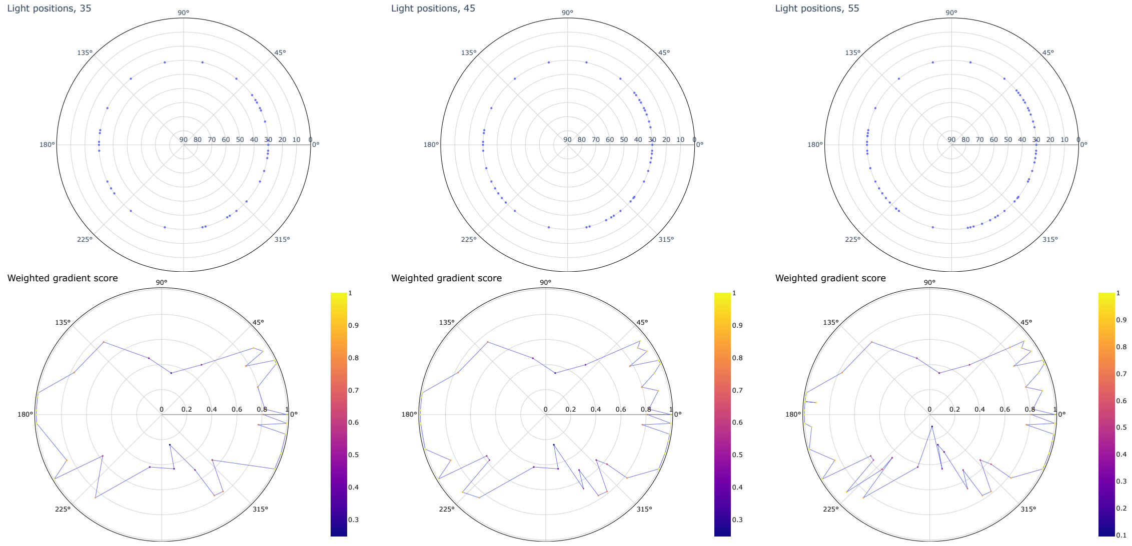
	Positive	Negative	
Positive	12	16	Sensitivity 0.42
Negative	41	291	Specificity 0.87
	Precision 0.22	Negative predictive value 0.94	Accuracy 0.84

Table C.2: Confusion matrix of the results for simple brushed metal surface at elevation 20˙

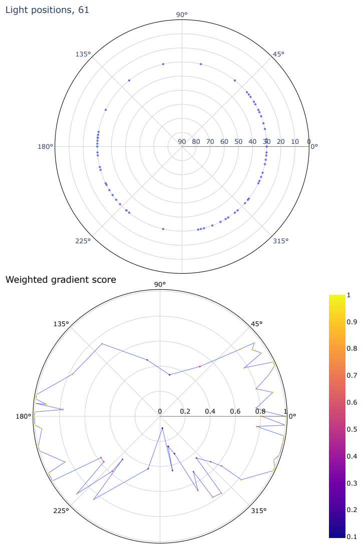
Elevation 30˙



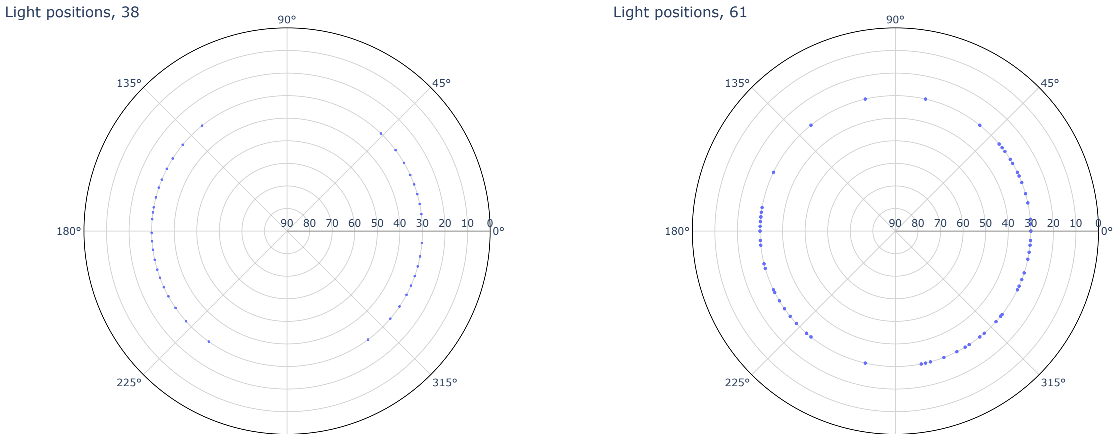
(a) Initial acquisition (b) 1^{st}iteration (c) 2^{nd}iteration



(d) 3^{rd}iteration (e) 4^{th}iteration (f) 5^{th}iteration



(g) 6^{th}iteration



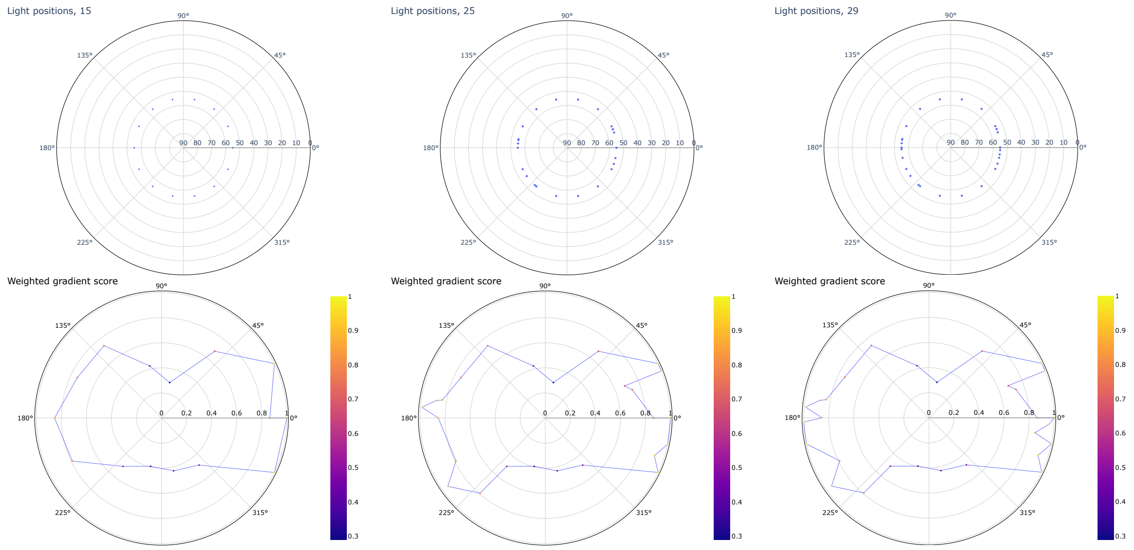
(h) Overall distribution of the light positions in ideal reference acquisition (left) and NBLP acquisition (right)

Figure C.3: NBLP acquisition on a simple brushed metal surface in the azimuthal space ring at an elevation of 30˙.

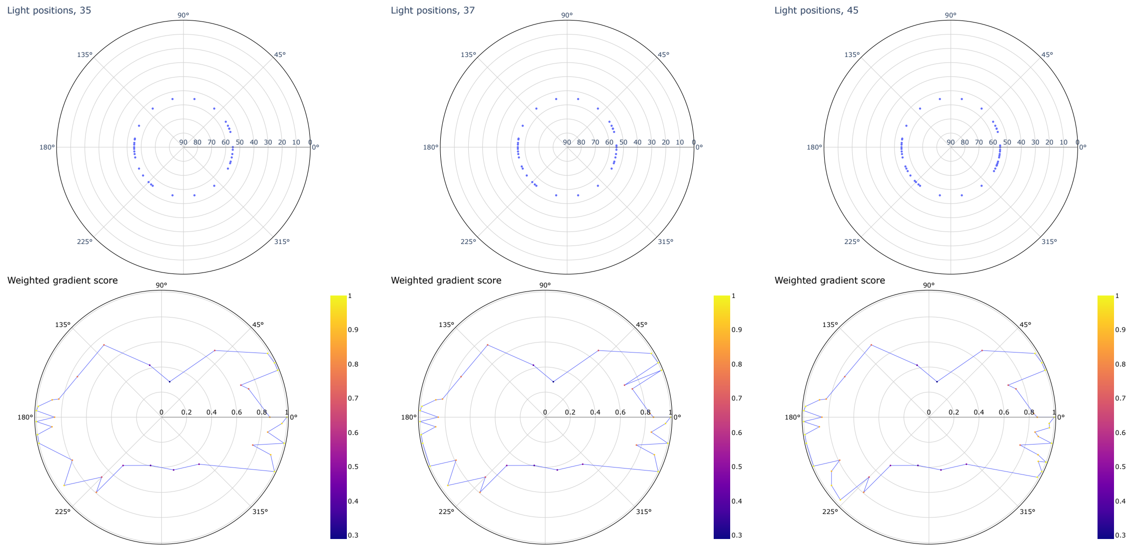
	Positive	Negative	
Positive	42	27	Sensitivity 0.60
Negative	44	247	Specificity 0.84
	Precision 0.48	Negative predictive value 0.90	Accuracy 0.80

Table C.3: Confusion matrix of the results for simple brushed metal surface at elevation 30˙

