



HAL
open science

Uncalibrated non-Lambertian photometric stereo. Application to microbial colonies surface reconstruction.

Khrystyna Glazunova Kyrgyzova

► To cite this version:

Khrystyna Glazunova Kyrgyzova. Uncalibrated non-Lambertian photometric stereo. Application to microbial colonies surface reconstruction.. Image Processing [eess.IV]. Université Paris Sud - Paris XI, 2014. English. NNT: 2014PA112155 . tel-01063285

HAL Id: tel-01063285

<https://theses.hal.science/tel-01063285>

Submitted on 11 Sep 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ PARIS-SUD
ECOLE DOCTORALE INFORMATIQUE PARIS-SUD

CEA LABORATOIRE ANALYSE DE DONNÉS ET INTELLIGENCE DES SYSTÈMES

DISCIPLINE : INFORMATIQUE

THÈSE DE DOCTORAT
Soutenue le 22 juillet 2014 par

Khrystyna KYRGYZOVA

**Stéréophotométrie non-calibrée de surfaces
non-Lambertiennes.**

**Application à la reconstruction de surface de
colonies microbiennes**

Directeur de thèse: M. Michaël AUPETIT Ingénieur Chercheur, HDR(CEA/LIST)
Encadrant: Mme. Lorène ALLANO Ingénieur Chercheur (CEA/LIST)

Composition du jury:

Rapporteurs: M. Gilles GESQUIÈRE Prof. (Université de Lyon 2/LIRIS)
M. Adrien BARTOLI Prof. (Université d'Auvergne/ISIT/ALCoV)
Président du jury : Mme. Anne VILNAT Prof. (Université Paris-Sud/LIMSI)
Examineur: M. Olivier AUBRETON MCF, HDR (Université de Bourgogne/LE2I)
Invité: M. Guillaume Perrin PhD., Ingénieur (bioMérieux)

There are no facts, only interpretations.

Friedrich Nietzsche

Acknowledgements

I would like to express my very great appreciation to my direct mentor Lorène Allano. Enthusiasm, energy, and diligence of this person inspired me all the time. Her contribution to this work as well as to my professional development is immeasurable.

I am grateful to Marine Depecker, Guillaume Sutra and Jérôme Gauthier that they kindly agreed to help me during my third and the most intensive year.

Also I would like to thank to my thesis advisor Michaël Aupetit provided me with valuable advises, remarques and suggestions.

I wish to acknowledge the help and the working atmosphere provided by my project colleges at the CEA and at bioMérieux Company. Particularly, I would like to thank bioMérieux company for funding this work and for the data providing.

Special thanks should be given to all my colleges from the lab CEA/LIST/LADIS with whom I was lucky to communicate each day.

Also I wish to thank to all my relatives and friends living in Ukraine and in France for their concern. I am particularly grateful for the assistance given by my parents. Thanks to their efforts and contribution to my education, I had an amazing possibility to write this work.

And my last and the most important acknowledgement I would like to express to my husband Olexiy who limitlessly supported me all these dynamic and interesting three years.

Abstract

Title: Uncalibrated non-Lambertian photometric stereo.
Application to microbial colonies surface reconstruction.

The PhD thesis work is dedicated to the problem of uncalibrated non-Lambertian photometric stereo surface reconstruction. The proposed approach consists in two phases: first we correct images of the input sequence from specularities in order to obtain images of pseudo-Lambertian surfaces, and then realize Lambertian photometric stereo reconstruction. In this work we proposed two original methods, respectively, for specularity correction and surface reconstruction with no prior information neither on light sources nor on surface properties. We apply the novel processing to Petri dish images for microbial colonies surface reconstruction.

Specularity is an optical effect of a complex physical nature. This effect is useful for human 3D objects perception but it affects automated image processing. In order to be able to apply the Lambertian photometric stereo model, specularities should be removed from the input images. We propose an original method for specular zones correction adapted to estimation of pseudo-Lambertian surface images and further reconstruction. This algorithm is able to detect specularities as abnormally elevated pixel intensity values in an image of the input sequence and to correct the found zones using information from all other images of the sequence and a specific continuous correcting function. This method allows removing specularities while still preserving all other particularities of shading important for the further surface reconstruction.

We then propose an original stereo photometric method for Lambertian surface reconstruction with no prior on illuminations. The implemented photometric stereo model consists of four components, two of them (albedo and normals) describe surface properties and the others (light sources intensities and directions) describe illumination. The proposed algorithm of the photometric stereo reconstruction uses the alternating optimization principle. Each model component is found iteratively fixing all variables but one and applying value and quality constraints for the optimization function. The original scheme of resolution allows separating of different information types included in input images. Thanks to such matrix factorization, the surface reconstruction is made with no prior information on lighting directions and the reconstructed objects properties. The applicability of the algorithm is proved using artificially created and open data-sets for which the ground truth information is available.

The last part of the thesis is dedicated to the application of the proposed uncalibrated non-Lambertian photometric stereo approach to the Petri dish images. Images are obtained using illuminating sources which are supposed to be unknown for the reconstruction process. Moreover, the reconstructed microbial colonies are very diverse, generally have small size, can be Lambertian or not, and their surface properties are not defined in advance. The results of reconstruction for such complex real-world data add value and importance to the developed approach.

Key-words: photometric stereo surface reconstruction, non-Lambertian reflectance models, specularity removal, alternating constrained optimization, surface recognition, bio-medical image processing.

Résumé de thèse

Titre: Stéréophotométrie non-calibrée de surfaces non-Lambertiennes.
Application à la reconstruction de surface de colonies microbiennes.

Cette thèse est dédiée à la résolution du problème de la stéréophotométrie non-Lambertienne sans connaissance a priori sur les conditions d'illumination et son application aux images de boîte de Pétri. Pour obtenir une bonne reconstruction de surfaces non-Lambertiennes, il est proposé de traiter une séquence d'entrée en deux étapes: premièrement il faut supprimer les effets spéculaires et obtenir ainsi des images de surface 'pseudo-Lambertienne'. Dans une deuxième étape une reconstruction stéréophotométrique Lambertienne sans aucune information préalable sur les directions d'illumination est effectuée à partir de ces images. Dans ce travail, nous proposons deux méthodes originales pour la suppression de spécularités et la reconstruction de surface sans information a priori, respectivement. Les méthodes proposées sont appliquées pour la caractérisation des colonies microbiennes.

La spécularité est un effet optique lié à la nature physique complexe des objets. Celui-ci est utile pour la perception humaine des objets 3D mais gêne le processus automatique de traitement d'images. Pour pouvoir appliquer le modèle Lambertien à la stéréophotométrie, les spécularités doivent être supprimées des images d'entrée. Nous proposons donc une méthode originale pour la correction des zones spéculaires adaptée à la reconstruction ultérieure. L'algorithme proposé est capable de détecter les spécularités comme des valeurs anormalement élevées d'intensité dans une image de la séquence d'entrée, et de les corriger en utilisant les informations des autres images de la séquence et une fonction de correction continue. Cette méthode permet de faire la suppression des spécularités en préservant toutes les autres particularités de distribution de lumière qui sont importantes pour la reconstruction de surface.

Nous proposons ensuite une technique de reconstruction stéréophotométrique de surfaces Lambertiennes sans connaissance a priori sur l'illumination. Le modèle mis en œuvre consiste en quatre composantes, deux composantes (albédo et normales) permettent de décrire des propriétés de surface et deux autres (intensités des sources de lumière et leurs directions) décrivent l'illumination. L'algorithme de reconstruction proposé utilise le principe de l'optimisation alternée. Chaque composante du modèle est trouvée itérativement en fixant toutes les autres variables et en appliquant des contraintes de structure, de valeurs et de qualité de la fonction d'optimisation. Un schéma original de résolution permet de séparer les différents types d'information inclus dans les images d'entrée. Grâce à cette factorisation de matrices, la reconstruction de surface est faite sans connaissance préalable sur les directions de lumière et les propriétés de l'objet reconstruit. L'évaluation de l'algorithme était effectuée pour des données artificielles et des images de bases publiques pour lesquelles la vérité terrain sur les surfaces des objets est disponible.

La dernière partie de la thèse est dédiée à l'application de la chaîne de traitement complète proposée pour les images de boîte de Pétri. Ces images sont obtenues en utilisant des sources de lumière complexe supposées inconnues pour le processus de reconstruction. L'évaluation de surfaces de colonies microbiennes s'est révélée être une étape importante pour l'analyse visuelle et automatique des colonies. La chaîne proposée est efficace pour ce type de données et permet de compléter des informations d'images par de la surface 3D.

Mots clefs: stéréophotométrie, modèles de réflexion non-Lambertiens, suppression de spéularités, optimisation alternée sous contraintes, reconnaissance de surface, traitement d'images biomédicales.

Contents

Acknowledgements	v
Abstract	vii
Résumé de thèse	ix
Abbreviations	xv
Notations	xvii
List of Figures	xix
List of Tables	xxiii
List of Algorithms	xxv
1 Introduction	2
1.1 'Shape-from-X' approaches	4
1.1.1 Passive methods	5
1.1.2 Active methods	11
1.2 Problem Statement	17
1.3 Thesis Contributions	18
1.4 Thesis Outline	19
2 Non-Lambertian surface reconstruction problem	22
2.1 Light-Surface-Camera interaction	22
2.1.1 Light source	23
2.1.2 Camera	23
2.1.3 Surface	24
2.2 Reflectance of opaque surfaces	31
2.2.1 Dichromatic reflectance model	31
2.2.2 Bidirectional reflectance distribution function	35
2.3 Specular surface reconstruction	40
2.3.1 Specular surface reconstruction methods	40
2.3.2 Outlier specular pixels	47
2.4 Proposed chain for uncalibrated non-Lambertian photometric stereo	51
2.4.1 Specularity correction	52

2.4.2	Uncalibrated photometric stereo reconstruction	53
3	Specularity correction	54
3.1	Specularity correction techniques	55
3.1.1	' $N - to - 1$ ' methods	56
3.1.2	' $1 - to - 1$ ' methods	57
3.1.3	' $N - to - N$ ' methods	58
3.1.4	Evaluation of specularity correction	59
3.2	Proposed specularity correction algorithm	61
3.2.1	Specularity evaluation	62
3.2.2	Strict specularity correction	65
3.2.3	Soft specularity correction	68
3.3	Influence of the initial images	72
3.3.1	Distribution of the illumination sources	72
3.3.2	The number of reference images	73
3.4	Experimental results for real-world images	76
3.5	Conclusions	78
4	Uncalibrated photometric stereo	80
4.1	Lambertian photometric stereo reconstruction	81
4.1.1	Basic photometric stereo model	81
4.1.2	Photometric stereo via matrix factorization	84
4.1.3	Uncalibrated photometric stereo via alternating optimization	86
4.1.4	Bas-relief ambiguity	89
4.1.5	Depth map reconstruction from normals	91
4.2	Proposed method of uncalibrated PS using alternating optimization	98
4.2.1	General model description	98
4.2.2	Alternating optimization scheme for photometric stereo	100
4.2.3	Constrained alternating optimization	109
4.2.4	Stopping criteria	117
4.2.5	Final reconstruction algorithm summary	121
4.3	Lambertian objects reconstruction	121
4.3.1	Separation of information	122
4.3.2	Influence of the spatial light sources distribution on reconstruction	123
4.3.3	Influence of the light sources number on reconstruction	125
4.4	Non-Lambertian objects reconstruction	127
4.5	Conclusions	129
5	Application to microorganisms surface reconstruction from Petri dish images	132
5.1	Work motivations	132
5.2	Imaging device	135
5.3	Microbial colonies surface reconstruction	138
5.3.1	Global Petri dish reconstruction	138

5.3.2	Microbial colonies surface reconstruction	139
5.3.3	Importance of specularities correction	142
5.3.4	Lambertian colonies surface reconstruction	144
5.3.5	Examples of different colonies reconstruction	145
5.4	Conclusions	149
6	Conclusions and Perspectives	150
6.1	Conclusions	150
6.2	Future Work	152
A	Depth map from normals reconstruction	154
B	Photometric stereo model without constraints	157
B.1	Albedo values matrix evaluation	157
B.2	Light sources intensities vector evaluation	158
B.3	Lighting sources directions matrix evaluation	159
B.4	Surface normals matrix evaluation	160
C	Photometric stereo model with constraints	161
C.1	Albedo values approximation with interval constraint	161
C.2	Light sources intensities approximation with non-negativity constraint	163
C.3	Lighting sources directions approximation with norm constraint . . .	165
C.4	Surface normals matrix approximation with norm, direction and smooth- ness constraints	167
D	The Levenberg-Marquardt Algorithm	170
	Bibliography	172

Abbreviations

AO	- Alternating optimization.
ALS	- Alternating least squares.
BRDF	- Bidirectional reflectance distribution function.
BTDF	- Bidirectional transmittance distribution function.
BSSRDF	- Bidirectional surface scattering reflectance distribution function.
CCD	- Charge-coupled device.
CMOS	- Complementary metal-oxide-semiconductor.
KKT	- Karush-Kuhn-Tucker conditions.
LED	- Light-emitting diode.
LSM	- Least squares method.
MSE	- Mean square error.
PS	- Photometric stereo.
RGB	- Red, green, blue color space.
RMSE	- Root mean square error.
SURF	- Speeded up robust features.
SVD	- Singular value decomposition.

Notations

(θ_i, ϕ_i)	- incident light direction in spherical coordinates.
(θ_o, ϕ_o)	- outgoing light direction in spherical coordinates.
$d\omega_i$	- differential solid angle of the incident light.
$d\omega_o$	- differential solid angle of the outgoing light.
L_i	- quantity of the incoming light.
L_o	- quantity of the outgoing light.
E_i	- quantity of light which reaches the camera.
ρ	- scalar value of albedo.
\mathbf{n}	- surface normal vector in 3D.
\mathbf{v}	- light source direction vector in 3D.
\mathbf{s}	- camera viewing direction vector in 3D.
\mathbf{r}	- direction vector of the perfect specular reflection in 3D.
\mathbf{I}	- matrix of image intensity values.
$\tilde{\mathbf{I}}$	- matrix of image intensity values with corrected specularities.
\mathbf{I}_{Spec}	- matrix of image intensity values of the specular component.
\mathbf{I}_{Diff}	- matrix of image intensity values of the diffuse component.
\mathbf{M}	- matrix of vectorized images intensity values.
\mathbf{A} and $\tilde{\mathbf{A}}$	- matrices of albedo values.
\mathbf{L} and $\tilde{\mathbf{L}}$	- matrices of light sources intensities.
\mathbf{V} and $\tilde{\mathbf{V}}$	- matrices of light sources directions.
\mathbf{N}	- matrix of surface normals.
\mathbf{W}_A	- binary structure matrix for albedo.
\mathbf{W}_L	- binary structure matrix for light sources intensities.
\mathbf{W}_V	- binary structure matrix for light sources directions.
\mathbf{Z}	- depth map.