Elevation 55˙



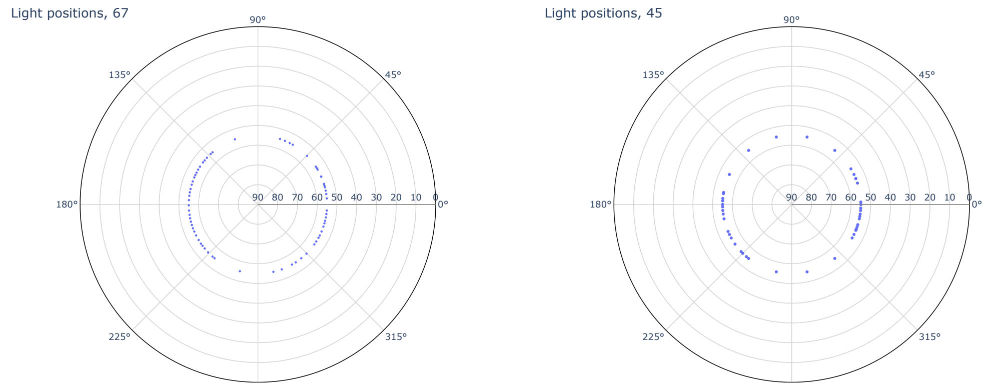
(a) Initial acquisition (b) 1stiteration (c) 2nditeration



(d) 3rditeration (e) 4thiteration (f) 5thiteration

	Positive	Negative	
Positive	30	69	Sensitivity 0.30
Negative	36	225	Specificity 0.86
	Precision 0.45	Negative predictive value 0.76	Accuracy 0.70

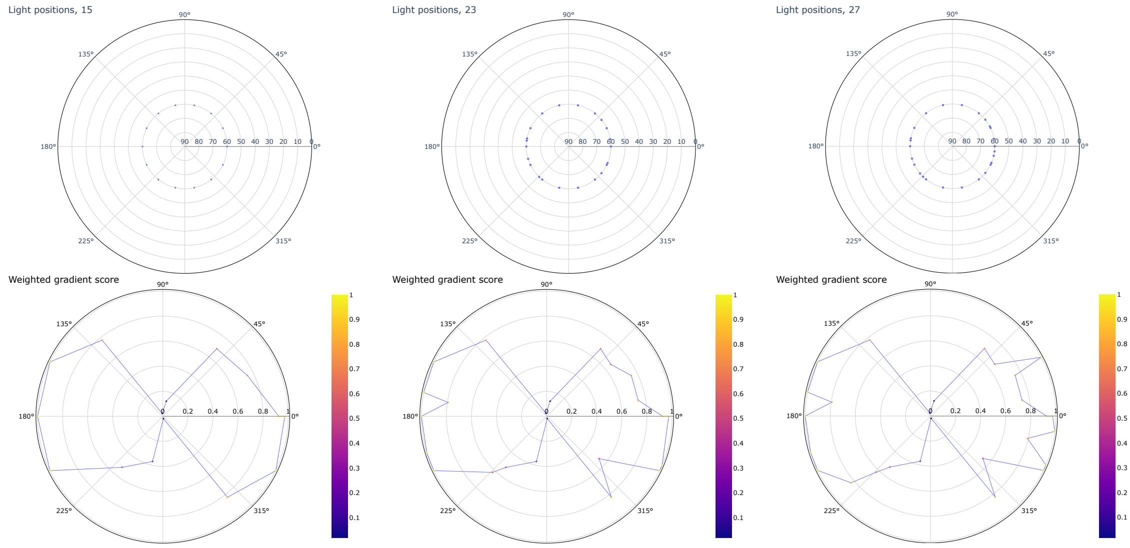
Table C.4: Confusion matrix of the results for simple brushed metal surface at elevation 55˙



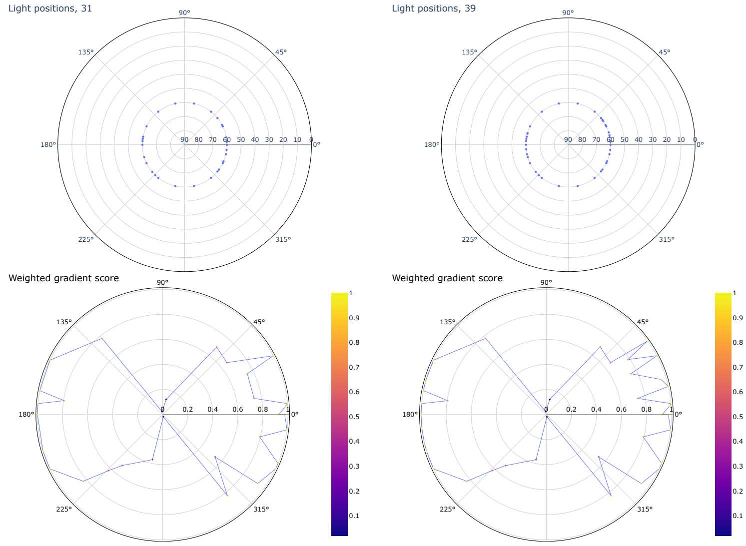
(g) Overall distribution of the light positions in ideal reference acquisition (left) and NBLP acquisition (right)

Figure C.4: NBLP acquisition on a simple brushed metal surface in the azimuthal space ring at an elevation of 55˙.

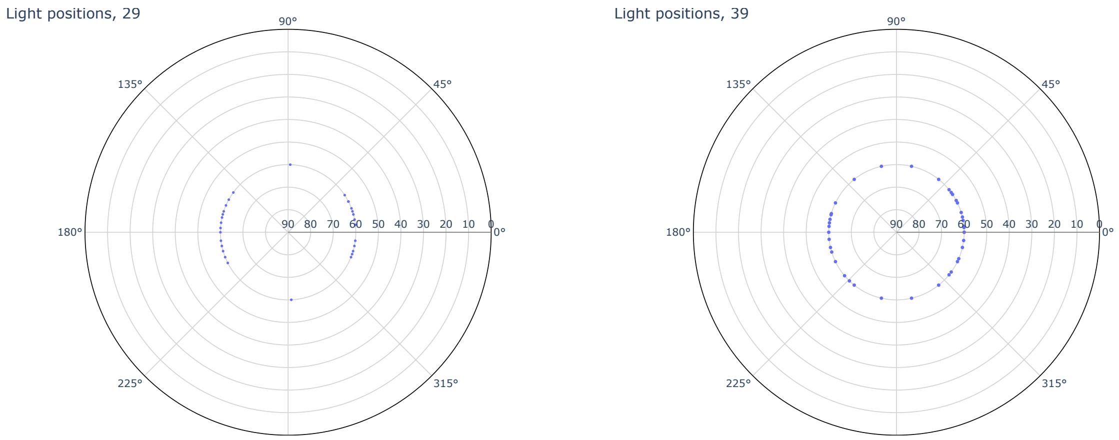
Elevation 60˙



(a) Initial acquisition (b) 1^{st}iteration (c) 2^{nd}iteration



(d) 3^{rd}iteration (e) 4^{th}iteration



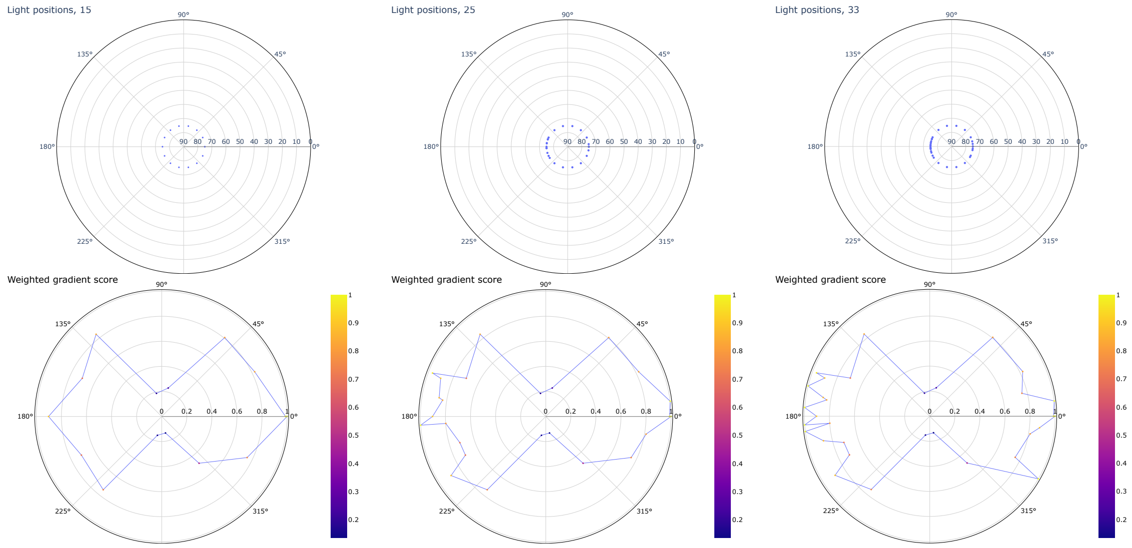
(f) Overall distribution of the light positions in ideal reference acquisition (left) and NBLP acquisition (right)

Figure C.5: NBLP acquisition on a simple brushed metal surface in the azimuthal space ring at an elevation of 60˙.

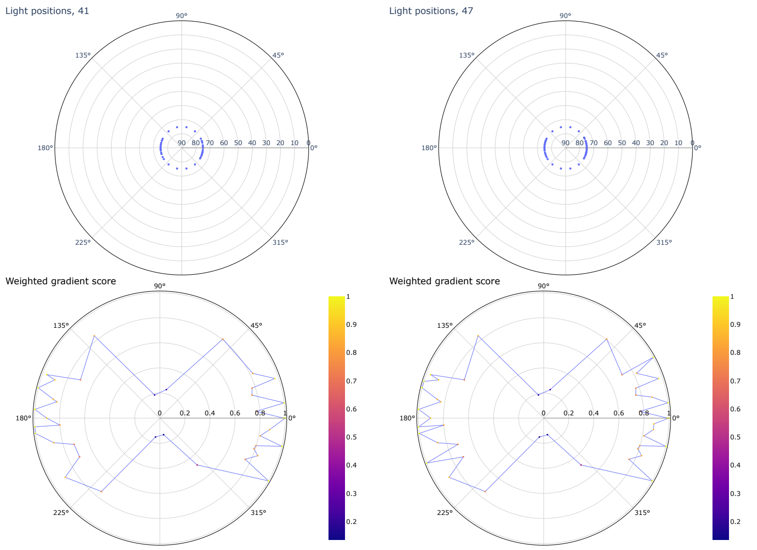
	Positive	Negative	
Positive	41	19	Sensitivity 0.84
Negative	48	252	Specificity 0.68
	Precision 0.46	Negative predictive value 0.92	Accuracy 0.81

Table C.5: Confusion matrix of the results for simple brushed metal surface at elevation 60˙

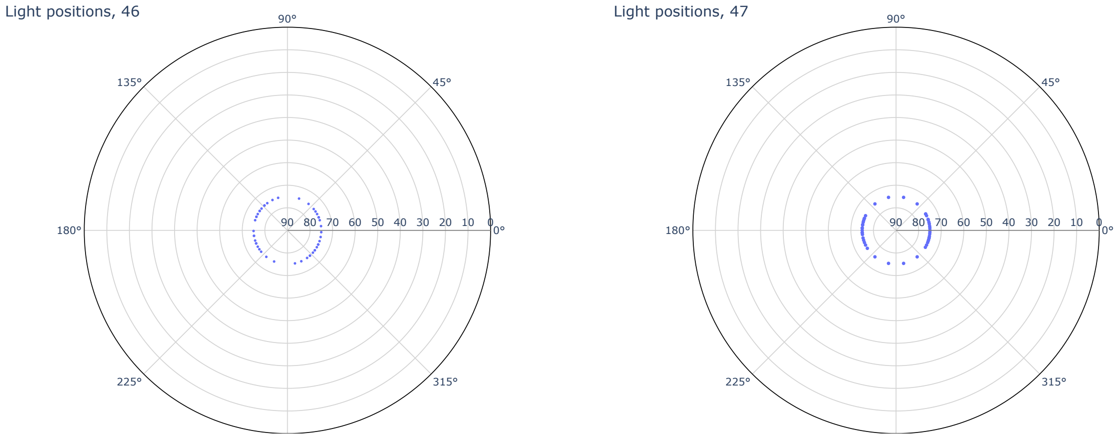
Elevation 75˙



(a) Initial acquisition (b) 1^{st}iteration (c) 2^{nd}iteration



(d) 3rditeration (e) 4thiteration



(f) Overall distribution of the light positions in ideal reference acquisition (left) and NBLP acquisition (right)

Figure C.6: NBLP acquisition on a simple brushed metal surface in the azimuthal space ring at an elevation of 75˙.

	Positive	Negative	
Positive	30	38	Sensitivity 0.44
Negative	52	240	Specificity 0.82
	Precision 0.36	Negative predictive value 0.86	Accuracy 0.75

Table C.6: Confusion matrix of the results for simple brushed metal surface at elevation 75˙

C.1.2 Brushed metal with a dent

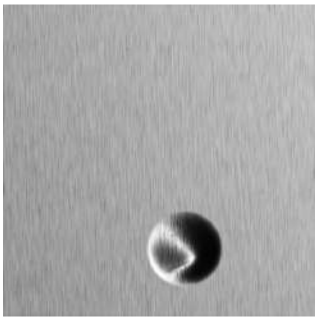


Figure C.7: Brushed metal surface with a dent (surface 2 in the rti-dataset)

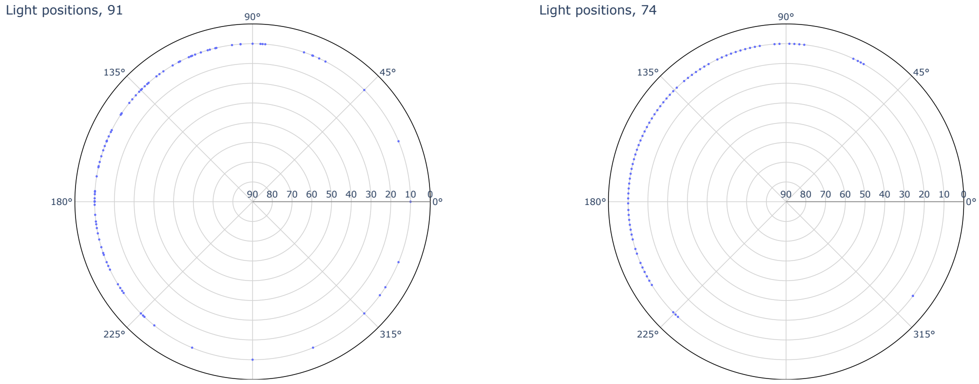


Figure C.8: Overall distribution of the light positions in ideal reference acquisition (left) and NBLP acquisition (right) of brushed metal surface with a dent at elevation 10˙

	Positive	Negative	
Positive	95	11	Sensitivity 0.89
Negative	28	226	Specificity 0.88
	Precision 0.77	Negative predictive value 0.95	Accuracy 0.89

Table C.7: Confusion matrix of the results for brushed metal surface with a dent at elevation 10˙

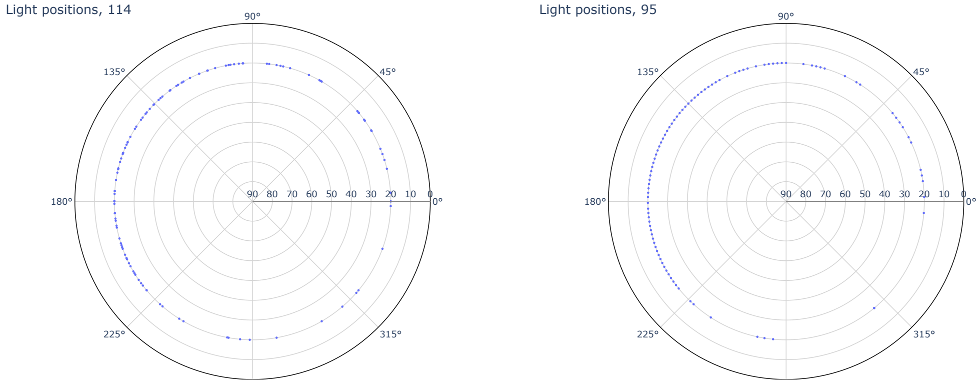


Figure C.9: Overall distribution of the light positions in ideal reference acquisition (left) and NBLP acquisition (right) of brushed metal surface with a dent at elevation 20˙

	Positive	Negative	
Positive	129	17	Sensitivity 0.88
Negative	23	191	Specificity 0.89
	Precision 0.84	Negative predictive value 0.91	Accuracy 0.88

Table C.8: Confusion matrix of the results for brushed metal surface with a dent at elevation 20˙

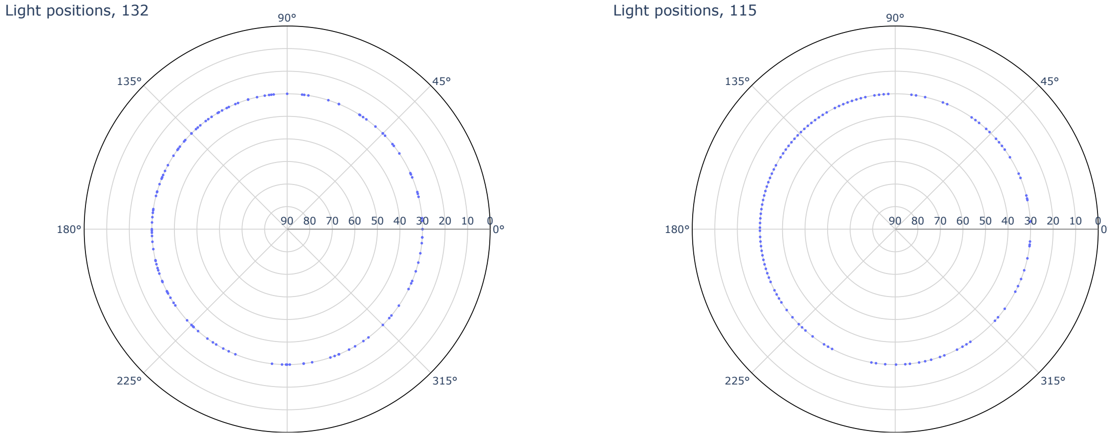


Figure C.10: Overall distribution of the light positions in ideal reference acquisition (left) and NBLP acquisition (right) of brushed metal surface with a dent at elevation 30˙

	Positive	Negative	
Positive	137	13	Sensitivity 0.91
Negative	23	187	Specificity 0.89
	Precision 0.85	Negative predictive value 0.93	Accuracy 0.90

Table C.9: Confusion matrix of the results for brushed metal surface with a dent at elevation 30˙