List of Figures

1.1	Example of Petri dish image with different illumination	3
1.2	Digital image acquisition process	4
1.3	Classification of the existing 'Shape-from-X' approaches	5
1.4	Passive stereo and triangulation principle	6
1.5	Slant and tilt angles	7
1.6	The Gaussian lens law	8
1.7	Shape-from-(de)focus	9
1.8	Shape-from-silhouettes	10
1.9	Active stereo reconstruction	12
1.10	Surface reconstruction using shading variation	13
1.11	Surface gradient space and reflectance map representation	15
2.1	Interaction of the incident light with the object surface	24
2.2	Shape-from-distortion for water surface	27
2.3	Transparent surface reconstruction by the structured light and shape-from-silhouettes	27
2.4	Specular effect formation	32
2.5	Specular component distribution models	33
2.6	Color formation due to the dichromatic reflection model	34
2.7	The BRDF reciprocity principle.	37
2.8	The BRDF energy conservation principle	38
2.9	Non-Lambertian and Lambertian surface BRDFs	38
2.10	Specular motion stereo	42
2.11	Shape-from-distortion setup	44
2.12	6-source PS with specularities and shadows discarding	50
2.13	Proposed uncalibrated non-Lambertian surface reconstruction	52
3.1	MSE results for specularity correction	60
3.2	Example of the specular image to be corrected and reference images	64
3.3	Ratio aggregation: mean vs median	64
3.4	Diffuse and specular component of the artificial image	66
3.5	Superposition of specular zones	66
3.6	1D example of the strict threshold	67
3.7	Example of the strict specular zone correction	67
3.8	Sigmoid function values depending on three parameters	69
3.9	Example of soft specularity correction	71

3.10	Specularity overlapping cases	73
3.11	Example of the specular cat statue images	77
3.12	Initial image of the specular cat's eye	77
3.13	Example of the specular cat's eye correction	78
3.14	Example of the specular fish correction	79
4.1	Alternating optimization general principle	87
4.2	A patch of depth map with appropriate values of surface normals.	92
4.3	Comparison of depth map reconstruction algorithms	97
4.4	The proposed PS model matrix representation	98
4.5	The proposed PS model matrix representation with structure constraints	99
4.6	Input images for initialization tests	100
4.7	Initialization of the normals map	101
4.8	Initialization of the matrix of light source intensities	102
4.9	Initialization of the matrix of albedo values	103
4.10	Testing of different alternating optimization algorithm configurations	104
4.11	Alternating optimization for the PS model evaluation	105
4.12	Reconstruction of an artificial hemisphere using AO without constraints	108
4.13	Initial images of a complex surface	108
4.14	Reconstruction of a complex object using AO without constraints	109
4.15	Post-implementation of value constraints	111
4.16	Embedded value constraints	114
4.17	Quality and value constraints post-implementation	116
4.18	Embedded quality and value constraints	117
4.19	Investigation of internal stopping criteria	119
4.20	Evolution of model and normals error for 10 iterations	120
4.21	Input images of folded paper used for the surface reconstruction	122
4.22	Demonstration of information separation by the proposed algorithm	123
4.23	Convergence of errors for the separation of information experiment	123
4.24	Images for the demonstration of the input data influence	124
4.25	Influence of the light sources distribution	125
4.26	Influence of the light sources number	126
4.27	Algorithm convergence and surface error depending on the input images	127
4.28	Artificial specular sphere images	128
4.29	Diffuse component of the artificial specular sphere images	128
4.30	Corrected artificial specular sphere images	129
4.31	Reconstructed surface of specular sphere	129
5.1	Examples of Petri dishes with different media and grown colonies	133
5.2	Visual analysis of a Petri dish with microbial colonies	134
5.3	The bioMérieux imaging device for the Petri dishes	136
5.4	The bioMérieux imaging device details	136
5.5	Examples of the Petri dish images with different illumination	137
5.6	RGB image with left side bottom illumination	137
5.7	Diversity of media and microbial colonies	138

5.8	Global Petri dish reconstruction	139
5.9	Processing chain for microorganisms surface reconstruction	140
5.10	The chosen image region inside the Petri dish to be processed	140
5.11	Image region with masked colony and masked background	141
5.12	Partial colony distortion related to specularities	143
5.13	Important colony distortion related to specularities	143
5.14	Lambertian surface reconstruction after specularities correction	144
5.15	Example 1 of the reconstructed small colony	146
5.16	Example 2 of the reconstructed small colony	146
5.17	Example of the reconstructed concave colony	147
5.18	Example 1 of the reconstructed colony of sophisticated shape	148
5.19	Example 2 of the reconstructed colony of sophisticated shape	148
5.20	Example 3 of the reconstructed colony of sophisticated shape	148

List of Tables

3.1	Specularity correction methods classification	59
3.2	Specular zones overlapping	74
3.3	Specularity soft correction depending on light source repartition in space	75
4.1	Global stopping criteria comparison	120

List of Algorithms

1	Strict speculariry correction	66
2	Soft speculariry correction	70
3	The model components evaluation by the AO without constraints . .	106
4	The model components evaluation by the AO with post-implementation of constraints	115
5	The model components evaluation by the AO with embedded constraints	118
6	Depth map from normals reconstruction	156

Chapter 1

Introduction

As we live in a 3D world, our brain is capable to interpolate the missing information on a 2D image in order to analyze the object surface. But this represents a complex issue to be automated. Researchers in image processing have been interested in reproducing this brain ability for many years. Reconstruction of 3D surfaces is used in many application fields, such as augmented reality, cinema, biomedical imaging, face recognition, etc. These techniques have become more popular from year to year with capturing devices evolution. The third dimension allows not only 3D visualization, but also the object surface features extraction.

Due to the popularity of the 3D technologies, many approaches of the surface reconstruction have been already proposed. 'Shape-from-X' is a generalized name for all existing algorithms which aim at 3D surface reconstruction from digital image(s). Some of these methods are widespread and well known (e.g., a multi-view stereo reconstruction), others are less popular and need specific acquisition devices alone with dedicated mathematical algorithms. The chosen method for 3D reconstruction is often limited by the acquisition device and physical properties of the reconstructed object.

The present thesis is part of a collaborative research project between CEA LIST and bioMérieux company, [1]. The main purpose of this project is to automate the process of microbiological analysis. One possibility to observe and study bacteria is to make them grow inside a special glass or plastic dish, called a Petri dish. After a certain time of growing inside the dish, microbial colonies can be analyzed visually by a technician. This analysis is very important for diagnosis, but is time-consuming and requires rigorous and continuous work from microbiologists. From the point of

view of information technology, the question arises on how to automate this complex process of manual and visual analysis, in order to provide the technicians with a computer-based decision support system. In this context, bioMérieux company, has developed a dedicated imaging device aiming at facilitating microbiological analysis, [2]. This device allows taking images of the Petri dish with different illumination conditions for a fixed position of a camera. Examples of these images are presented in Fig. 1.1. After that, microbial colonies analysis can be automated thanks to image processing algorithms. Such an automation includes tasks of different diagnostic interest such as microbial colonies characterization or enumeration.

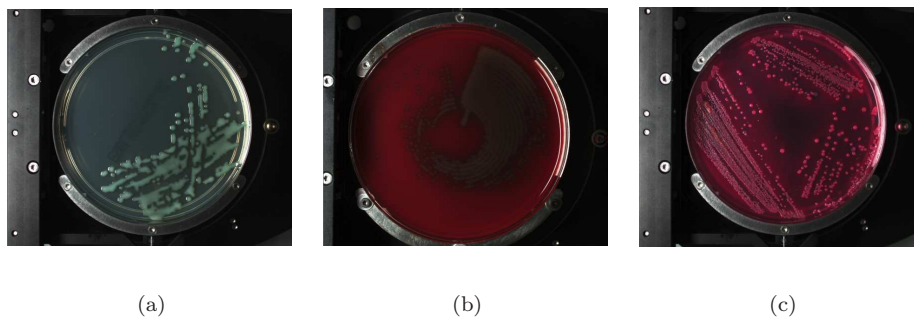


FIGURE 1.1: Example of Petri dish image with different illumination.

Microbial colony surface is also of high interest for microbiologist. The present thesis work is dedicated to the task of microbial colonies surface reconstruction from the Petri dish images. Our purpose is to propose an efficient algorithm of 3D reconstruction applicable to objects with complex optical properties, such as microorganisms, using the data of the existing device. Moreover, the result of such an algorithm has to be easy to exploit for a further surface analysis. Thus, the developed reconstruction technique is limited by the structure of the already existing device and particularities of the microbial colonies (e.g., small size, specular surface, etc.).

In our work we deal with an industrial system, structure of which can evolve with new requirements, development of lighting or capturing technologies, etc. One of the goals is to use as few prior knowledge on the acquisition device as possible. On the other hand, it is not known in advance which type of bacteria will grow inside a dish. Thus, the properties of the reconstructed surface are not pre-defined as well.

Taking into account the limitations and particularities of the task, we are going, first of all, to analyze the existing methods of the 3D reconstruction in order to choose the

techniques which can be potentially used for the microbial colonies reconstruction in our context.

1.1 'Shape-from-X' approaches

The history of the 3D technologies begins in 1844 when David Brewster introduced the Stereoscope, his device to take stereo images. After this important historical event, scientists attempt to better understand 3D world representation by a 2D projection to an image plane. With the advent of digital photos, the process of image acquisition can be simplistically seen as it is presented in Fig. 1.2. A camera lens which defines an optical image plane receives light reflected from a 3D surface. It is supposed that there is no optical distortion determined by the lens structure (orthographic projection), and all straight lines in the real world are transformed to straight lines in the image. After that, the optical image is converted to an electrical signal. Level of this signal depends on the quantity of photons reflected from the photographed object and captured by a matrix of sensors (e.g., CCD, CMOS, etc.). The sensor discretizes the input signal into 2^p levels (from 0 to $2^p - 1$), where p is a number of bits per a memory cell. Each sensor of the matrix corresponds to the resulting digital image pixel. The obtained digital image is analyzed and processed using different computer vision and image understanding algorithms in order to extract necessary information. For the case of surface reconstruction, the sought information is the missing third dimension of the photographed objects.

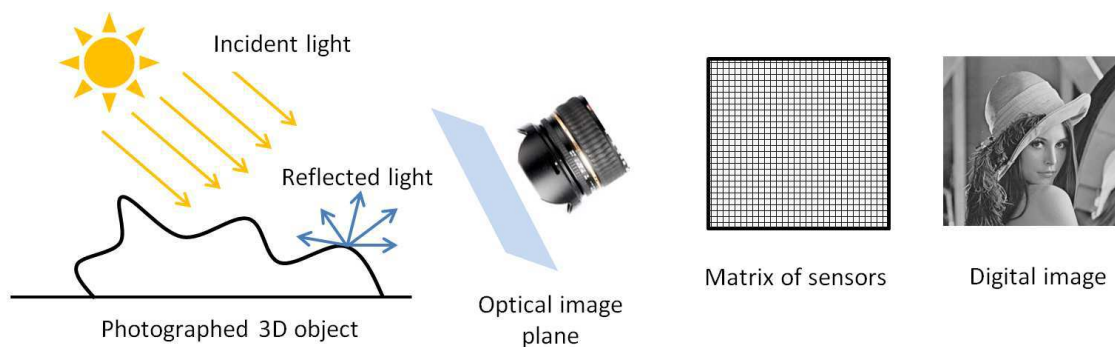


FIGURE 1.2: Digital image acquisition process.

Existing 'shape-from-X' techniques may be classified as proposed in [3] and presented in Fig. 1.3, according to the presence of controlled illumination sources (active) or without them (passive). Another criterion of classification is a single or multiple cameras positions (vantage points) usage for algorithms input images acquisition.

Active techniques usually need special devices for the acquisition. On the other hand, algorithms with multiple vantage points require the input data volume several times higher than algorithm with a single vantage point. The presented scheme is one of the possible classifications, it reflects the opinion of the authors of [3] which we share, but any other classification of the presented methods is possible. We present the most widespread 'shape-from-X' techniques (in red in Fig. 1.3), analyze these techniques advantages, deficiencies and needs in particular data or acquisition devices. Here we aim at general principles presentation with appropriate references for more detailed reviewing.

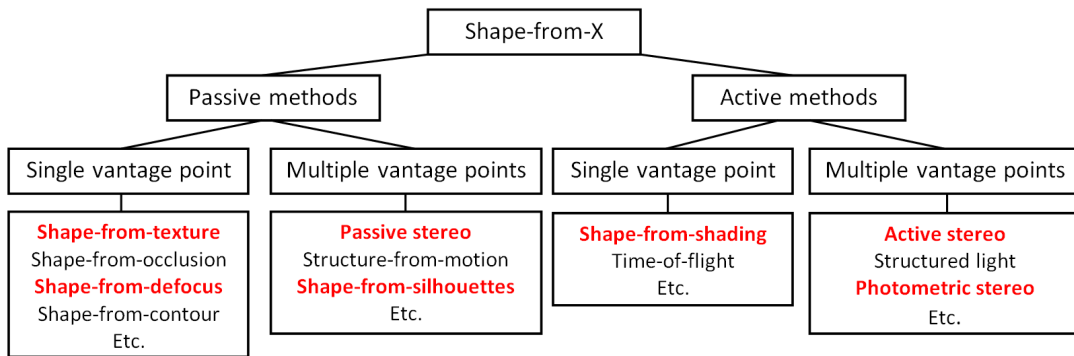


FIGURE 1.3: Classification of the existing 'Shape-from-X' approaches.

1.1.1 Passive methods

Passive shape-from-stereo

This is the oldest and the mostly used method of surface reconstruction. It is based on the geometrical principle of triangulation and can be applied to images taken from several vantage points. In the classical version of the passive stereo, surface reconstruction is made on the basis of two images taken at the same time from two different viewing positions, Fig. 1.4. It is called a stereo setup. Thus, the 3D point position is found as intersection of two projections of the same point s to the planes of two images. After that, the following simple geometrical principle is used (triangulation):

$$\begin{aligned}
 l &= \frac{d}{\tan(\alpha)} + \frac{d}{\tan(\beta)}, \\
 \frac{1}{d} &= \frac{1}{l} \left(\frac{1}{\tan(\alpha)} + \frac{1}{\tan(\beta)} \right), \\
 \frac{1}{d} &= \frac{\sin(\alpha+\beta)}{l \sin(\alpha)\sin(\beta)}.
 \end{aligned} \tag{1.1}$$

The usage of the described triangulation principle requires knowledge about the cameras, such as their relative position (l), orientations (α and β), and also focal lengths, as presented in Fig. 1.4. Special process of these parameters estimation is called calibration, it is required before surface reconstruction.

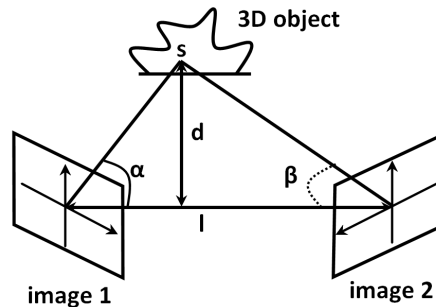


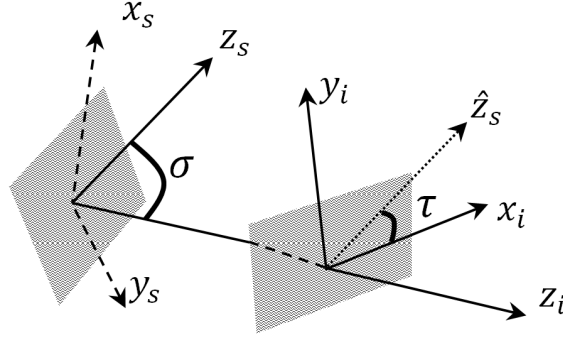
FIGURE 1.4: Passive stereo and triangulation principle.

Another important problem, which should be solved to use the passive stereo, is the point correspondence problem. All points of the second image which correspond to specific points of the first image should be found, and vice versa. This is the most difficult challenge of the passive stereo. Analysis and comparison of existing local and global correspondence methods are given in [4]. All of these methods are rather complex and not universal for all types of images.

For the algorithms shape-from-motion (or structure-from-motion) initial images are taken in a sequence from many spatially close positions. If images are taken with short time intervals, correspondence between similar points are easier to be found. Moreover, this approach does not need cameras calibration procedure as for the classical passive stereo.

Shape-from-texture

Gibson was the first to use the term texture gradient in order to show that parts of the same texture are perceived differently according to their various orientations and distances to the observer, [5]. Any texture can be divided into texels (patterns with similar statistical or geometrical characteristics). The original texture should be maximally isotropic, so local inhomogeneities of the texture are considered as changes in surface orientation. Slant (σ) and tilt (τ) angles of each texel defines surface orientation for the image region occupied by the texel, Fig. 1.5. Slant denotes the amount and tilt denotes the direction of the slope of the planar surface (x_s, y_s) projected to the image plane (x_i, y_i).

FIGURE 1.5: Slant (σ) and tilt (τ) angles.

In order to better understand the principle of shape-from-texture reconstruction let us present it in terms of the article [6]. The affine transformation $T_{f \rightarrow i}$ of the frontal texel f to the texel i can be written as the following Singular Value Decomposition (SVD):

$$T_{f \rightarrow i} = \mathbf{U}\mathbf{D}\mathbf{V}^T, \quad (1.2)$$

where \mathbf{U} and \mathbf{D} defines tilt and slant directions respectively. These matrices can be rewritten as:

$$\mathbf{U} = \begin{pmatrix} \mathbf{U}_{1,1} & \mathbf{U}_{1,2} \\ \mathbf{U}_{2,1} & \mathbf{U}_{2,2} \end{pmatrix} = \begin{pmatrix} \cos(\tau) & \sin(\tau) \\ -\sin(\tau) & \cos(\tau) \end{pmatrix}, \quad (1.3)$$

$$\mathbf{D} = \begin{pmatrix} \mathbf{D}_{1,1} & \mathbf{D}_{1,2} \\ \mathbf{D}_{2,1} & \mathbf{D}_{2,2} \end{pmatrix} = \begin{pmatrix} r \cos(\sigma) & 0 \\ 0 & r \end{pmatrix}, \quad (1.4)$$

where r is a scaling factor. The surface normal in 3D space then is evaluated as:

$$\begin{pmatrix} n_x \\ n_y \\ n_z \end{pmatrix} = \begin{pmatrix} \pm \cos(\tau) \sin(\sigma) \\ \pm \sin(\tau) \sin(\sigma) \\ \cos(\sigma) \end{pmatrix}. \quad (1.5)$$

Tilt and slant angles can be respectively evaluated as the followings:

$$\tau = \arctan(\mathbf{U}_{1,2}/\mathbf{U}_{1,1}), \quad \sigma = \arccos(\mathbf{D}_{1,1}/\mathbf{D}_{2,2}). \quad (1.6)$$

The main challenge of the shape-from-texture methods is, first of all, to estimate the frontal texel, and then to find transformation $T_{f \rightarrow i}$, $\forall i$. The methods of this category can be applied only to texture images.

The principle of the shape-from-contour approach is rather similar to the shape-from-texture. Observing an ellipse, one can make an assumption that in reality this is a circle projected to the image plane with certain slant and tilt angles, [7].

Another possible application for shape-from-contour principle is to search such slant and tilt angles which maximize the object contours symmetry, [8] and [9]. These techniques are applicable for objects contours which can be easily extracted from the input image, forms of this contours are usually particular and all constraints depend on prior knowledge.

Shape-from-(de)focus

This method is based on the limited depth-of-field of the camera. It is known that only points which are located at a certain distance (position of focus) will be captured to the image plane with a sharp projection, Fig. 1.6. Other object points will be blurred. When certain object regions are blurred it means that these regions are either closer or farther from the camera than the focus position. Using this information about defocused points from several images taken from the same fix position and varying focal lengths, shape-from-(de)focus methods are capable to reconstruct the object depth. We present the principle of focused and defocused images formation and basic shape-from-(de)focus technique as in [10].

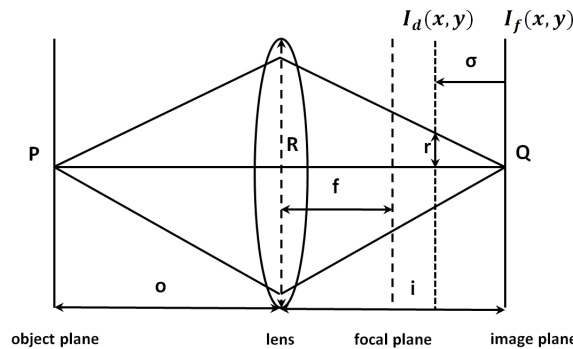


FIGURE 1.6: Demonstration of the Gaussian lens law.

Fig. 1.6 illustrates the Gaussian lens law which is the foundation of the shape-from-(de)focus. Let us have a convex lens of negligible thickness (thin lens) and of radius R and a focal length f . There is a point P on the object plane which is projected to the image plane into point Q . Let o denotes a distance between the lens and the projected object, and i is a distance from the lens to the image plane. As a consequence of triangles similarity we obtain the following equation:

$$\frac{1}{o} + \frac{1}{i} = \frac{1}{f}. \quad (1.7)$$

It can be seen that each point of the object plane is projected to a single point of the image plane. Therefore, the focused image $I_f(x, y)$ can be obtained only when the matrix of sensor is located at the same distance from the lens as the image plane. When it is not hold, we obtain blurred, or defocused image $I_d(x, y)$. When the matrix of sensor is placed on distance σ from the image plane, also using similar triangles, the degree of blur can be obtained by:

$$r = \frac{\sigma R}{i}. \quad (1.8)$$

Thus, the defocused image is a convolution of the focused image $I_f(x, y)$ and a circular symmetric function $p(x, y)$, and it can be presented as:

$$I_d(x, y) = p(x, y) * I_f(x, y), \quad \text{where } p(x, y) = \begin{cases} \frac{1}{\pi r^2}, & \text{if } x^2 + y^2 \leq r^2 \\ 0, & \text{otherwise.} \end{cases} \quad (1.9)$$

The principle proposed by authors in [10] is presented in Fig. 1.7. Let us have a surface S with a point s , which we want to reconstruct. This surface is situated on a movable platform, or translation stage. There is also a reference plane in the image scene. The camera is placed above the surface and the translation stage can be moved toward the camera. The camera lens focal distance is d_f , the depth of the point s is d_s , and the distance between the translation stage and the reference plane at each instant is d . Distances d_f and d are supposed to be known. Thus, when the platform is moved toward the focused plane, the degree of focus of image is increased. The best focusing is obtained for the maximum focus distance, and $d_s = d_f - d$. The most important question is how to evaluate the degree of focus from the image. It is done on the basis of focus measure operators, which are broadly analyzed and compared in [11].

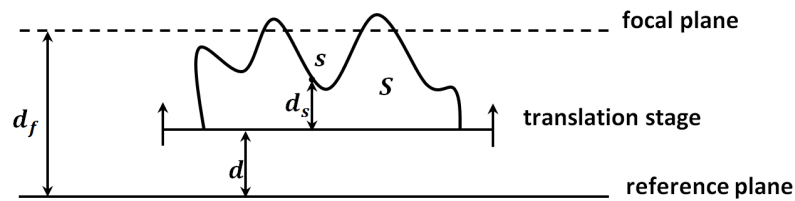


FIGURE 1.7: Shape-from-(de)focus principle illustration.

The difference between shape-from-focus and shape-from-defocus techniques is defined by the possibility to dynamically change parameters of the camera (e.g., focal distance). For the shape-from-focus we can change these parameters, at the same

time in shape-from-defocus methods we process already defocused images, using known parameters of the camera and modeling the defocusing.

By the means of the shape-from-(de)focus methods rather qualitative surfaces can be obtained. The main disadvantage of this technique is that the image acquisition process (parameters of the acquisition device and the camera) should be known in details and well controlled.

Shape-from-silhouettes

In order to capture input images for shape-from-silhouettes technique, one possibility is to use a turntable with an object and a camera which takes pictures of this object from different rotation angles. It is the same, if we have a static table with a set of cameras around the photographed object, Fig. 1.8. Each of camera captures the image from its own vantage point, and after that the object silhouettes are extracted for each obtained image. It is supposed that from each vantage point the silhouette forms a cone of projection rays. The results of intersection of all projections give us an approximate shape, which is also called a visual hull. This method can also be used for 3D video and motion capturing.

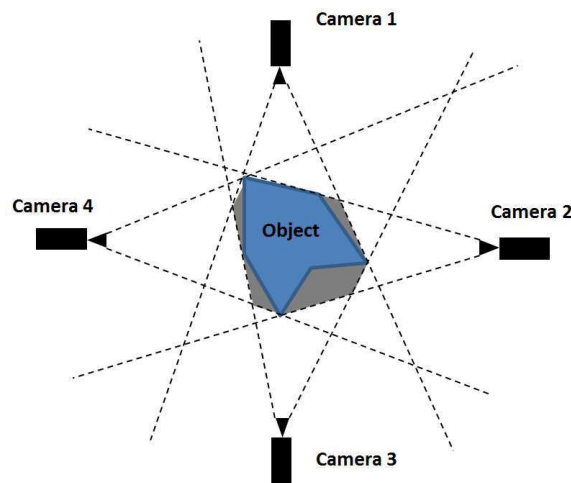


FIGURE 1.8: Illustration of the shape-from-silhouettes technique.

Resulting surface obtained using this method can be useful only for applications with no need in object surface details, e.g. [12]. Otherwise, enormous number of images should be taken from spatially close positions around the reconstructed object, or special refinement procedure should be applied to precise resulting surface details, [13].

In conclusion, passive reconstruction methods are powerful mathematical instruments for 3D surface evaluation. Nevertheless, each of them needs either a specific acquisition device and prior information on this device structure and/or characteristics (e.g., shape-from-stereo and shape-from-focus), or strong hypotheses on the reconstructed surface properties (e.g., shape-from-texture). Passive stereo algorithms can be rather efficient for colonies reconstruction. The main disadvantage is that in context of our problem we are limited to a single camera with fixed position, which makes impossible this technique application. Shape-from-texture methods utilization is restricted by presence of regular texture on the surface. Not all bacterial colonies have particular texture, moreover this texture can be non-uniform for different object parts, which can cause problems during reconstruction. The methods of the category 'shape-from-contours' are not applicable as well, because contours of microbial objects inside the Petri dish may be blurred and sometimes are very difficult to detect. From the existing device, camera parameters can change from one version of the device to another. So, we aim at being independent from these parameters and suppose them unknown. Shape-from-silhouettes algorithms do not give exact reconstruction, which can be crucial for further recognition process. The size of objects inside the dish is too small, colonies do not have significant silhouettes, and sometimes can be even almost flat.

In the context of our problem, we try to be as independent as possible from prior knowledge on acquisition device and reconstructed object. Thus, the utilization of passive methods in our case is difficult.

1.1.2 Active methods

In contrast to passive methods, active stereo approaches are based on images with one or several controlled illumination sources.

Active shape-from-stereo

In order to facilitate correspondence finding in passive shape-from-stereo tasks, the active stereo methods can be used. For these methods one of the cameras in a stereo setup is replaced by a projection device, Fig. 1.9. In other words, we obtain a combination of an illumination source and one camera. If the camera and the emitting device (e.g., laser) positions, orientations and parameters are known, the

position of the 3D object point can be found using exactly the same calculations as for the passive stereo, Fig. 1.4.

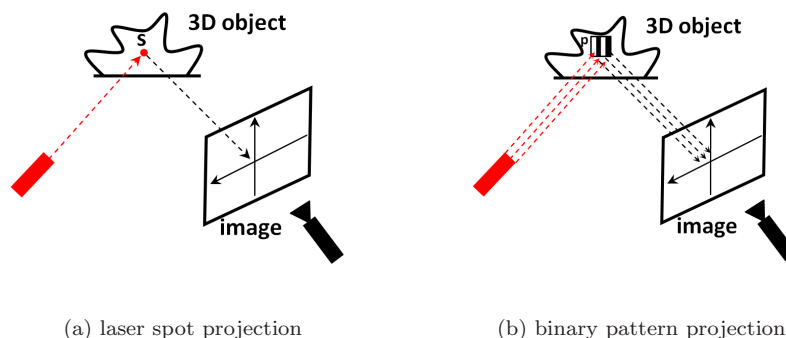


FIGURE 1.9: Active stereo reconstruction.

The main problem is that one point projection, Fig. 1.9(a), is not sufficient to obtain the whole surface reconstruction. Many points should be projected and captured. The acquisition device with 'scanning' laser needs complex mechanics and is time-consuming. One of the possible solutions is to replace the laser spot by some more complex pattern, for example a line of laser light with a cylindrical lens. But even 'scanning' with such a line remains long.

Another existing solution is to use some kind of binary or color codes into a two-dimensional projection patterns, in Fig. 1.9(b). For example, if the pattern is presented as parallel lines with alternating colors, the phase shift method can be efficiently used to obtain the resulting surface, [14].

All these techniques of patterns projections are often referred to as structured light techniques. They can be used either with visible light or light in infrared range as it is done in the famous Kinect device, [15], developed for human body 3D reconstruction and actions recognition. These techniques also turn to be efficient for metallic highly specular objects 3D scanning due to changes in metallic object temperature under action of laser (scanning-from-heating), [16], [17].

Shape-from-shading

When we look at an illuminated object, we are capable to extract information on surface orientation on the basis of shading variation. It becomes possible thanks to the physical principle that the shading of the object surface point depends on the incident light direction and this point orientation in 3D space (normal), Fig. 1.10(a).