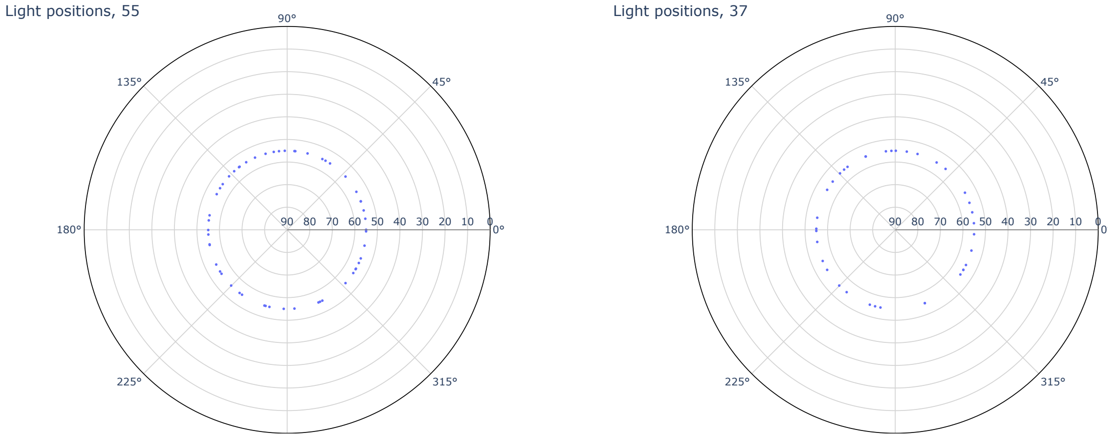


Figure C.11: Overall distribution of the light positions in ideal reference acquisition (left) and NBLP acquisition (right) of brushed metal surface with a dent at elevation 55˙

	Positive	Negative	
Positive	45	31	Sensitivity 0.59
Negative	73	211	Specificity 0.74
	Precision 0.38	Negative predictive value 0.87	Accuracy 0.71

Table C.10: Confusion matrix of the results for brushed metal surface with a dent at elevation 55˙

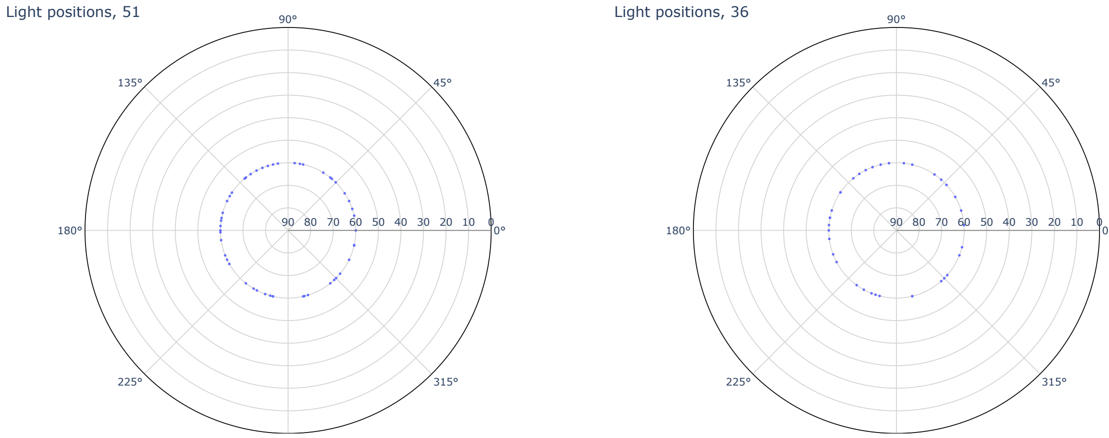


Figure C.12: Overall distribution of the light positions in ideal reference acquisition (left) and NBLP acquisition (right) of brushed metal surface with a dent at elevation 60˙

	Positive	Negative	
Positive	47	26	Sensitivity 0.64
Negative	69	218	Specificity 0.75
	Precision 0.40	Negative predictive value 0.89	Accuracy 0.73

Table C.11: Confusion matrix of the results for brushed metal surface with a dent at elevation 60˙

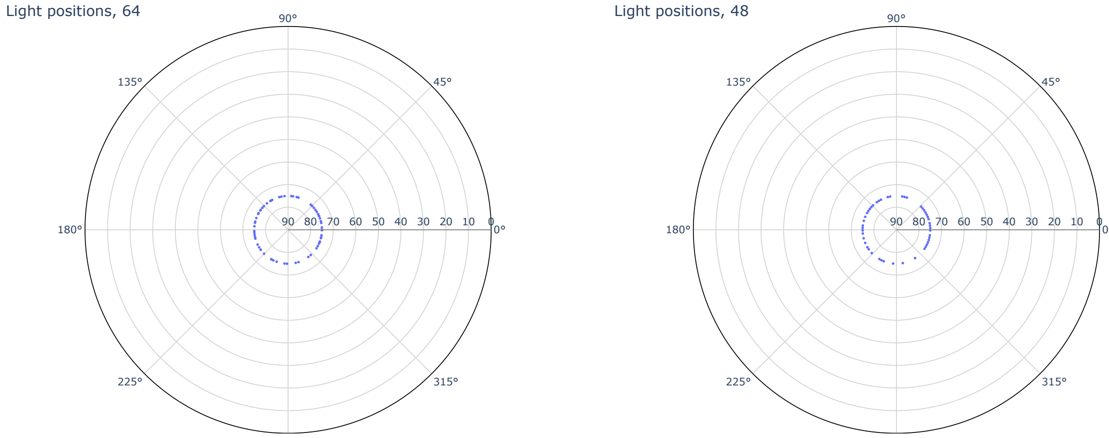


Figure C.13: Overall distribution of the light positions in ideal reference acquisition (left) and NBLP acquisition (right) of brushed metal surface with a dent at elevation 75˙

	Positive	Negative	
Positive	86	24	Sensitivity 0.78
Negative	38	212	Specificity 0.84
	Precision 0.69	Negative predictive value 0.89	Accuracy 0.82

Table C.12: Confusion matrix of the results for brushed metal surface with a dent at elevation 75˙

C.1.3 Ancient coin 1

Here, we present the comparison of overall NBLP estimated light positions and the reference light positions for the ancient coin 1 surface(surface 14 in our dataset).

Figure C.14: ancient coin 1 (surface 2 in the rti-dataset)

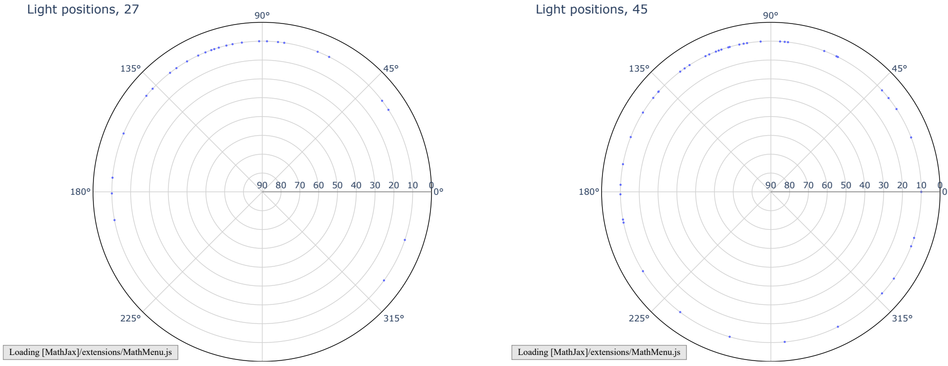


Figure C.15: Overall distribution of the light positions in ideal reference acquisition (left) and NBLP acquisition (right) of Ancient coin 1 at elevation 10˙

	Positive	Negative	
Positive	61	17	Sensitivity 0.78
Negative	46	236	Specificity 0.83
	Precision 0.57	Negative predictive value 0.93	Accuracy 0.82

Table C.13: Confusion matrix of the results for ancient coin 1 at elevation 10˙

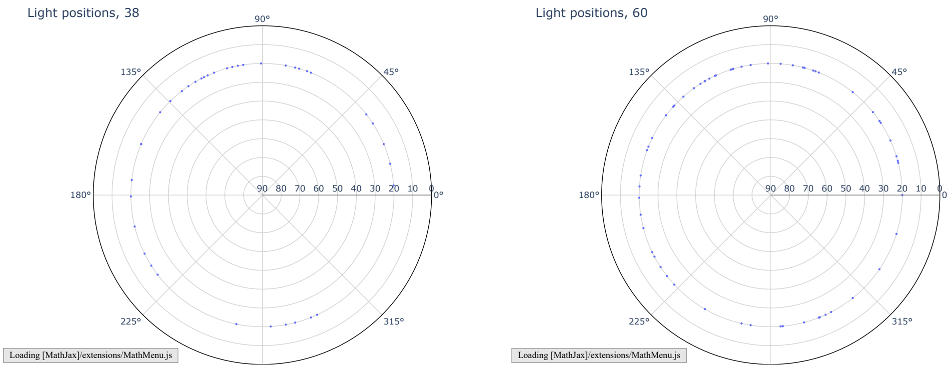


Figure C.16: Overall distribution of the light positions in ideal reference acquisition (left) and NBLP acquisition (right) of ancient coin 1 at elevation 20˙

	Positive	Negative	
Positive	66	20	Sensitivity 0.76
Negative	57	217	Specificity 0.79
	Precision 0.53	Negative predictive value 0.91	Accuracy 0.78