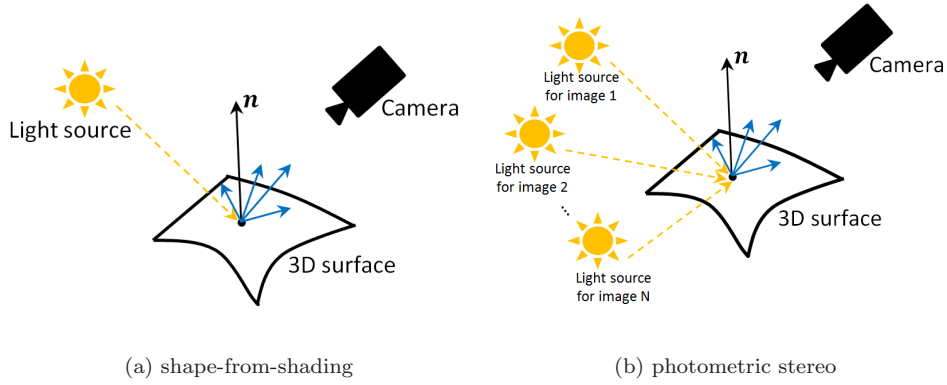


FIGURE 1.10: Surface reconstruction using shading variation.

Shape-from-shading consists in recovering the shape of the surface on the basis of intensities (or shading) variation, [18]. It is considered that all the reconstructed surface has a uniform reflectance factor (albedo), which rarely corresponds to reality. At the same time, the light source direction is either known a priori or can be calibrated using a reference object of known geometry. The light source is supposed to be a distant (point). According to all these assumptions, irradiance \mathbf{I} equation due to the shape-from-shading has the following formulation:

$$\mathbf{I}(x, y) = R(p(x, y), q(x, y)), \quad (1.10)$$

where $(p(x, y), q(x, y)) = (\mathbf{z}_x(x, y), \mathbf{z}_y(x, y))$ are the depth map derivatives, $R(p(x, y), q(x, y))$ is called the reflectance map. Reflectance map equation depends on the surface material properties and the light source direction.

Illustration of reflectance maps in surface gradient space was proposed by Horn in [18]. Let us have a surface point M , Fig. 1.11(a). If we consider that 3D surface around M is locally linear, then all neighbor surface points (x, y, z) can be projected to the plane with the coefficients A , B , C and D in the form of equation:

$$\begin{aligned} Ax + By + Cz + D &= 0, \\ z &= -\frac{A}{C}x - \frac{B}{C}y - \frac{D}{C}. \end{aligned} \quad (1.11)$$

Partial derivatives of the surface depth z can be written as:

$$\begin{cases} \frac{\partial z}{\partial x} = -\frac{A}{C} = p, \\ \frac{\partial z}{\partial y} = -\frac{B}{C} = q. \end{cases} \quad (1.12)$$

Using tangent to the surface vectors $(1, 0, p)^T$ and $(0, 1, q)^T$, value of the normal \mathbf{n} in the point M can be written as:

$$\hat{\mathbf{n}} = \frac{\mathbf{n}}{|\mathbf{n}|} = \frac{(1, 0, p)^T \times (0, 1, q)^T}{\sqrt{1 + p^2 + q^2}} = \frac{(-p, -q, 1)^T}{\sqrt{1 + p^2 + q^2}}. \quad (1.13)$$

In order to present a reflectance map, cosine of the incident light direction θ_i should be defined. In terms of the surface gradient, this cosine is presented as:

$$\cos(\theta_i) = \frac{pp_s + qq_s + 1}{\sqrt{1 + p^2 + q^2} \sqrt{1 + p_s^2 + q_s^2}}, \quad (1.14)$$

where $(-p, -q, 1)^T$ and $(-p_s, -q_s, 1)^T$ define the surface point orientation and the lighting direction in gradient space, respectively. The reflectance map then can be written as:

$$R(p(x, y), q(x, y)) = \rho \cos(\theta_i) = \frac{\rho (pp_s + qq_s + 1)}{\sqrt{1 + p^2 + q^2} \sqrt{1 + p_s^2 + q_s^2}}, \quad (1.15)$$

where ρ is the material reflective power (albedo).

The reflectance map theory proposed in [19] states that the reflectance $R(p, q)$ which defines the surface-intensity relation for the image can be seen graphically as a map of iso-contours, Fig. 1.11(b). All iso-contours are centered around the direction of lighting (p_s, q_s) which corresponds to the most elevated intensity value of the object image. Each iso-contour $R(p, q)$ corresponds to all possible combinations of p and q values which define a certain intensity value of the image pixel. Thus, the reflectance map in gradient space does not give a unique correspondence between the image intensity, Eq. 1.10, and the surface gradient (p, q) . In shape-from-shading methods these ambiguity is resolved by the means of additional constraints on surface, camera and light source properties.

A broad survey on the shape-from-shading methods is presented in [20]. Conclusions made in this survey are rather disappointing. None of the discussed shape-from-shading algorithms gives satisfactory results. In most of cases, the found solution corresponds to local minima in the search space, and the reconstructed surface is of bad quality. A certain revolution in the domain of shape-from-shading problems was made with works of Barron and Malik, [21], [22] and [23]. The authors managed to obtain qualitative surfaces from a single grayscale or color image with artificial or natural illumination using a lot of priors on reflectance, shape and illumination and exploiting a complex multi-scale optimization technique.

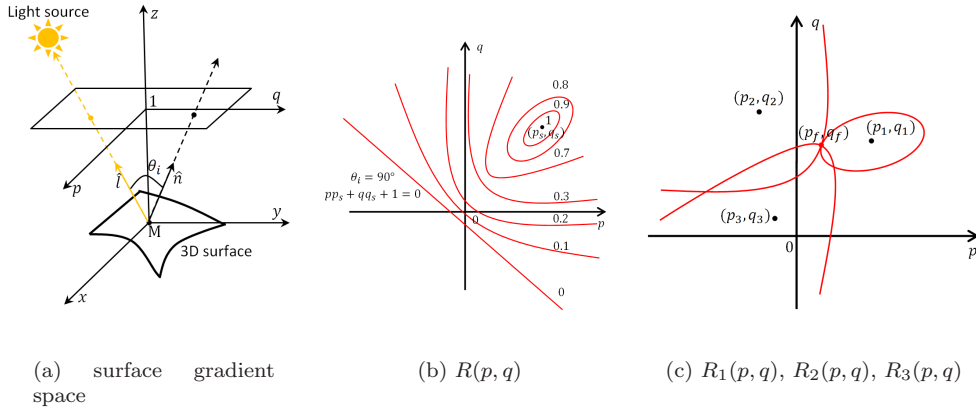


FIGURE 1.11: Surface gradient space and reflectance map representation.

Photometric stereo

Another possibility to use the information on shading variation for surface reconstruction is to exploit several images made from a fixed camera position and varying illumination directions for each image acquisition, Fig. 1.10(b). This technique is called photometric stereo (PS), [24], because many positions of light sources play the same role as several positions of cameras in passive stereo methods. According to our consideration, shape-from-shading and photometric stereo have common physical foundations. Shape-from-shading can be presented as a special case of photometric stereo when it is used only one image for the 3D reconstruction. Similarly to the shape-from-shading, we can present a reflectance map for each of N initial images: $R_1(p(x, y), q(x, y))$, $R_2(p(x, y), q(x, y))$, ..., $R_N(p(x, y), q(x, y))$. Surface $(p, q, 1)$ and albedo ρ can now be found using at least 3 input images if all other parameters of the reflectance maps are known. For the easiest case of Lambertian surface, for the image pixel (x, y) using Eq. 1.10 we obtain the following system of equations:

$$\begin{cases} \mathbf{I}^1(x, y) = \rho \mathbf{n}(x, y) \mathbf{v}_1, \\ \mathbf{I}^2(x, y) = \rho \mathbf{n}(x, y) \mathbf{v}_2, \\ \dots \\ \mathbf{I}^N(x, y) = \rho \mathbf{n}(x, y) \mathbf{v}_N, \end{cases} \quad (1.16)$$

where ρ denotes the albedo value, N is the number of initial images, $\mathbf{I}^i(x, y)$ is the intensity value of the pixel (x, y) for image i , $\mathbf{n}(x, y)$ is a 3D normal to the surface point which corresponds to the pixel (x, y) and \mathbf{v}_i denotes the 3D direction of the light source for the i^{th} image. If the information on light sources positions is

available, this system of equations can be solved using linear least squares method, [25].

In terms of reflectance maps for Lambertian surface, Fig. 1.11(c), each input image of the photometric stereo technique has its own set of iso-contours centered in reflectance map point corresponding to light source intensity. The intersection of iso-contours obtained for different lighting directions, for example, of three initial images, uniquely determines the orientation of the surface points. Such an intersection is illustrated in Fig. 1.11(c) by the point (p_f, q_f) . This property of the reflectance maps formed the basis of the PS theory, firstly introduced in [24].

When the reconstructed surface has complex reflective properties, the reconstruction process becomes more difficult. There is a need in supplementary knowledge on camera position and more sophisticated reflectance equations and reconstruction algorithm usage. The main deficiency of the photometric stereo is a presence of a bas-relief ambiguity, [26], which can be partially solved using shape-from-interreflection algorithms, [27], or properties of perspective projection, [28]. And even the most complex reflectance models do not cope with inner scattering for transparent surfaces. The result of the photometric stereo is not an absolute depth but a normal map with surface orientation for each image pixel. This normal map should be supplementary integrated to obtain the depth reconstruction.

As for the case of passive methods, active methods also require in most cases a specific device and prior knowledge on this device configuration. To solve our task, we cannot utilize active stereo, because in the capturing device of bioMérieux company, which is used in context of our problem, there is no projection system. But even if we imagine the presence of a necessary device, bacterial colonies can be of a very small size, so structured light algorithms may be not sufficient for their detailed surfaces reconstruction. The algorithms of shape-from-shading can also potentially be used for the colonies surface reconstruction, but if we consider having no prior on acquisition process and object to reconstruct, it becomes difficult to resolve a shape-from-shading problem. Realized this overview and analysis, we can conclude that only utilization of active reconstruction methods, such as photometric stereo algorithms, corresponds to the work needs and is fully compatible with the existing capturing device.

1.2 Problem Statement

Returning to the motivations of the work, we aim at microbial colonies surface reconstruction from Petri dish images. Microbial colonies have rather specific surface properties, which potentially may influence reconstruction and become a real challenge for this process. First of all, colonies are very diverse. They can be of a small size, sometimes less than 1 mm in diameter, and have complex reflecting characteristics. In addition, it is not known a priori which colonies will be obtained, so there is no prior information on surface properties for the reconstruction algorithm. Moreover, colonies are living organisms, thus, the expected analysis should be non-destructive. This gives an additional possibility for the further colonies growing and study. There exist many methods for microbial colonies analysis. By bioMérieux company it was chosen to use a non-destructive technology based on images of Petri dishes. These images should be used as input data for the surface reconstruction algorithm.

The choice of the acquisition device is also dictated by the idea of the further industrialization process. Thanks to modern cameras, it is possible to make fast acquisitions which increases the capacity of the device and gives a possibility to process a large number of Petri dishes with a reasonable time delay. Also this choice is associated with relatively small prices of the device components, there is no need in usage of expensive non-conventional optical systems. Moreover, the domain of image processing is vast and well developed with many automated algorithms which can be adapted for different microbiological needs. Thus, it has been finally decided to exploit the imaging device with a single fixed camera and illumination of visible range of wavelengths. The illumination direction varies from one acquisition of the Petri dish to another, as a result, for each dish we obtain a set of images with different lightings. Different light directions exploited for acquisition allow obtaining diverse information on microbial colonies color, size, shape, contours, elevation, etc. This set of images can be also further exploited for the photometric stereo reconstruction.

The main limitation of the existing photometric stereo algorithms (see Section 1.1) is in requirements of additional prior information either on lighting conditions or on surface orientation and properties. This information is not stable or even not available in most of the real applications, which is the case for our problem. The acquisition device creates illumination conditions difficult to model, the light sources may evolve in time or be not calibrated. Thereby, in our problem we suppose them to be

unknown. Information on the reconstructed surface properties (microbial colonies) is not available a priori as well.

The purpose of the thesis work is to develop photometric stereo approach which could cope with this lack of knowledge. On the one hand, the proposed method should be applicable to any sequence of images with a fix camera and varying unknown directions of light sources. On the other hand, this method should be able to provide a qualitative reconstruction for non-Lambertian objects, which include microbial colonies, without prior information on the surface material.

1.3 Thesis Contributions

In this thesis work we propose an uncalibrated non-Lambertian photometric stereo approach. This original approach is realized in two stages: specularities correction in order to pass from non-Lambertian to 'pseudo-Lambertian' surface, which is further used for the second stage of uncalibrated photometric stereo reconstruction. The main advantage of my approach consists in its independence from prior knowledge on the light sources directions and intensities used during the input images acquisition as well as on the properties of the surface to be reconstructed.

The first stage is dedicated to specularities correction in order to obtain images of 'pseudo-Lambertian' surface and to be able to apply Lambertian reflectance model. This correction is made using an original algorithm, which is able to detect specularities as abnormally elevated pixel intensity values in an image of the input sequence, and to correct the found zones using information from all other images of the sequence and a specific continuous function. It copes even with large specular zones. This method allows removing specularities while still preserving all other particularities of shading important for the further surface reconstruction.

After the described pre-processing, we propose an original algorithm of uncalibrated Lambertian photometric stereo reconstruction. Uncalibrated method supposes that there is no prior knowledge on light sources directions. Information on the light sources can be changed from one acquisition device to another. Thus, with the hypothesis of unknown illuminations we can apply the proposed algorithm to any existing photometric stereo databases, or any sequences of images made with a fixed camera position and varying illuminations. The implemented photometric stereo model has four components: two of these components (albedo and normals) describe

surface properties and the other two (light sources intensities and directions) describe illumination. The proposed algorithm of the photometric stereo reconstruction uses the alternating optimization principle with value and quality constraints. The original resolution scheme allows separating the different information types included in the input images. Thanks to such matrix factorization the surface reconstruction is realized with no prior information neither on lighting directions nor on the properties of the object to be reconstructed. The algorithm is evaluated using artificial and open image datasets for which the ground truth information is available.

The last contribution of this thesis is the application of the proposed reconstruction approach to Petri dish images taken with the specific device developed by bioMérieux company. This thesis part adds value to the proposed algorithms usage for surface reconstruction of non-Lambertian colonies. The reconstruction results can be further used for the colonies visualization, and the obtained 3D surface completes the information provided by the images.

Different parts of this work are published in two conference papers and two patents:

1. K.Kyrgyzova, L.Allano, M.Aupetit. Highlights and shadows multi-image correction while remaining objects shading. Le 24me Colloque de GRETSI, Brest, France, 2013.
2. K.Kyrgyzova, L.Allano, M.Aupetit. Alternating optimization for Lambertian photometric stereo model with unknown lighting directions. The 15th International Conference on Computer Analysis of Images and Patterns (CAIP), York, Royaume-Uni, 2013.
3. L.Allano, K.Kyrgyzova. Method, system and computer program product for improving the quality of an image, EP13305673.9, 2013.
4. L.Allano, K.Kyrgyzova. Method, system and computer program product for producing a raised relief map from images of an object, EP13305672.1, 2013.