Table C.14: Confusion matrix of the results for ancient coin 1 at elevation 20˙

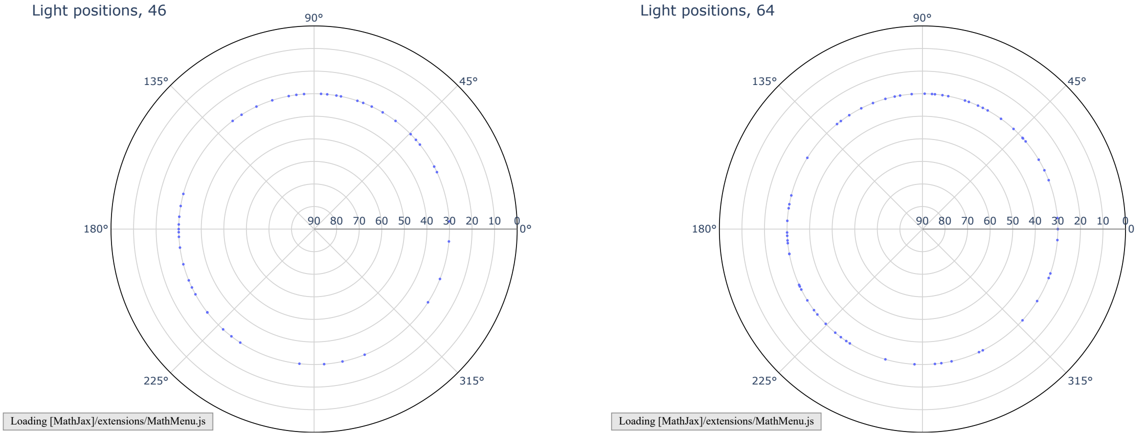


Figure C.17: Overall distribution of the light positions in ideal reference acquisition (left) and NBLP acquisition (right) of ancient coin 1 at elevation 30˙

	Positive	Negative	
Positive	78	17	Sensitivity 0.82
Negative	45	220	Specificity 0.83
	Precision 0.63	Negative predictive value 0.92	Accuracy 0.82

Table C.15: Confusion matrix of the results for ancient coin 1 at elevation 30˙

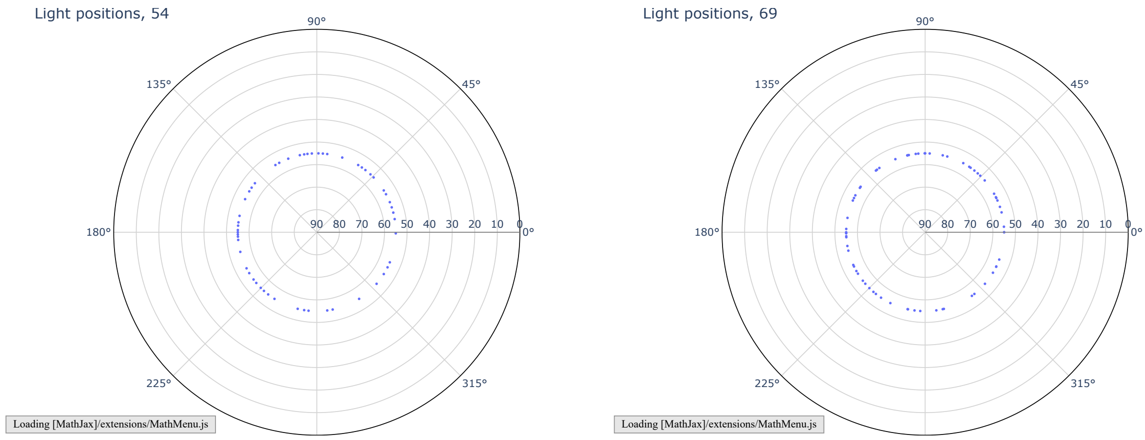


Figure C.18: Overall distribution of the light positions in ideal reference acquisition (left) and NBLP acquisition (right) of ancient coin 1 at elevation 55˙

	Positive	Negative	
Positive	76	35	Sensitivity 0.81
Negative	45	204	Specificity 0.68
	Precision 0.62	Negative predictive value 0.85	Accuracy 0.77

Table C.16: Confusion matrix of the results for ancient coin 1 at elevation 55˙

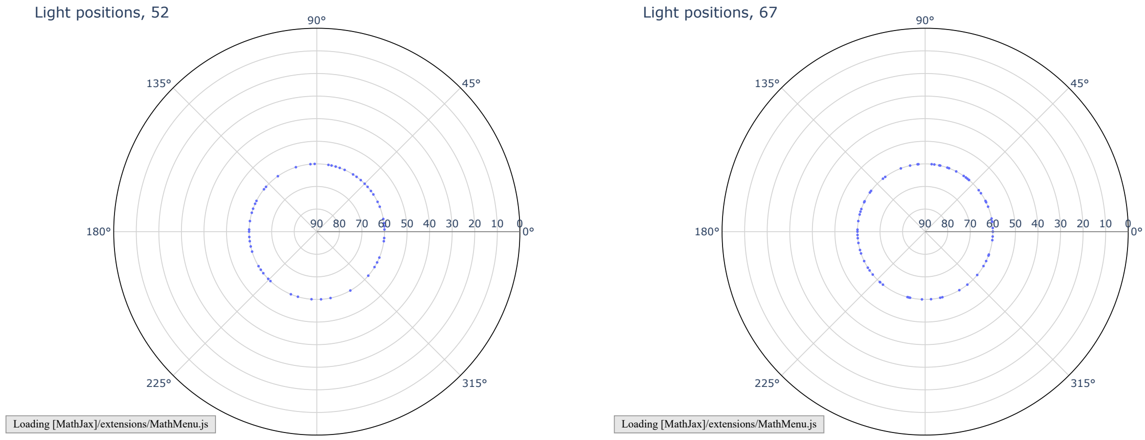


Figure C.19: Overall distribution of the light positions in ideal reference acquisition (left) and NBLP acquisition (right) of ancient coin 1 at elevation 60˙

	Positive	Negative	
Positive	77	33	Sensitivity 0.81
Negative	47	203	Specificity 0.70
	Precision 0.62	Negative predictive value 0.86	Accuracy 0.77

Table C.17: Confusion matrix of the results for ancient coin 1 at elevation 60˙

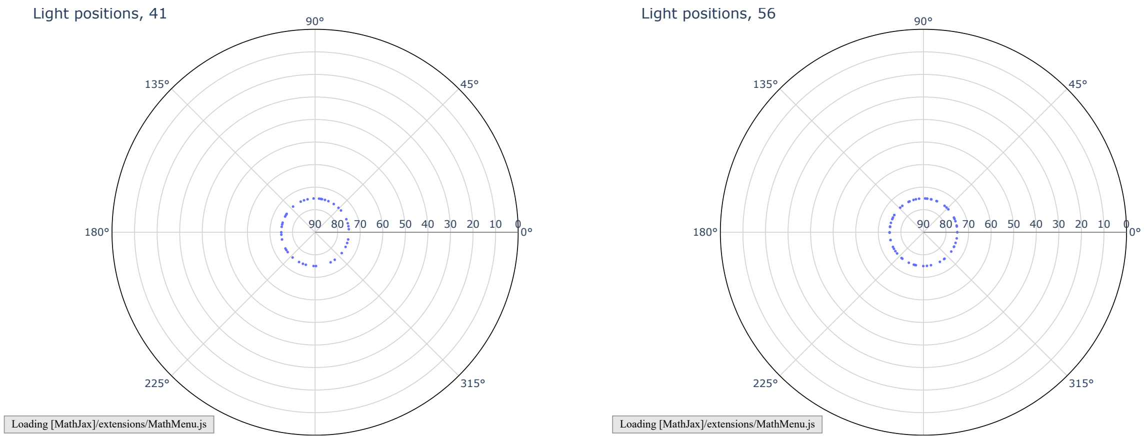


Figure C.20: Overall distribution of the light positions in ideal reference acquisition (left) and NBLP acquisition (right) of ancient coin 1 at elevation 75°

	Positive	Negative	
Positive	52	33	Sensitivity 0.76
Negative	66	209	Specificity 0.61
	Precision 0.44	Negative predictive value 0.86	Accuracy 0.72

Table C.18: Confusion matrix of the results for ancient coin 1 at elevation 75°

C.1.4 Canvas painting

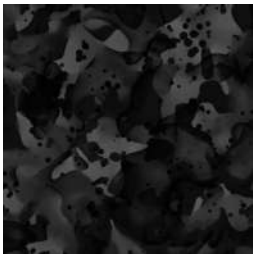


Figure C.21: canvas painting (surface 2 in the rti-dataset)

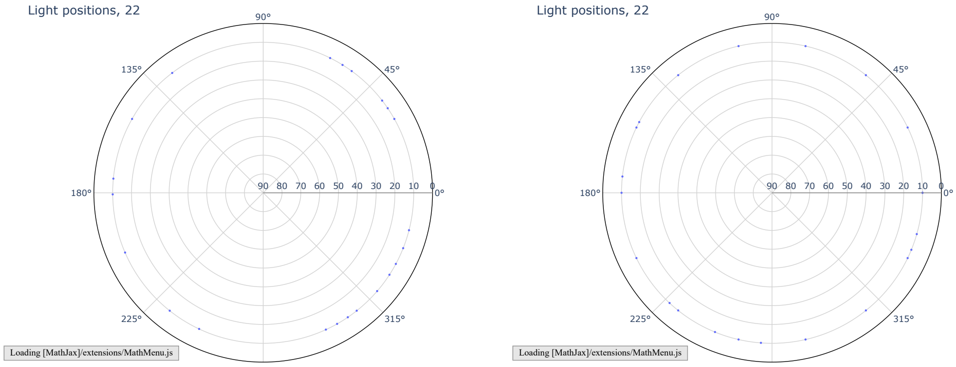


Figure C.22: Overall distribution of the light positions in ideal reference acquisition (left) and NBLP acquisition (right) of canvas painting at elevation 10°

	Positive	Negative	
Positive	11	24	Sensitivity 0.31
Negative	40	285	Specificity 0.87
	Precision 0.21	Negative predictive value 0.92	Accuracy 0.82

Table C.19: Confusion matrix of the results for canvas painting at elevation 10˙

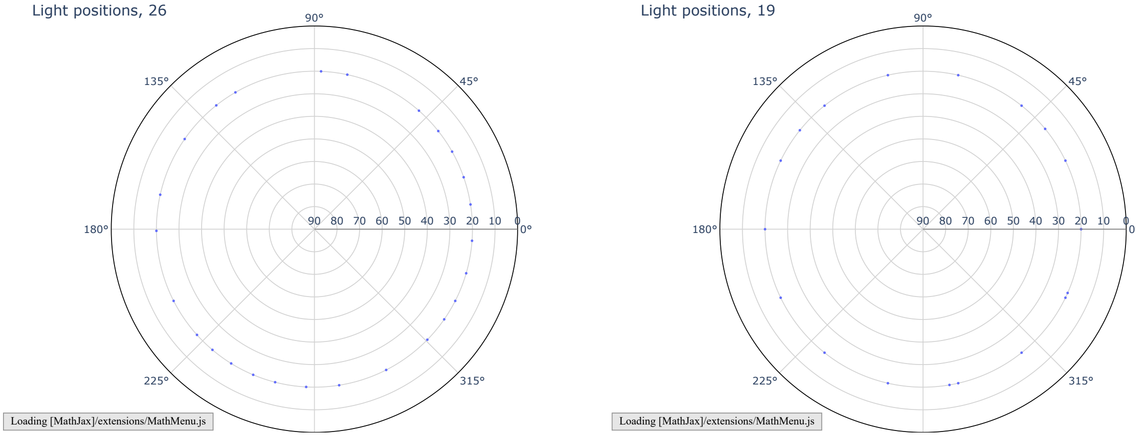


Figure C.23: Overall distribution of the light positions in ideal reference acquisition (left) and NBLP acquisition (right) of canvas painting at elevation 20˙

	Positive	Negative	
Positive	12	34	Sensitivity 0.26
Negative	42	272	Specificity 0.86
	Precision 0.22	Negative predictive value 0.88	Accuracy 0.78

Table C.20: Confusion matrix of the results for canvas painting at elevation 20˙

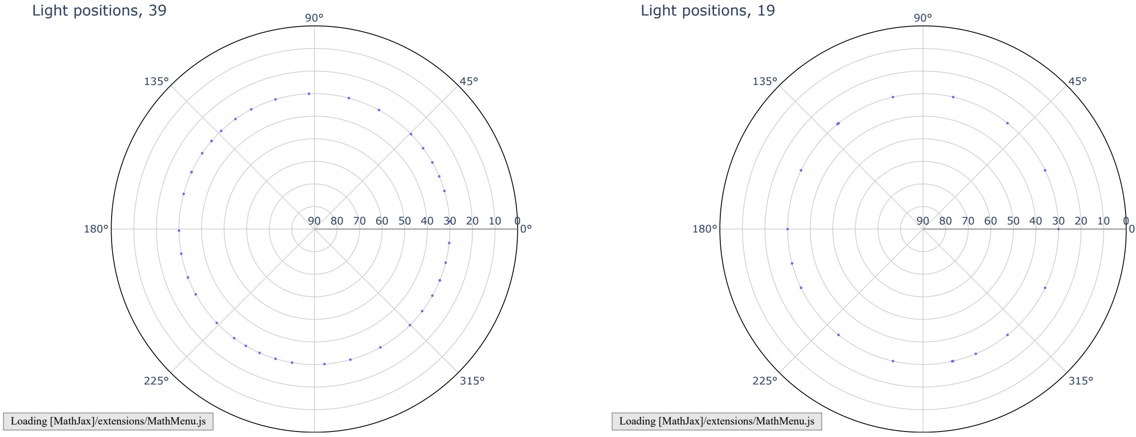


Figure C.24: Overall distribution of the light positions in ideal reference acquisition (left) and NBLP acquisition (right) of canvas painting at elevation 30˙

	Positive	Negative	
Positive	8	60	Sensitivity 0.11
Negative	32	260	Specificity 0.89
	Precision 0.2	Negative predictive value 0.81	Accuracy 0.74

Table C.21: Confusion matrix of the results for canvas painting at elevation 30°

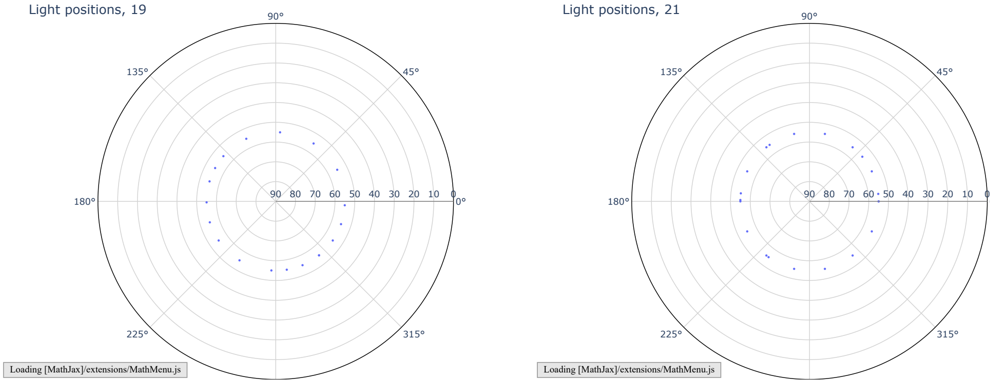


Figure C.25: Overall distribution of the light positions in ideal reference acquisition (left) and NBLP acquisition (right) of canvas painting at elevation 55°

	Positive	Negative	
Positive	3	25	Sensitivity 0.10
Negative	63	269	Specificity 0.81
	Precision 0.04	Negative predictive value 0.91	Accuracy 0.75

Table C.22: Confusion matrix of the results for canvas painting at elevation 55°

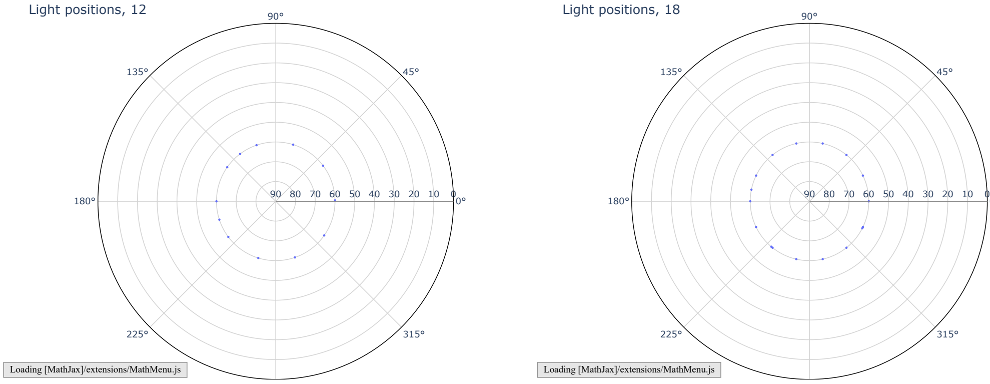


Figure C.26: Overall distribution of the light positions in ideal reference acquisition (left) and NBLP acquisition (right) of canvas painting at elevation 60°

	Positive	Negative	
Positive	5	18	Sensitivity 0.21
Negative	47	290	Specificity 0.86
	Precision 0.09	Negative predictive value 0.94	Accuracy 0.81

Table C.23: Confusion matrix of the results for canvas painting at elevation 60˙

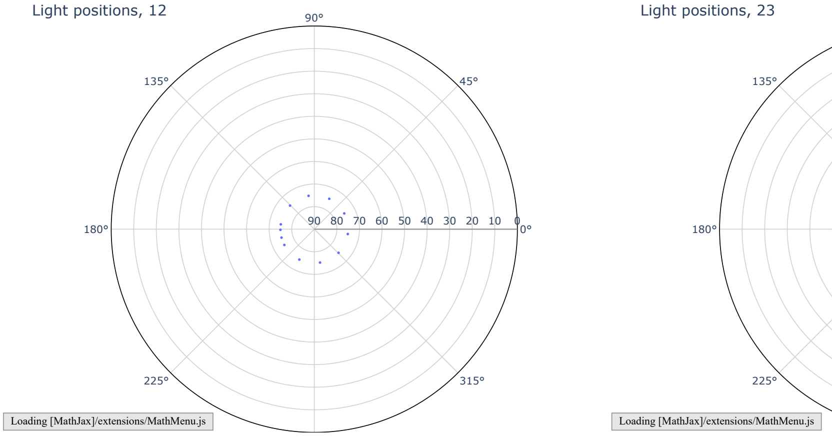


Figure C.27: Overall distribution of the light positions in ideal reference acquisition (left) and NBLP acquisition (right) of canvas painting at elevation 75˙