1.4 Thesis Outline

The thesis manuscript consists of six chapters. Chapter 1 is introductory, it includes a review of existing 3D surface reconstruction methods, discusses their limits, and argues for the choice of the photometric stereo technique to our problem of the microbial colonies surface reconstruction. Chapter 2 is dedicated to the foundations of the image formation process and the relation of this process to existing photometric

stereo algorithms for specular objects. We present the hypotheses on the properties of the acquisition system and surface which serve as a basis for the proposed approach. In Chapter 3 we disclose details of the proposed method for specularities correction. This correction allows obtaining images of 'pseudo-Lambertian' surfaces which are suitable for the further photometric stereo reconstruction. The proposed algorithm of uncalibrated photometric stereo reconstruction is presented in Chapter 4. Thanks to this algorithm, it is possible to obtain the depth map without prior knowledge on light sources and surface properties. And finally Chapter 5 is dedicated to the application of the proposed uncalibrated non-Lambertian photometric stereo reconstruction to microbiological data. we conclude and discuss the possible perspectives of this work in Chapter 6.

Chapter 2

Non-Lambertian surface reconstruction problem

Photometric stereo (PS) is an existing image processing technique which is widely known and used since the 1980s. This approach was invented by Woodham in [24] on the basis of works by Horn on the reflectance map theory, [19], and the shape-from-shading reconstruction, [18]. Compared with other existing 'shape-from-X' approaches, PS reconstruction is based on an interaction between light source(s), camera and surface of a reconstructed object. This interaction is a complex physical process which can be described by the means of reflectance distribution functions. Modeling of these functions requires knowledge on the acquisition device (camera position and lighting directions) and on the reconstructed surface material properties. In this chapter we make a review of state-of-the-art reflectance distribution models and existing algorithms of surface reconstruction which use these models. We also present hypotheses related to the stated problem and limited by the structure of the Petri dishes imaging device, as well as our own approach of the processing which can be used for the non-Lambertian PS.

2.1 Light-Surface-Camera interaction

We are able to see objects (as a camera is able to capture their colors intensities), which do not emit the light themselves, thanks to the incident illumination arriving from other emitting objects (light sources). Let us discuss in details each actor of

this interaction and list hypotheses related to the acquisition system for Petri dish images and particularities of the reconstructed surface.

2.1.1 Light source

The first essential element of the image formation process is the light source. The photometric stereo is based on images acquired with different illumination directions. Illumination used during each image acquisition can be complex and may consist of several light sources or light sources surrounded by a diffuser. Obviously, the complexity of a light source structure increases the complexity of an image formation process modeling. In our approach, the light source used for each image acquisition is supposed to be of a directional nature. Such a light source is of a negligible size and can be characterized by a single direction, intensity and color. Intensity of the directed light source is related to the quantity of photons emitted by the light source, and color of the source depends on the wavelength of the emitted light. Emitted light can have a visible range of wavelengths (e.g., white light), as well as it can have the wavelengths of ultra-violet (UV) or infra-red (IR) ranges.

In our approach we make an assumption of a single light source, which may vary in direction and intensity from one acquisition to another. This light source is of a visible wavelength range.

2.1.2 Camera

The second element of the image formation process is the camera. The camera is a device designed to capture the reflected or emitted light. It is characterized by the optical lens parameters, its spatial position and the camera response function.

In our approach, we suppose that the photographed object is far from the camera and close to the camera optical axis. This hypothesis permits to use an orthographic projection as a linear approximation of the perspective projection of an object to the image plane. The camera is also supposed to be fixed from one image acquisition to another, which provides a total correspondence between pixels of different images of the initial sequence and there is no need to solve the additional problem of point correspondence.

The camera response function refers to how the radiance (quantity of photons) which arrives to image is transformed to the pixel intensity value. In the context of the industrial project of Petri dish images processing, we work with an imaging device prototype. This prototype contains a CCD camera with a linear response function. Thereby, in our photometric stereo modeling we do not consider a specific function for the camera. However, with the evolution of the prototype, the evolution of this parameter is also possible. The camera response function can strongly impact the reconstruction results. For such a case, an appropriate transformation should be applied to the image intensity values in order to avoid artifacts related to the camera response function.

2.1.3 Surface

The third and the most complex actor of the image formation process is the surface. Once the light arrives at the surface point, there are three possible types of interaction between the incident light and the object surface: the light can be reflected, refracted or scattered. Reflection can be seen as an external surface event, refraction (or transmission) is purely an internal surface event, and scattering is a combination of external and internal events. All these phenomena are schematically presented in Fig. 2.1. Let us discuss more in details each of these light events and existing distribution functions for their modeling.

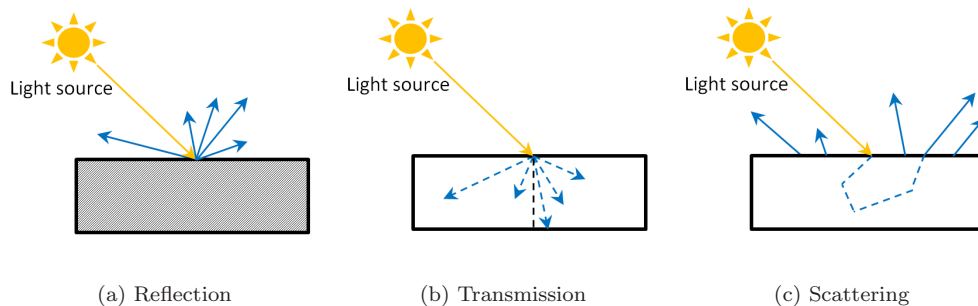


FIGURE 2.1: Interaction of the incident light with the object surface.

Reflectance distribution functions

In Fig. 2.1(a) the light which arrives to the surface point is immediately reflected in all possible spatial directions from the same point to which it arrived. The quantity of

light reflected in each direction is usually different, it strongly depends on the object material properties and on the direction of illumination. Reflection can be explicitly modeled by Bidirectional Reflectance Distribution Functions (BRDFs) which take into account all affecting properties of the light source(s), the object surface and the viewer (or camera). In other words, a BRDF describes how much light is reflected in the camera direction when it interacts with a surface made of a certain material depending on source direction. This function is widely presented in Section 2.2.2. It can be applied to model the image formation process for opaque objects or objects which can be roughly considered as opaque. For the opaque objects all the light is either absorbed by the material or reflected. The reflected light can be seen by the viewer (or captured by the camera), while the light absorbed by the material is neither reflected nor transmitted. The absorption is measured by the means of the material reflective power (albedo).

The second type of light event, transmission, Fig. 2.1(b), is dedicated to transparent or translucent objects. Transmission event can be modeled by the means of the Bidirectional Transmittance Distribution Functions (BTDFs). The BTDF describes how much light is transmitted when it interacts with a material. The BTDFs are not used alone for the image formation process modeling, they are usually used in combination with BRDFs to describe the third light interaction with the surface, scattering. Indeed, to make an object visible it is mandatory to have a certain amount of light reflected directly from the surface or emitted from the surface after going through an internal path.

Scattering, Fig. 2.1(c), is the most complex incident light interaction with the surface. In this process it is supposed that the incident light is transmitted inside the surface material, interacts with the material and gets out from the material but not necessarily at the point of its arrival. The scattering thus consists in taking into account simultaneously transmissions and reflections. The functions for its modeling, Bidirectional Surface Scattering Reflection Distribution Functions (BSSDRFs) are combinations of the BRDFs and the BTDFs. The BRDF can also be seen as a simplified version of the BSSRDF for opaque materials with reflections only in the incident light arrival point.

Different colors of the light spectrum are reflected, transmitted and absorbed differently. It depends on the physical properties of the surface material. Thus, BRDF, BTDF and BSSRDF are also functions of wavelengths and ideally should be presented separately for each image color channel. The surface of the object can be

also composed of several materials or the light distribution function can be inhomogeneous for different object parts.

It should be mentioned that all the discussed functions (BRDF, BTDF and BSSRDF) are used not only for the photographed object surface reconstruction, but mostly for 3D objects rendering for realistic objects visualization (e.g., [29], [30], [31]). The direct task of rendering supposes using pre-defined model and parameters. Thus for this problem, one can apply more sophisticated models than for photometric stereo.

Transparent surface reconstruction

Task of transparent (and translucent) surfaces reconstruction with transmission and scattering effects is a complex inverse problem. There are too many unknown factors which influence the incident light inner scattering and can be unpredictable or even immeasurable for some cases. We can find certain bibliographic sources on this subject, however the existing algorithms for the most part use non-conventional data acquisition methods and then only solve the stated problem mathematically or algorithmically. In this brief review we do not take into consideration semi-transparent dynamic objects reconstruction (such as fire or smoke), they have very particular physical nature which is taken into account during the reconstruction process and is not present in our surface obtaining problem.

Work on water surface reconstruction presented in [32] can be considered as one of the first approaches on the refractive surface reconstruction by the means of the model:

$$r\mathbf{v}_w - \mathbf{v}_a = \mathbf{n}(r \cos(\beta) - \cos(\alpha)), \quad (2.1)$$

where r is a refraction index, \mathbf{v}_w denotes a unit vector of the incident ray in water, \mathbf{v}_a is a unit vector of the incident ray in air, \mathbf{n} corresponds to the surface normal, α and β angles as displayed in Fig. 2.2. The algorithm proposed by the authors is of the shape-from-distortion category. The camera is placed perpendicular to the water surface with a 2D pattern on the bottom. Then a sequence of images distorted by water movement is taken. After that, on the basis of the optical flow analysis and the pixels average trajectory, it is suggested to reconstruct the gradient map and integrate it to the final surface.

In [33], for transparent surfaces reconstruction and rendering, the authors used a special acquisition setup consisting of a rotating platform with a reconstructed object,

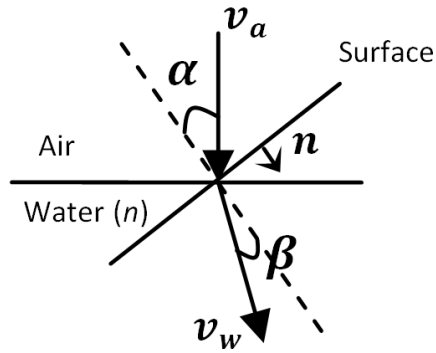


FIGURE 2.2: Shape-from-distortion for water surface from [32].

4 light sources connected to the platform, 6 cameras and 2 monitors for creating color patterned backdrops from the side and from below, Fig. 2.3. Then 4×11 images are taken from 6 fixed cameras and the reconstruction of the surface is made on the basis of the opacity hull construction as it is done in shape-from-silhouettes technique, slightly presented in Section 1.1.1. This algorithm can be considered as a combination of the structured light and the shape-from-silhouettes techniques.

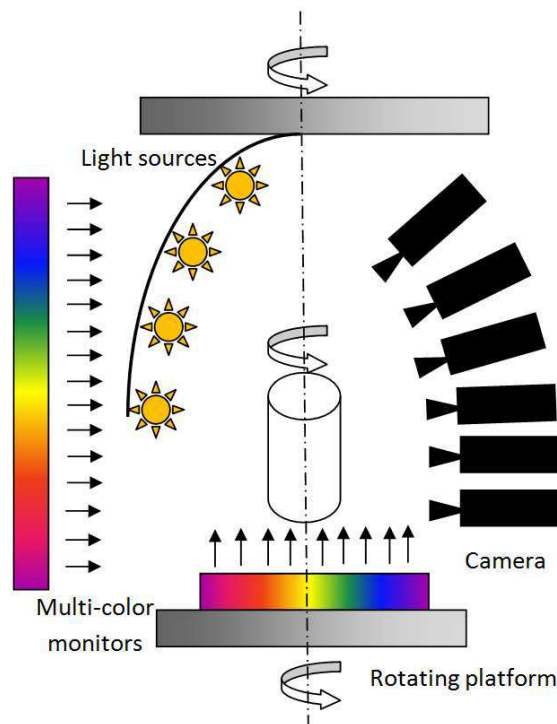


FIGURE 2.3: Transparent surface reconstruction by the structured light and shape-from-silhouettes from [33].

Polarization ray-tracing approach for transparent objects reconstruction is presented by the authors in [34]. The so-called acquisition system 'Cocoon' consists of a

camera, a light polarizer, a geodesic dome with light sources which surrounds a spherical plastic diffuser with a transparent object inside. Four images with 0° , 45° , 90° and 135° rotations of polarizer are taken. After that, the proposed iterative algorithm estimates the surface orientation minimizing the difference between the data rendered by the polarization ray-tracing and the polarization data obtained during the acquisition. The measured degree of polarization for a pixel (x, y) is defined as:

$$\eta(x, y) = \frac{\mathbf{I}_{max}(x, y) - \mathbf{I}_{min}(x, y)}{\mathbf{I}_{max}(x, y) + \mathbf{I}_{min}(x, y)}, \quad (2.2)$$

where $\mathbf{I}_{max}(x, y)$ and $\mathbf{I}_{min}(x, y)$ correspond to the maximal and minimal intensity values obtained for different rotations of the polarizer. At the same time, the theoretical value of the degree of polarization for the given refraction index r and is defined from the Snell's law:

$$\eta(x, y) = \frac{2\sin(\phi)\tan(\phi)(r^2 - \sin^2(\phi))^{1/2}}{r^2 - \sin^2(\phi) + (\sin)^2(\phi)\tan^2(\phi)}, \quad (2.3)$$

where ϕ corresponds to the orientation of the plane of incidence. Substituting the measured η from Eq. 2.2, the angle ϕ which describes the surface can be found. The weakness of this algorithm consists in prior knowledge of the refractive index which is untruthfully for many applicative tasks. Light polarization principle is also efficiently used in works [35], [36] and [37] with various evaluation algorithm but using the same physical principle.

The authors of [38] proposed to use the tomographic reconstruction for transparent object surfaces. They also invented a new setup for images acquisition. It consists of a turntable, a glass cylinder with a fluid, a uniformly illuminated backdrop and a camera. First of all, a geometrical calibration is made in order to obtain the light paths corresponding to each image pixel. After that, a semi-transparent or a totally transparent object (the authors discuss both cases) is placed inside the cylinder filled with the fluid with a refractive index close to the index of the reconstructed transparent object. The mounted camera takes a significant number of images (from several dozen to a few hundred depending on the angle step of the turntable). As in the case of x-ray tomography, the pixel values of the obtained images are line integrals of the visible light absorption along the corresponding calibrated ray. To obtain the final object 3D representation, a standard tomographic problem is formulated and solved. First, the absorption of light along a ray in the logarithmic space is described

as:

$$\log A = - \int \alpha(t) dt \approx - \sum_i \alpha_i, \quad (2.4)$$

where α_i corresponds to the density of the i^{th} voxel (regular grid 3D point). For each 3D slice and each incident ray, the distance of the ray to all voxels can be computed. This distance is used to define a weight w_j for each voxel j in the proximity of the ray. Thus, the new k estimates of the absorption and difference with total absorption along the ray are:

$$\log A^{(k)} = \frac{\sum_j w_j \alpha_j^{(k)}}{\sum_j w_j}, \quad (2.5)$$

$$\Delta A = \log A - \log A^{(k)}. \quad (2.6)$$

The updated density of the voxel is then calculated as:

$$\alpha_j^{(k+1)} = \alpha_j^{(kj)} + \lambda \frac{\sum_j w_j \Delta A}{\sum_j w_j}, \quad (2.7)$$

where λ is a relaxation factor for the optimization procedure. The iterations of the algorithm continue till a certain convergence criterion is valid. The algorithm from [38] requires a large number of initial images for precise results. The computation time strongly depends on the size of initial data, but for not time-critical tasks this algorithm turns out to be rather efficient.