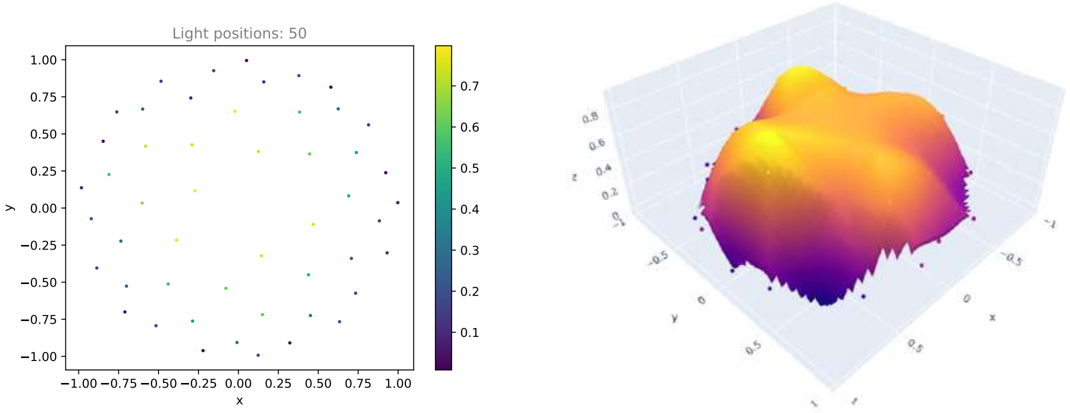
	Positive	Negative	
Positive	9	11	Sensitivity 0.45
Negative	51	289	Specificity 0.85
	Precision 0.15	Negative predictive value 0.96	Accuracy 0.82

Table C.24: Confusion matrix of the results for canvas painting at elevation 75˙

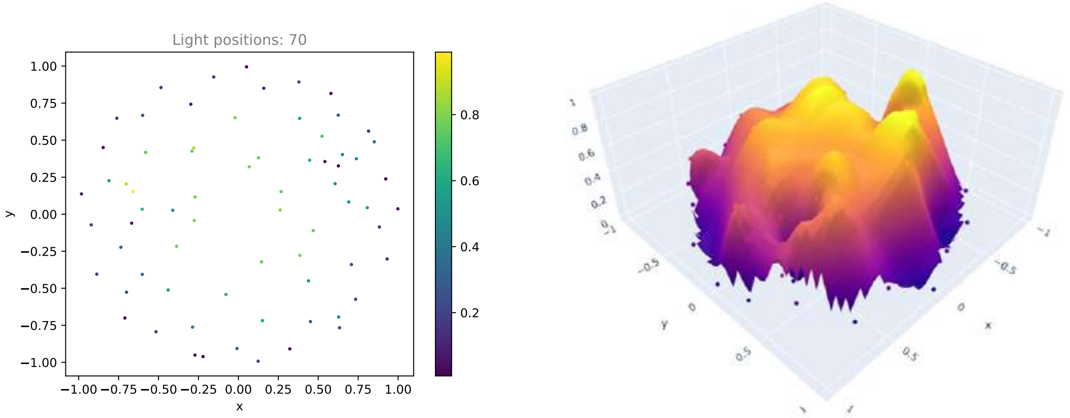
C.2 Results of Azimuth-Elevation space NBLP

Here we present the progresses of the NBLP acquisition in Azimuth-Elevations space for the surfaces. We show the iterations by showing the light positions (on left) and the normalized gradient scores plot (right) in each iteration. Following the iterations, we present the NBLP light positions and the reference ideal light positions side by side for comparison.

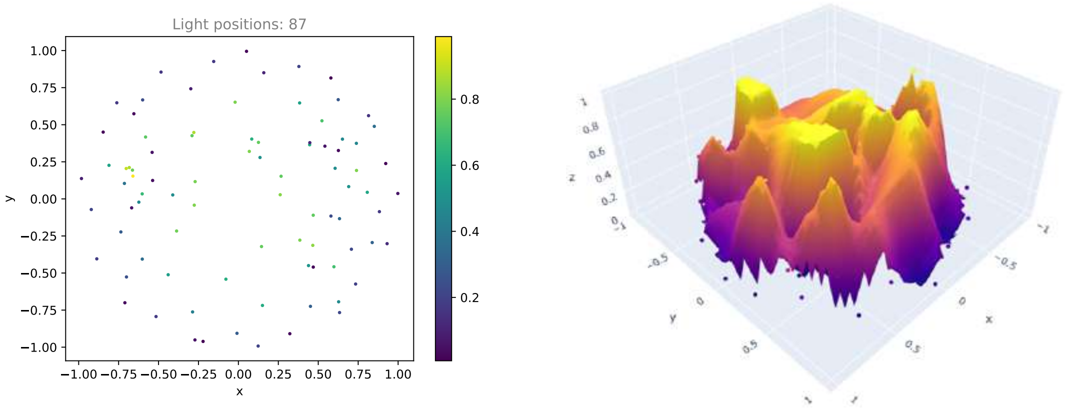
C.2.1 Brushed metal surface with a dent



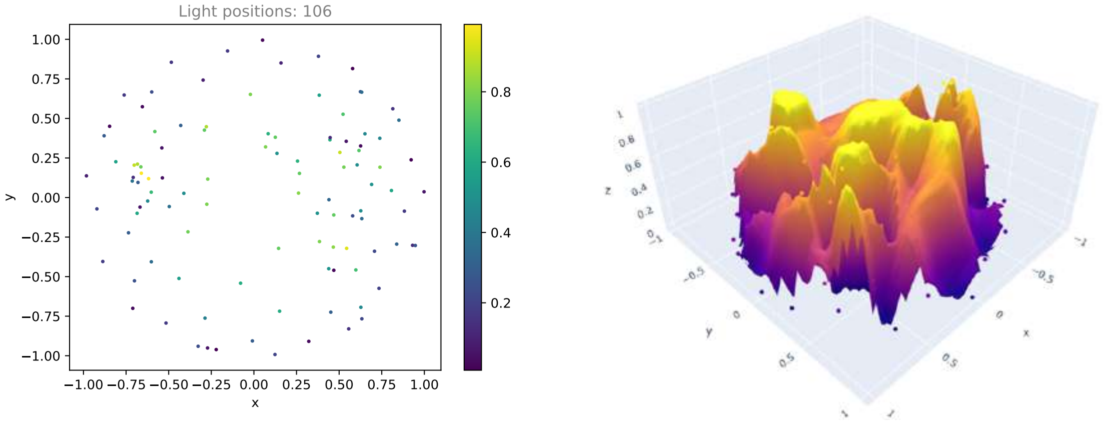
(a) Initial acquisition



(b) 1^{st} iteration. (Left) Light positions, (Right) Normalized gradient scores

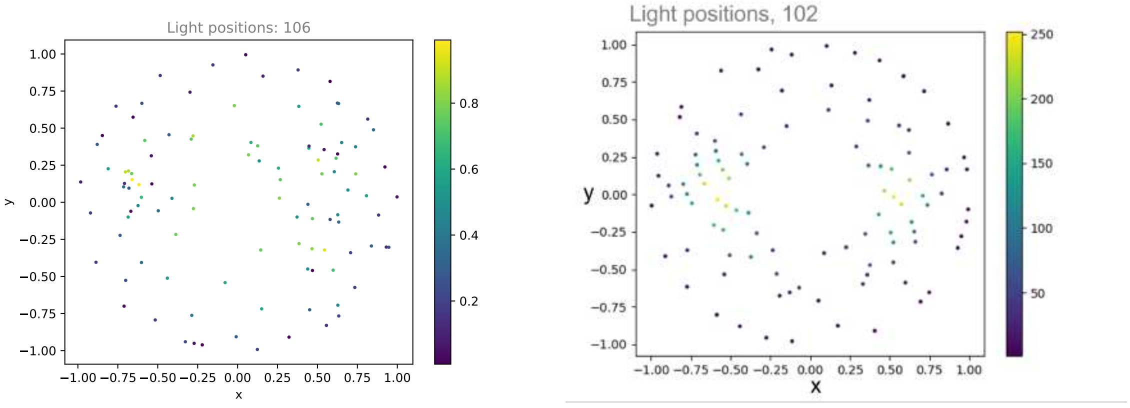


(c) 2^{nd} iteration. (Left) Light positions, (Right) Normalized gradient scores



(d) 3^{rd} iteration. (Left) Light positions, (Right) Normalized gradient scores

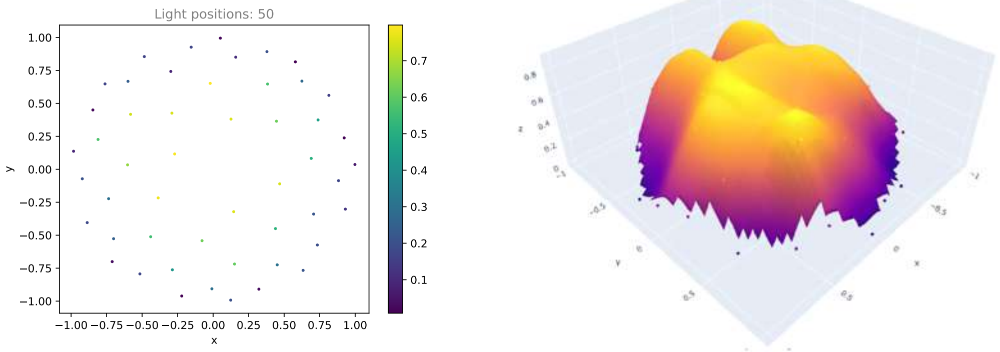
Figure.C.29 compares the light positions estimated by the implemented NBLP method with the best light positions estimated from dense acquisition in Chapter. 4, Figure.4.18. As seen, the light positions are spread such that it is in general towards the ideal light directions but still far from achieving the same result as the reference ideal light directions.