The authors of [39] propose an algorithm which is capable to reconstruct exterior surface of a transparent object with inhomogeneous inner content (e.g., a plastic bottle with apple juice). In order to take the algorithm input images it is proposed to use a one-view setup with dense positions of the light sources on a regular grid. Thus the reflected light captured by the camera consists of the direct surface reflection and some additional light reflection related to the light inner scattering (so-called light transport). In the proposed algorithm the direct reflection component and the indirect lighting effects are separated thanks to the physical properties of the light transport. There are two main constraints used for surface reconstruction: the incident light rays employed for each image acquisition converge towards the surface point and the reflection point should be located on the (known) viewing direction. Under these strict constraints the authors provide very good results of transparent surfaces reconstruction.

One of the possible way for transparent surfaces acquisition is to use an ultra-violet laser source, [40], which during the acquisition creates white fluorescence points not

affected by the transmission. This method was developed concretely for classical glasses which have a strong absorption band and appear opaque within the ultraviolet range.

In our work we do not consider transmittance and scattering for the reconstructed objects. These events certainly occur for microbiological objects inside a Petri dish, but a simplification to the opaque materials and a modeling with the BRDF function is already not an easy task for solving in the conditions of unknown surface and lighting properties. Such a simplification obviously leads to distortions for the reconstructed object and an imprecision of the resulting surface. This imprecision can be neglected if we use only general surface curvature for the analysis. However, such a distortion related to the model simplification can have negative consequences for tasks related to accurate surface details analysis (e.g., roughness extraction).

Despite all the described techniques, it can be concluded that the task of transparent surfaces reconstruction has not been wholly solved yet in terms of the PS approach with a simple fixed camera position and varying illuminations. Hypothesis of unknown lighting directions and even basic surface properties, such as a refraction index, makes impossible to use the methods from the literature. Non-conventional illumination (e.g., polarization, UV, etc.) are not included to our acquisition device. That is why, a transparent nature of reconstructed object is not taken into account in our proposed approach; and all surfaces are considered as opaque. For the opaque surface, it is supposed that there is only a reflection light event which can be described by the BRDF.

We can conclude this section, summarizing all the hypotheses for the stated problem:

- Light source**
- the light sources used for each image are supposed to be of a directional nature;
 - the illumination direction and intensity vary from one image to another;
 - there is no prior knowledge on light sources directions and intensities.
- Camera**
- the camera is fixed;
 - orthographic projection of the object surface to the image plane is supposed;
 - linear camera response function.
- Surface**
- the reconstructed object is supposed to be opaque and only light reflection is considered for the image formation process;
 - there is no prior knowledge on material properties.

The interaction between an opaque surface and lighting can be described by a bidirectional reflectance distribution function (BRDF). This function depends on the direction of illumination, the camera position and the reconstructed surface reflecting properties.

2.2 Reflectance of opaque surfaces

A surface is considered as opaque if it does not let an incident light pass through, the light can be absorbed by the material or reflected. Thanks to the reflection of light, opaque objects are visible. Opaque surfaces can be described by the BRDF which has different properties depending on the surface material.

2.2.1 Dichromatic reflectance model

All the existing BRDFs can be generalized by the means of the dichromatic reflectance model, first introduced in [41]. The model is called dichromatic because it presents reflected light consisting of two additive components: interface (specular) and body (diffuse) reflections. A third component, which presents the ambient light, may also be included to the reflectance model. In this work, we suppose to have no ambient light, due to the structure of the imaging device as a closed cupboard.

Reflection of light from other objects inside the dish is supposed to be negligible and is not taken into account as well. In further descriptions we will use interface and specular reflections, as well as body and diffuse reflections as synonymous terms.

The interaction between the incident light and the object surface, as shown in Fig. 2.4, depends on three vectors: the incident light direction, the surface normal and the emergent direction. The incident light is presented by the vector \mathbf{v} and the angle between the incident light direction and the surface normal, or the angle of incidence, (θ_i, ϕ_i) . This angle, as well as all others, is presented in spherical coordinates for a unit vector. The surface normal is represented by the vector \mathbf{n} . The emergent light is shown by the vector \mathbf{s} . There are two angles which characterize the emergent light: the viewing direction, or the angle of reflectance between vectors \mathbf{s} and \mathbf{n} , (θ_o, ϕ_o) , and the angle between the incident and the emergent light directions, or the phase angle, (θ_s, ϕ_s) . In Fig. 2.4 we present also the vector of the perfect specular reflection \mathbf{r} and corresponding angle between this vector and the surface point normal \mathbf{n} , (θ_r, ϕ_r) .

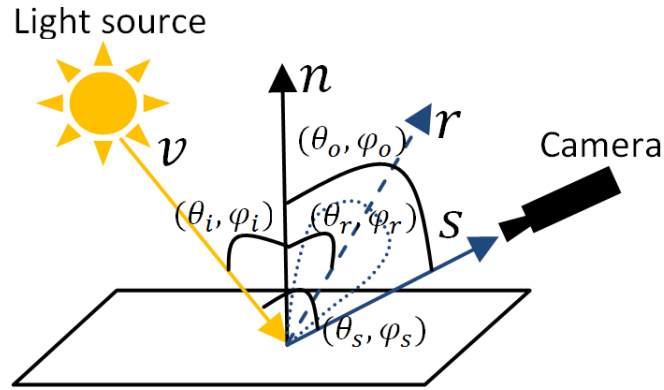
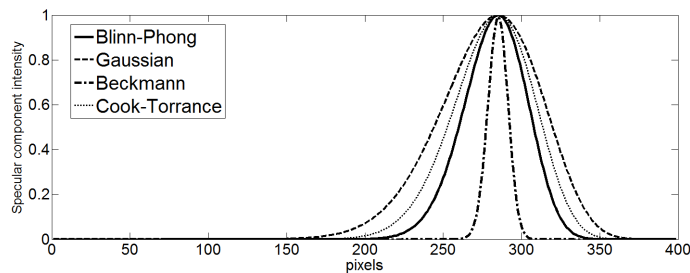
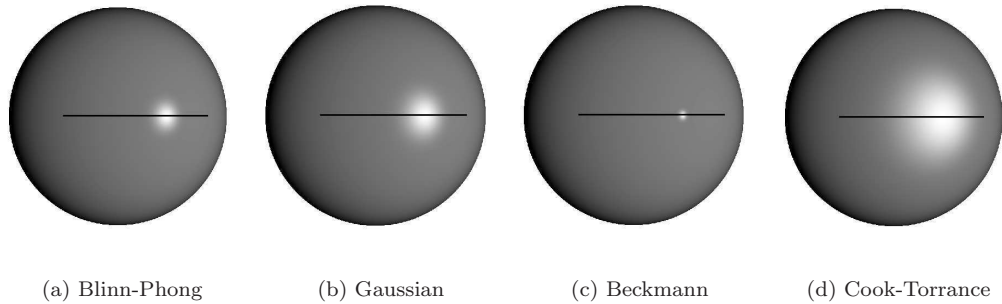


FIGURE 2.4: Specular effect formation.

According to the nature of the specular effect formation, the most intense specularity is obtained when the surface normal bisects the phase angle (θ_s, ϕ_s) . The specularity is formed not only for one surface orientation, but for a set of reflectance angles. This is called specular lobe, which is usually of continuous nature. The modeling of specularity in images corresponds to the specular lobe representation by different types of intensity values distributions depending on the relative position of a camera and the perfect specular direction.

There exist many conceptions for the specular lobe modeling. Several of them, the most widespread in Computer Graphics (such as the Blinn-Phong model, [42], the

Gaussian model, [43], the Beckmann model, [44], and the Cook-Torrance model, [45]) are presented in Fig. 2.5. In the given examples there are four different specular distributions on the surface of a sphere with the same viewing direction and direction of illumination. It can be seen that the form of the specular lobes of different distributions are similar, Fig. 2.5(e). Yet it should be mentioned that they can depend on the same (e.g., illumination and viewing directions) or on the different parameters (e.g., Beckmann model includes the surface material roughness parameter).



(e) distribution of specularity for the chosen image line

FIGURE 2.5: Specular component distribution models.

Not all the incident light is reflected to the perfect specular direction because of irregular microfacets composing the surface even for surfaces that appear quite smooth. Each microfacet, due to their orientations, reflects light in its own perfect specular direction. Many microfacets on a surface contribute to a single pixel of the final image. Such a reflection can be seen as an aggregate of the microfacets reflections, and is called light diffusion (or body radiance).

The dichromatic reflection model presents the reflected light and the light captured by a camera as a combination of two independent components, interface and body radiances. This type of light is not characterized by a direction, but only by an intensity, and it arrives to the object simultaneously from different sides. The dichromatic

reflectance model for angles (θ_i, ϕ_i) , (θ_o, ϕ_o) and (θ_s, ϕ_s) can be written as, [41]:

$$\mathbf{L}(\lambda) = \mathbf{L}_i(\lambda) + \mathbf{L}_b(\lambda) = m_i \mathbf{C}_i(\lambda) + m_b \mathbf{C}_b(\lambda), \quad (2.8)$$

where \mathbf{L} corresponds to the reflected light, \mathbf{L}_i and \mathbf{L}_b are radiances related to interface and body reflections respectively, \mathbf{C}_i and \mathbf{C}_b are spectral power distributions which depend only on wavelength of the light, and finally, m_i and m_b are geometrical scale factors which depend only on surface geometry.

Lambertian surfaces can be defined as surfaces which do not form interface reflection. Under this assumption, the dichromatic reflection model, Eq. 2.8, can be simplified by eliminating the specular \mathbf{L}_i component, which leads to consider only the diffuse component.

Another interesting property of the dichromatic reflection model is the ability to interpret a pixel color formation in terms of the spectral distribution power and the geometrical scale factor, [41]. Let us have the measured colors of a pixel $\mathbf{C}_L(x, y)$, which can be defined as:

$$\mathbf{C}_L(x, y) = m_i(x, y) \mathbf{C}_i + m_b(x, y) \mathbf{C}_b, \quad (2.9)$$

where $m_i(x, y)$ and $m_b(x, y)$ are geometrical scale factor of the pixel for the interface and the body reflection respectively, \mathbf{C}_i and \mathbf{C}_b are the colors of the interface and the body reflection of the material. The geometrical representation of this equation is given in Fig. 2.6. Having no ambient light, if we consider that $\mathbf{C}_L(x, y)$ contains a set of all possible colors of the surface, then values of $m_i(x, y)$ and $m_b(x, y)$ are different for surface points depending on their geometry, but vectors \mathbf{C}_i and \mathbf{C}_b stay the same for all the surface. Thus, the value of each element in \mathbf{C}_L can be presented as a linear combination of \mathbf{C}_i and \mathbf{C}_b with $m_i(x, y)$ and $m_b(x, y)$ coefficients different at each point.

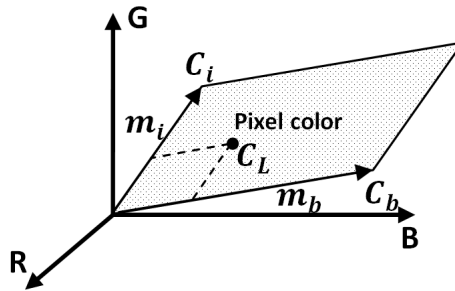


FIGURE 2.6: Color formation due to the dichromatic reflection model.

2.2.2 Bidirectional reflectance distribution function

As it has been already mentioned, the dichromatic reflectance model is a generalization of all the existing BRDFs. In this section, we aim at discussing the main definitions and properties of the BRDF which are useful to understand the image formation process. The explanations presented here are based on works of [45] and [46]. For a more detailed study, we send the reader to primary sources.

Definition

A BRDF describes how much light is reflected in the camera direction when it interacts with a surface made of a certain material. There are two categories of BRDFs in the literature: isotropic and anisotropic. The isotropic BRDF is invariant when the surface is turned around its normal vector. It means that if we have a flat smooth surface illuminated by a light source and photographed by a camera, intensities of object pixels are preserved by rotations of this surface around its normal vector, thus the BRDF is invariant to these rotations. And vice versa, the anisotropic BRDF is changed during such a rotation. Most real-world materials have anisotropic BRDFs, but to simplify modeling they are frequently supposed to be isotropic (wood, plastic, etc.). We also consider surface to be isotropic.

The light resulting from different wavelengths is absorbed and reflected differently, thus the BRDF depends on the wavelength λ of the incoming light. Moreover, the BRDF can be position variant and different from one surface point (x, y) to another even for the same material. In the following explanation we consider these two properties but simplify the representation of the BRDF as $\text{BRDF}_\lambda(\theta_i, \phi_i, \theta_o, \phi_o, x, y) = \text{BRDF}(\theta_i, \phi_i, \theta_o, \phi_o)$, where (θ_i, ϕ_i) and (θ_o, ϕ_o) are respectively the incident and the reflected light directions in spherical coordinate system for the unit vectors. This is a classical presentation of the BRDF. For the following sections, the BRDF is always supposed to be spatially (or position) variant. Moreover, the BRDF remains dependent on the incident light wavelengths, even if it is not shown in notation or mentioned explicitly in the given descriptions and explanations.

The BRDF is defined as the ratio of the quantity of light L_o reflected in the considered direction (θ_o, ϕ_o) to the quantity of light E_i that reached the surface for the incoming direction (θ_i, ϕ_i) :

$$\text{BRDF}(\theta_i, \phi_i, \theta_o, \phi_o) = \frac{L_o}{E_i}, \quad (2.10)$$

where (θ_i, ϕ_i) and (θ_o, ϕ_o) present respectively the incident and the reflected light directions in a spherical coordinates system. The quantity of light which reaches the surface can be evaluated as:

$$E_i = L_i \cos(\theta_i) d\omega_i, \quad (2.11)$$

where $d\omega_i$ is a differential solid angle of the incident light corresponding to (θ_i, ϕ_i) . Having the definition of the BRDF we can present the general lighting equation which defines the illumination of the surface point as seen by an observer (or captured by a camera) as the following:

$$L_o = \int_{\Omega} L_{io}(\theta_i, \phi_i, \theta_o, \phi_o) d\omega_i, \quad (2.12)$$

where $L_{io}(\theta_i, \phi_i, \theta_o, \phi_o)$ is the amount of light received from the direction (θ_i, ϕ_i) and reflected to the direction (θ_o, ϕ_o) , Ω denotes a hemisphere of the incoming light directions. In the discrete case, the integral can be replaced by the sum of light quantities from all directions corresponding to a solid angle $d\omega_i$.

For M light sources present around the object surface, the general lighting equation is defined as a superposition of the light from each source reflected in the considered direction (θ_o, ϕ_o) :

$$L_o = \sum_{j=1}^M \text{BRDF}(\theta_i^j, \phi_i^j, \theta_o, \phi_o) L_i^j \cos(\theta_i^j). \quad (2.13)$$

Depending on the task, the BRDF can be also presented as a function of two vectors: vector of normal \mathbf{n} and vector of light source direction \mathbf{v} , $\text{BRDF}(\mathbf{n}, \mathbf{v})$, or vector of light source \mathbf{v} and vector of viewing \mathbf{s} , $\text{BRDF}(\mathbf{v}, \mathbf{s})$. It can also include a vector of the parameters which are not related to the light direction or surface normal (e.g., surface roughness), and then be written as $\text{BRDF}(\mathbf{n}, \mathbf{v}, \alpha)$.

Properties

All BRDFs have physical properties which can be also taken into consideration during their usage. As many physical characteristics, the BRDF is always positive:

$$\text{BRDF}(\theta_i, \phi_i, \theta_o, \phi_o) \geq 0. \quad (2.14)$$

Another important BRDF property is the reciprocity principle which is also called Helmholtz reciprocity, illustrated by Fig. 2.7. This principle states that the BRDF remains the same if the incident and the emergent directions are exchanged:

$$\text{BRDF}(\theta_i, \phi_i, \theta_o, \phi_o) = \text{BRDF}(\theta_o, \phi_o, \theta_i, \phi_i). \quad (2.15)$$

The reciprocity principle is very important and can be used for the surface reconstruction, [47]. Surface reconstruction using reciprocity requires a specific images acquisition device which is capable to capture multiple reciprocal image pairs which are further used for the resolution of the system of constraints.

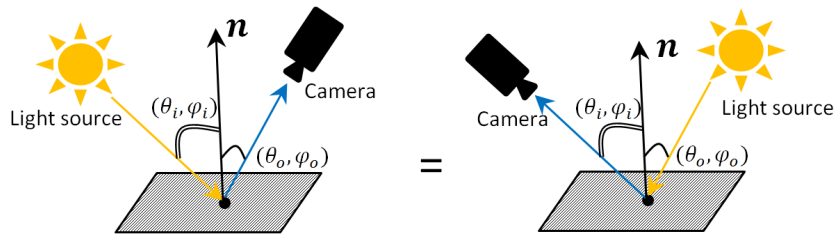


FIGURE 2.7: The BRDF reciprocity principle.

In order to be physically plausible, the BRDF must obey the law of energy conservation, Fig. 2.8. Light vectors lengths in this figure are proportional to the quantity of light reflected into different directions. This principle states that for each unit of light energy which arrives to the surface point, no more than one unit of the light energy can be reflected in all possible spatial directions. It can be written as:

$$\int_{\Omega} \text{BRDF}(\theta_i, \phi_i, \theta_o, \phi_o) \cos(\theta_o) d\omega_o \leq 1, \quad (2.16)$$

where \int_{Ω} represents an integral over a hemisphere, (θ_i, ϕ_i) characterizes the incident light direction in spherical coordinates, (θ_o, ϕ_o) denotes the reflected light direction in spherical coordinates, and $d\omega_o$ is a differential solid angle of the outgoing light corresponding to (θ_o, ϕ_o) .

Theoretical modeling of the BRDF is not always possible and not defined for all types of materials, nevertheless it can be measured empirically using a measuring device such as a gonioreflectometer, [48]. The BRDF measurements can be very useful for applicative tasks when one aims at modeling certain materials (e.g., 3D rendering of painted objects made of pre-defined materials).

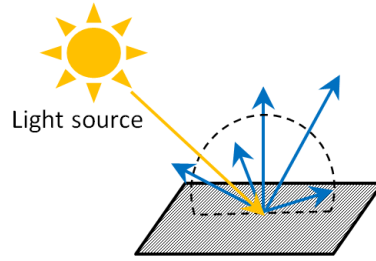


FIGURE 2.8: The BRDF energy conservation principle.

Categorization of surfaces into non-Lambertian and Lambertian is the most common one when we are talking about specular component consideration or not. In this section we present a standard representation of Lambertian model vs an example of the non-Lambertian Phong model BRDF.

Lambertian BRDF

The BRDF of a Lambertian surface can be seen as a specific simplified case of the general BRDF. The surface is Lambertian if it reflects the incident light uniformly in all possible spatial directions, Fig. 2.9(a). It means that the BRDF of the Lambertian surface is invariant to the reflecting directions (θ_i): for the same lighting direction, the camera receives the same quantity of light in any spatial position. The Lambertian surface model is the most popular one found in the literature. Despite its simplicity, it can be efficiently used for many surface reconstruction tasks, especially when there is little prior information on image(s) acquisition process.

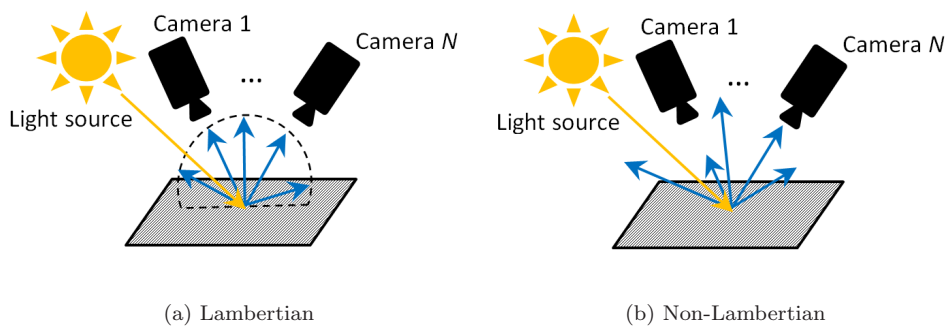


FIGURE 2.9: Lambertian (a) and non-Lambertian (b) surface BRDFs.

The equation for a single light source for the Lambertian surface is the following:

$$L_o = L_i \frac{\rho}{\pi} (\mathbf{v} \cdot \mathbf{n}) = L_i \frac{\rho}{\pi} \cos(\theta_i), \quad (2.17)$$

where \mathbf{v} is the lighting direction, \mathbf{n} is the surface normal in 3D space, $\|\mathbf{v}\| = \|\mathbf{n}\| = 1$, θ_i is the angle of the incident light and ρ denotes the surface albedo (ability of the surface to reflect/absorb the incident light). The fact that Lambertian surfaces do not depend on the viewing position \mathbf{s} considerably simplifies the equation and makes possible surface reconstruction (evaluation of \mathbf{n}) with a very small number of input images (1 for the shape-from-shading technique and minimum 3 for the PS) and without a large amount of prior information on the acquisition system.

The BRDF of the Lambertian surface can therefore be written as:

$$\text{BRDF}_{\text{Lamb}}(\theta_i, \phi_i) = \frac{\rho}{\pi}. \quad (2.18)$$

Thus the Lambertian BRDF is constant and proportional to the surface albedo. The surface reconstruction is realized using Eq. 2.17, in which L_o is supposed to be captured by the camera and presented by digital images.

The Lambertian BRDF as any other BRDF depends on the incident light wavelength and can vary from one surface point to another. Therefore during the Lambertian surface reconstruction, calculations are channel-wise and pixel-wise as well.

Non-Lambertian BRDFs

The Phong model, [49], is one of the most widespread non-Lambertian BRDFs, Fig. 2.9(b), which is used in computer graphics for 3D objects rendering. An equation for a single light source, due to the Phong model, is of the following form:

$$L_o = L_i(k_d(\mathbf{v} \cdot \mathbf{n}) + k_s(\mathbf{r} \cdot \mathbf{s})^n), \quad (2.19)$$

where \mathbf{v} and \mathbf{s} are the lighting and the viewing directions respectively, \mathbf{n} is the surface normal in 3D space, \mathbf{r} is the normalized vector in the direction of the angular bisector of \mathbf{s} and \mathbf{v} , k_d and k_s are coefficients which define portions of diffuse and specular component for the reflected quantity of light, and n is a power term to regulate the width of the specular highlight. Thus, the Phong model allows to represent not only diffuse surface material but also glossy materials with specular component. The vector \mathbf{r} is called a perfect reflecting direction. Here we can just mention that for \mathbf{n} bisecting the angle between the light incidence and reflectance to the camera direction, specular component coefficient k_s is the highest. Thanks to [46] and some

mathematical transformations, the BRDF of the Phong model can be calculated as:

$$\text{BRDF}_{\text{Phong}}(\theta_i, \phi_i, \theta_o, \phi_o) = \frac{k_d(\mathbf{v} \cdot \mathbf{n}) + k_s(\mathbf{r} \cdot \mathbf{s})^n}{\cos(\theta_i)} d\omega_i. \quad (2.20)$$

The Phong model became popular thanks to its few necessary parameters and an understandable tuning of these parameters to achieve a desired visual effect. However, this model is not the only one for simultaneous diffuse and specular reflection representation, there are also Cook-Torrance model, [45], which is popular for highly specular metallic objects rendering, Oren-Nayar model, [50], which is convenient for rough materials reflectance description, etc.

2.3 Specular surface reconstruction

In this section we present the existing techniques for non-Lambertian opaque surface reconstruction with their advantages and limitations. We chose to describe the state-of-the-art methods using a division into two categories: reconstruction methods taking into account specularities as a part of an image formation process and methods considering specular pixels as outliers of the Lambertian model. Approaches presented in this section seem to us the most representative, but for a broader review as well as for a more detailed classification of methods we invite the reader to study [51] and all primary sources cited below.

2.3.1 Specular surface reconstruction methods

Specularities, so useful for human 3D shape perception, affect the surface reconstruction process. As described in Section 2.2.1, the specular lobe can be modeled by the means of different mathematical formulations such as Phong, Cook-Torrance and Beckmann models, Fig. 2.5. However the model which best fits the specular effect on the image as well as the reconstructed surface normals depends on object surface properties and is not known a priori. Methods for surface reconstruction based on specularities exploit different reflectance properties of the specular effect and different models of specular luminance distribution as supplementary information to resolve ambiguities arising during the surface normals estimation.

Categorization of the PS approaches for specular surface is not an easy task. In most cases, authors implement a dedicated reconstruction algorithm for a specific

material or use a particular acquisition device. We categorize all the presented techniques depending on the method employed to obtain the initial images, with different light source-surface-camera interactions. PS permits to obtain a 3D surface on the basis of images with different repartitions of light on the reconstructed object surface. The variety of shading can also be obtained by changing the positions of the acquisition process actors. Each of the presented methods considers one or a combination of the following assumptions:

- changing a camera position (fixed light source(s) and object);
- object movement with a special mechanism (fixed light source(s) and camera);
- changing lighting directions (fixed camera and object).

In the standard photometric stereo approach, it is supposed that the position of camera and object are fixed and the lighting directions vary from one acquisition to another. So, only the last category of this list can be considered as initial data for photometric stereo.

Specular surface reconstruction with camera or object movement

An interesting work on specular surfaces reconstruction is proposed in [52]. The authors of this article study an interdependence between a movement of specular highlights and a movement of the viewing direction (an imaging sensor position). For this method we have a camera which changes its position from one acquisition to another. The linear system of the mentioned dependency can be expressed as:

$$2\mathbf{s} \cdot (\mathbf{MH} - k_{sv}I)\mathbf{x} = w, \quad (2.21)$$

where \mathbf{s} is the viewing direction, $\mathbf{M} = \begin{pmatrix} \sec(\sigma) & 0 \\ 0 & \cos(\sigma) \end{pmatrix}$, σ denotes the angle of reflection, $\mathbf{H} = \begin{pmatrix} z_{xx} & z_{xy} \\ z_{yx} & z_{yy} \end{pmatrix}$ is the Hessian matrix of the surface, $k_{sv} = \frac{1}{2} \left(\frac{1}{\|\mathbf{s}\|} + \frac{1}{\|\mathbf{v}\|} \right)$, \mathbf{v} is the vector of lighting direction, I is the image intensity, $\mathbf{x} = (x, y)$ and $w = (-d_1 + d_3 \tan(\sigma) - d_2)$, $\mathbf{d} = (d_1, d_2, d_3)$ describes the viewer movement. Combining information from different viewing directions, it is possible to estimate surface curvature. Convex-to-concave surface ambiguity is solved by the means of the surface

Hessian matrix analysis. For the presented algorithm the viewing and the lighting directions should be controlled and accurately known for each input image to estimate the correct surface.

In [53] the authors proposed an algorithm of specular surface reconstruction which uses the images made with radial light sources illuminations, Fig. 2.10. These sources are located around the object and the acquisition is made by a fixed camera. After that, rotating the object, one can observe changes in highlight stripes. This information is accumulated and then used for the object surface geometry estimation (it is the so-called specular motion stereo).

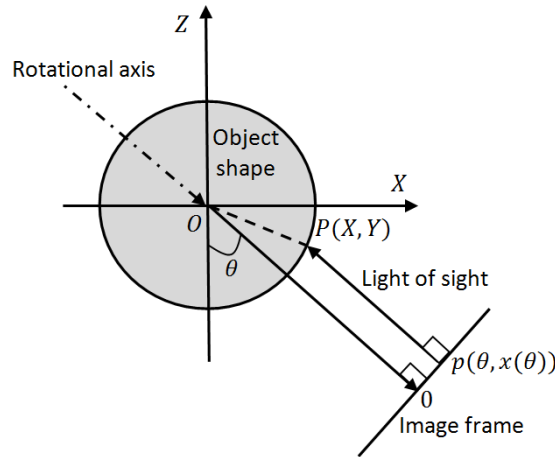


FIGURE 2.10: Specular motion stereo from [53]

Helmholtz reciprocity principle usage (Section 2.2.2) for surface BRDF extraction (Helmholtz stereopsis) is presented in the algorithms from [54] and [47]. In the reciprocity principle it is supposed that both camera and illumination source can change their position, while the photographed object does not move. These methods allow making no prior assumptions on the surface BRDF which is useful for real-world tasks. The image intensity generated in camera c when the surface point \mathbf{p} is illuminated by the light source l can be given by:

$$I_{c,l} = \rho \text{BRDF}(\mathbf{v}_l, \mathbf{v}_c) \mathbf{n} \cdot \mathbf{v}_l \frac{1}{\|\mathbf{o}_l - \mathbf{p}\|^2}, \quad (2.22)$$

where ρ represents the albedo value, \mathbf{v}_l is the direction of lighting, \mathbf{v}_c corresponds to the direction of viewing, \mathbf{n} denotes the surface point normal and $\frac{1}{\|\mathbf{o}_l - \mathbf{p}\|^2}$ is the specular falloff. The light source and the camera positions are exchanged for the input data acquisition. If the reciprocal case is considered, then the image irradiance can be

rewritten as:

$$I_{l,c} = \rho BRDF(\mathbf{v}_c, \mathbf{v}_l) \mathbf{n} \cdot \mathbf{v}_c \frac{1}{\|\mathbf{o}_c - \mathbf{p}\|^2}. \quad (2.23)$$

The intensity of image pixel (the BRDF) for such replacement should remain the same:

$$BRDF(\mathbf{v}_l, \mathbf{v}_c) = BRDF(\mathbf{v}_c, \mathbf{v}_l). \quad (2.24)$$

It is thus possible to eliminate the value of the BRDF and find the surface normals \mathbf{n} from the equation:

$$\left(I_{l,c} \mathbf{v}_l \frac{1}{\|\mathbf{o}_l - \mathbf{p}\|^2} - I_{c,l} \mathbf{v}_c \frac{1}{\|\mathbf{o}_c - \mathbf{p}\|^2} \right) \mathbf{n} = 0. \quad (2.25)$$

Unlike the basic PS methods, the Helmholtz stereopsis is a multi-view method, and the camera position is changed as well as the lighting direction for the input images acquisition. For each camera-light source position a pair of images is taken. Thus, combining the conventional stereo and the Helmholtz reciprocity, the authors provide an efficient reconstruction algorithm. A combination of the binocular and the PS is also presented in the work of [55]. In this paper, the authors propose a new measurement for the surface reconstruction (Light Transport Constancy) which permits to introduce a rank constraint on multi-view stereo matching when the scene is photographed with varying lighting for each acquisition. Their algorithm needs neither a priori knowledge on the light source positions nor a calibration object on the scene.