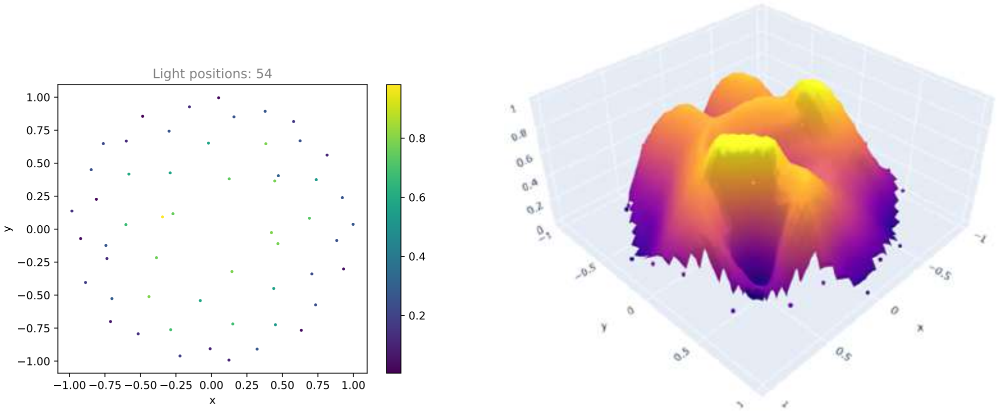
(a) NBLP estimated light position (b) Best light positions derived in chapter. 4

Figure C.29: Comparison of NBLP estimated light positions with that of the best light positions estimated from dense acquisition in chapter. 4 for the brushed metal surface with a dent.

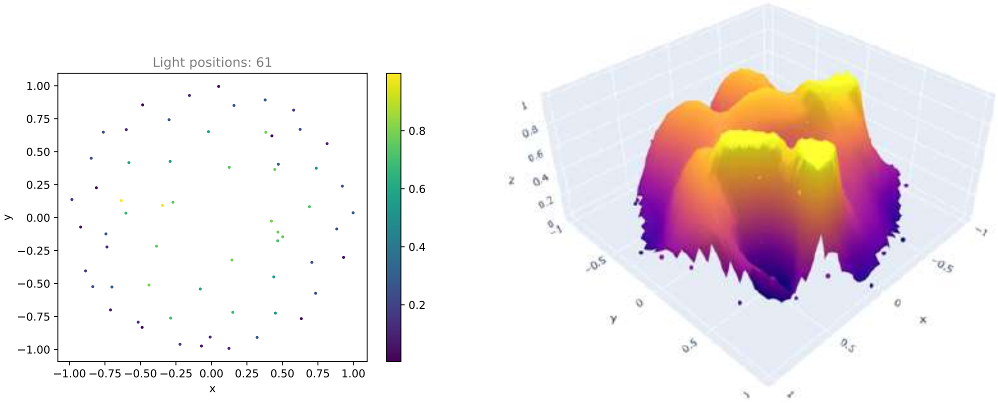
C.2.2 Ancient coin 1



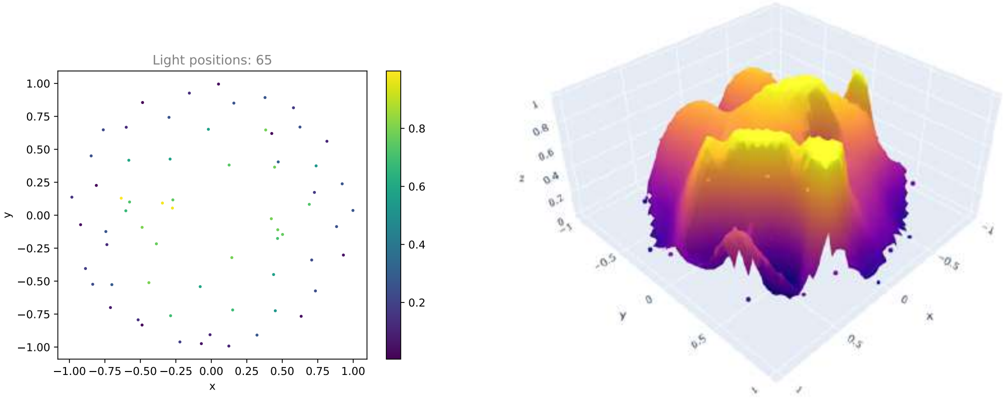
(a) Initial acquisition



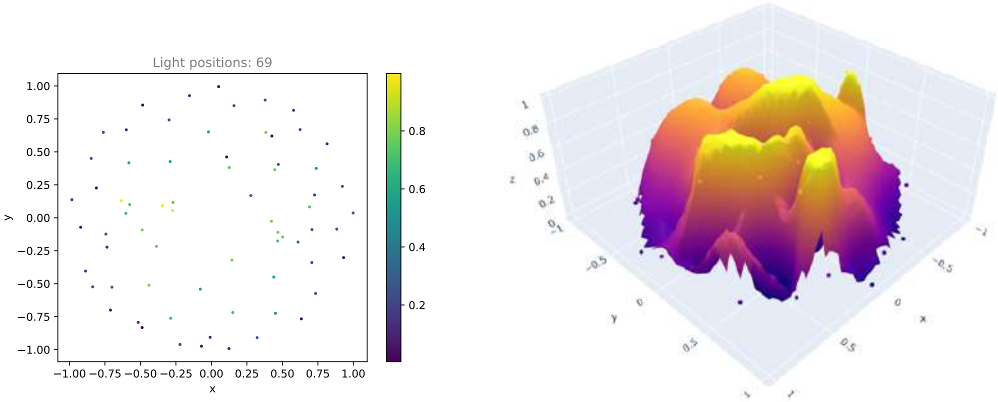
(b) 1^{st} iteration. (Left) Light positions, (Right) Normalized gradient scores



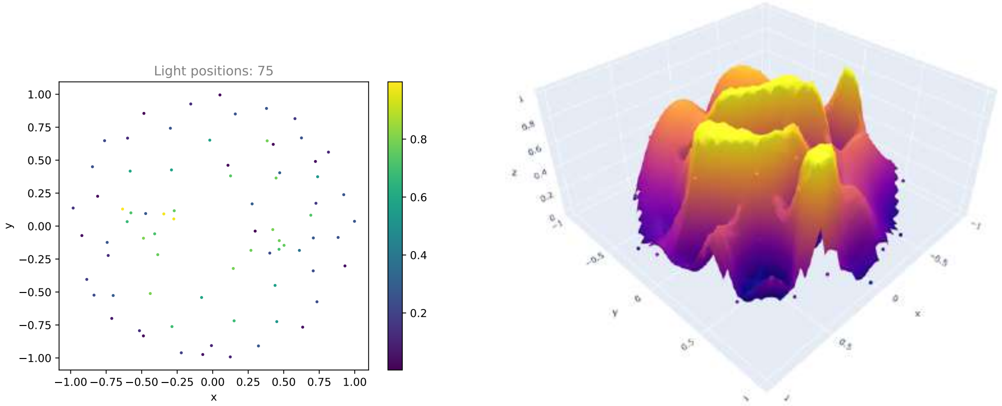
(c) 2^{nd} iteration. (Left) Light positions, (Right) Normalized gradient scores



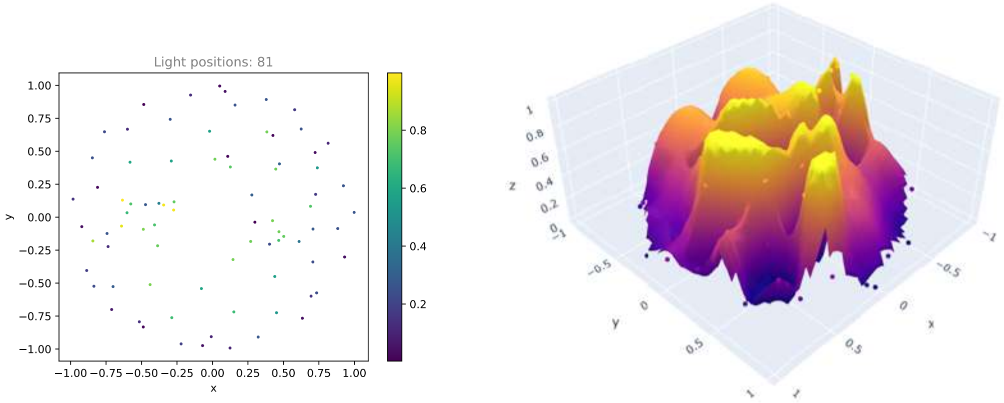
(d) 3^{rd} iteration. (Left) Light positions, (Right) Normalized gradient scores



(e) 4^{th} iteration. (Left) Light positions, (Right) Normalized gradient scores



(f) 5th iteration. (Left) Light positions, (Right) Normalized gradient scores



(g) 6th iteration (termination). (Left) Light positions, (Right) Normalized gradient scores

Figure C.30: NBLP iterations in acquisition of the ancient coin1 surface in the rti-dataset

Figure.C.30 compares the NBLP acquisition result with the reference ideal acquisition light positions for the ancient coin1 surface in the rti-dataset. The NBLP resulted in undersampling for this surface, unlike the redundant sampling for the brushed metal surface. The approximation of linearity in elevation space for correction significantly affects the NBLP estimation for this surface. This approximation suppresses the masking and shadowing of points from it 3D geometry, which can hinder the identification of certain critical light directions.