The so-called shape-from-distortion method is presented in [56] and extended in [57]. Totally specular surfaces (mirror, smooth metallic surface, etc.) reflect surfaces of other objects around. If these reflected surfaces are planar with a regular texture, then the distortion of the reflected patterns can be related to the orientation of the specular reflecting surface, as shown in Fig. 2.11. The most difficult part of the algorithm is to find correspondence between the original and reflected pattern. Unfortunately, not all specular surfaces can reflect surrounding objects in this way. In the majority of cases specularities lead to intensity over-saturation for some object parts, all other parts reflect the incident light almost as surfaces of Lambertian nature. The shape-from-distortion technique is presented in [58] as well. But the algorithm in this paper can be classified as active stereo more than photometric stereo. The authors proposed to project 2D patterns of different frequencies onto a mirroring object, after that the distortion obtained for all patterns is used for the normal and the depth map reconstruction. Shape-from-distortion methods require

only one image with fixed object, lighting source and camera positions, however they could be applied only for totally mirroring object reconstruction.

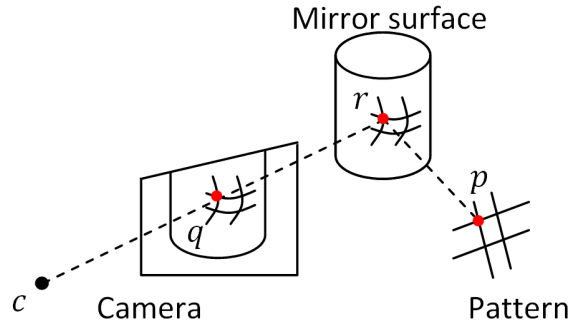


FIGURE 2.11: Shape-from-distortion setup from [57].

An idea for the surface representation as a linear combination of several materials (Lambertian or non-Lambertian) is illuminated in works ([59], [60]):

$$I_{i,p,c} = \sum_m \gamma_{p,m} \cdot BRDF_c(\mathbf{n}_p, \mathbf{v}_i, \alpha_m), \quad (2.26)$$

where $I_{i,p,c}$ is the intensity of the p^{th} pixel in the c^{th} channel of the i^{th} image, m denotes different materials, $\gamma_{p,m}$ is the weight of the m^{th} material in the p^{th} pixel, the reflectance function $BRDF_c$ corresponds to the c^{th} color channel, \mathbf{n}_p is the pixel normal, \mathbf{v}_i is the lighting condition and α_m is the vector of the BRDF parameters. This approach is based on the observation that the majority of objects surfaces can be presented as a composition of a small number of fundamental materials. In general, the BRDFs of composing material are non-Lambertian and may include specularities. These algorithms permits not only to recover the orientation and the BRDF value for each point surface, but also the values of weights for all composing materials, which can be useful for further automated processing (e.g., clustering). The algorithm from [59] needs preliminary light calibration while the authors in [60] require no calibration, but demand presence of the reference object on the scene which is made of the same material or at least a material with close enough reflectance properties. Reconstruction is then performed on the images taken with a fixed object and camera but with varying illuminations.

In the article [61] the authors use a specific imaging device presented earlier in [62] in order to scan fine-scale details of specular or partially specular surfaces. The mentioned imaging device consists of a concave parabolic mirror to obtain multiple angles in a single image and a camera. Different incident light directions are created

using parallel light which is shifted by a movable aperture. Thus we can obtain dense BRDF 2D slices. The surface normals can be reconstructed by the means of the slices with maximal intensity values (corresponding to specularities). Texture color and surface orientation are measured simultaneously by the presented acquisition device.

The authors of [63] were the first to introduce the concept of specular flow. An image motion observed by a moving camera can be seen as a combination of the diffuse and the specular flows (according to the diffuse and the specular surface components present on the image). The authors propose an Expectation-Maximization algorithm to relate probabilistic mixture of the diffuse and specular flows and the 3D surface orientation. In [64] the principal of the specular flow is used as well. Due to their approach only the environment map should move, and the relative camera-object position should remain unchanged. The authors proposed an algorithm based on the resolution of non-linear partial differential equations for the special case of rotation around the camera optical axis to obtain the surface orientation of the environment which is neither calibrated nor known.

In [65] the authors exploit three properties of the reflectance (monotonicity, visibility and isotropy) with respect to the varying lighting direction and the fixed surface orientation. The image pixel intensity I_p can be represented by:

$$I_p = f(L BRDF(\mathbf{n}_p, \mathbf{v}_i) + A), \quad (2.27)$$

where f is a radiometric response function, L is the light source intensity, \mathbf{n}_p denotes the surface point normal, \mathbf{v}_i is the lighting source direction and A corresponds to the ambient light component. Monotonicity (\mathfrak{N}_1) is expressed as:

$$\mathfrak{N}_1 = \left\{ \mathbf{n} \in \mathbb{R}^3 \mid \bigcap_{i,j} ((\mathbf{v}_i - \mathbf{v}_j) \cdot \mathbf{n} > 0) \right\} \quad (2.28)$$

for pairs of \mathbf{v}_i and \mathbf{v}_j that satisfy $BRDF(\mathbf{n} \cdot \mathbf{v}_i) > BRDF(\mathbf{n} \cdot \mathbf{v}_j)$. Visibility (\mathfrak{N}_2) can be defined as:

$$\mathfrak{N}_2 = \left\{ \mathbf{n} \mid \bigcap_i (\mathbf{n} \cdot \mathbf{v}_i > 0) \right\}. \quad (2.29)$$

Isotropy (\mathfrak{N}_3) from m -sets lighting directions is expressed as:

$$\mathfrak{N}_3 = \left\{ \mathbf{n} \mid \mathbf{n} = \sum_m a_m \mathbf{n}_m, a_m \leq 0 \right\}. \quad (2.30)$$

Each of these properties (constraints) already gives a space of solutions for the surface orientation estimation. The intersection of the solution spaces (so-called consensus PS) provides the resulting surface orientation:

$$\mathfrak{N} = \mathfrak{N}_1 \cap \mathfrak{N}_2 \cap \mathfrak{N}_3. \quad (2.31)$$

This method can give rather accurate results (accuracy of less than one degree) providing that a large number of input images is used (50 and more).

In conclusion, all the presented methods of specular surface reconstruction with camera or object movement require specific devices for the algorithm initial images acquisition. Some of the methods need a pixel correspondence problem solution which is considered as a complex supplementary procedure. Moreover, almost all of the discussed approaches require prior information either on illumination directions, or on the reconstructed object properties.

Photometric stereo reconstruction of specular surface

One of the approaches for the specular surfaces photometric stereo reconstruction is presented by Healey and Binford in [66]. The camera position and the surface are fixed for this method, while lighting varies to generate the different input images. The authors of this article use the physical Cook-Torrance model for specular highlights and show how information on the curvature of the surface part (magnitudes and directions) can be extracted from a single highlight. The formal definition of the Cook-Torrance model for the specular image component is the following:

$$\mathbf{I}_{Spec} = FDA, \quad (2.32)$$

where F is the Fresnel coefficient, D denotes the facet orientation distribution function and A is a geometrical attenuation factor. The variable A is derived from the Cook-Torrance model and can be expressed as:

$$A = \frac{G}{\mathbf{n} \cdot \mathbf{s}}, \quad (2.33)$$

$$G = \min \left\{ 1, \frac{2(\mathbf{n} \cdot \mathbf{r})(\mathbf{n} \cdot \mathbf{s})}{(\mathbf{s} \cdot \mathbf{r})}, \frac{2(\mathbf{n} \cdot \mathbf{r})(\mathbf{n} \cdot \mathbf{v})}{(\mathbf{s} \cdot \mathbf{r})} \right\}, \quad (2.34)$$

where \mathbf{n} is the surface point normal, \mathbf{s} represents the viewing direction, \mathbf{v} corresponds to the lighting direction and \mathbf{r} is the angular bisector of \mathbf{s} and \mathbf{v} . Under the

assumption about equality of viewing direction \mathbf{s} (which is supposed to be known) and ideal specular reflecting direction, the authors introduce an algorithm for the estimation of the surface normals \mathbf{n} . In the article it is also supposed that the regions of glossy surfaces which do not contain specularities are probably planar, and the surface corners and edges can be extracted from specularities. In the context of the PS, this algorithm can exploit a lot of input images with different specular zone location for each image related to the different light sources positions used for each acquisition. The surface regions not covered by specularities can be interpolated using normals found on the basis of specular pixels. The more complex the surface we aim at reconstructing, the more difficult it is to make this interpolation and to obtain the proper surface.

It should be mentioned that even if the methods of specular surface reconstruction are widely presented in the literature, each of them was designed for a particular type of data or capturing device and uses certain prior knowledge on the acquisition system or the observed surface which facilitates the evaluation process but deprives the algorithm universality. Due to the used acquisition device, we are limited to a standard set of images for PS for which the camera and the reconstructed object are fixed, while illumination directions are varying to create the images. This limitation significantly decreases the number of methods with a direct consideration of specularities which can be exploited for microbial colonies reconstruction. If we also include lack of prior knowledge on lighting directions and surface properties, as well as an absence of reference patterns in the scene, none of the presented methods is sufficient to solve our problem of non-Lambertian surface reconstruction. Therefore, we redirected our attention to the methods treating specularities as outliers and then using Lambertian PS principle.

2.3.2 Outlier specular pixels

In the literature we can find several algorithms of non-Lambertian surfaces reconstruction which nevertheless utilize the described Lambertian reflectance model. In these algorithms the authors suppose specularities to be outliers of the image pixel intensity values, and hence Lambertian surface representation (normal and albedo values) for specular pixels has significant deviations which should be detected and processed. All the methods of this type use a classical setup for initial image acquisition: camera and reconstructed object are fixed, while light is varying from one image acquisition to another. We can distinguish three categories of these methods:

specularities post-processing, non-specular pixel intensities selection and specularities correction before reconstruction.

Specularities post-processing

The work of Coleman and Jain, [67], can be boldly considered as the first reconstruction approach presented for non-Lambertian surface treatment. They propose to use four images of specular surface made from a fixed camera position and varying illuminations (4-source photometry). The basic Lambertian PS model, [24], is applied to each possible triplet ([1,2,3], [2,3,4], [1,3,4], [1,2,4]) of input images to evaluate surface normals and albedo value for each image pixel. The albedo value for the i^{th} triplet is evaluated as:

$$\rho_i(x, y) = |[\mathbf{V}_i]^{-1} \hat{I}(x, y)|, \quad (2.35)$$

where $\hat{I}(x, y)$ is a vector of the three image intensity values and \mathbf{V}_i is the matrix of light sources known directions for the i^{th} triplet of images. The pixels for which the resulting albedo have the deviation superior to a pre-defined threshold are considered as specular ones. The deviation is calculated as:

$$\rho_{dev}(x, y) = \frac{\sum_{i=1}^4 (\rho_i(x, y) - \rho_{mean}(x, y))}{4\rho_{min}(x, y)}, \quad (2.36)$$

where $\rho_{mean}(x, y) = \text{mean}(\rho_i(x, y))$ and $\rho_{min}(x, y) = \min(\rho_i(x, y))$, $i = 1, \dots, 4$. Thus, surface normals are equal to the average value of the normals calculated on the basis of all possible triplets for each pixel of Lambertian nature (i.e., without the albedo outlier), or as a normal of the image permutations with the smallest resulting albedo value for specular surface regions. The algorithm of [67] needs either initial information of the lighting directions or the preliminary calibration procedure for these direction approximations. It is not efficient for large specular image zones and may cause resulting object shape deformations on the border between specular and non-specular regions. This method can also be seen as a posterior specularity filtering. It is rather resource consuming: it demands the albedo value evaluation for each possible triplet (i.e., to apply four times the PS algorithm) before making the choice on normals which are not affected by specularities.

The authors of [68] proposed to expand the Coleman and Jain 4-source photometry by adding a processing of RGB color images:

$$I^k(x, y) = (I_1^k(x, y), I_2^k(x, y), I_3^k(x, y))^T = (\mathbf{V}^k \mathbf{n}(x, y)) \cdot \mathbf{C}(x, y), \quad (2.37)$$

where $\mathbf{C} = (C_r, C_g, C_b)$ is a vector of the surface point colors, k is a triplet of the input images, $I^k(x, y)$ is the image pixel intensity, \mathbf{V}^k is the matrix the light source directions for the k^{th} triplet and \mathbf{n} is the surface point 3D normal. The proposed algorithm permits not only specular but also shadowed pixel detections in order to eliminate them from consideration in the surface reconstruction procedure. Specularities and shadows are detected thanks to the chromaticity of the object and the color source analysis. This method needs knowledge both on the lighting directions used during image acquisitions and on the lighting source colors. Decision on elimination is based on the chosen threshold values which should be empirically defined. The authors of [69] propose to expand the 4-source photometry to an arbitrary number of light sources, iteratively discarding the pixels of images related to shadows or specularities.

To conclude, the presented methods of this category can be rather efficient for specular surfaces, but finally each of them exploits basic Lambertian photometric stereo scheme with prior knowledge on light sources directions. Furthermore, these methods are time-consuming, as it is necessary to realize photometric stereo algorithm several times before deciding which pixel intensities do not contain specularities.

Non-specular pixel intensities selection

The second category of specular processing method consists in non-specular pixels selection before the Lambertian PS application. Two algorithms from [70] and [71] of 5- and 6-source photometry can be considered as methods of this category. The pre-processing consists in simply discarding specular pixels before the surface normal estimations, Fig. 2.12. The final decision about the pixels belonging to the class of specular or shadowed values is made on the basis of relative pixel intensity values comparing between different images. The lowest and the highest intensities are not further used for reconstruction. The main goal pursued by the authors of the presented algorithm is to preserve the Lambertian model simplicity.

For the algorithms of specularities post-processing and non-specular pixels selection, it is necessary to have more than three input images for the reconstruction algorithm.

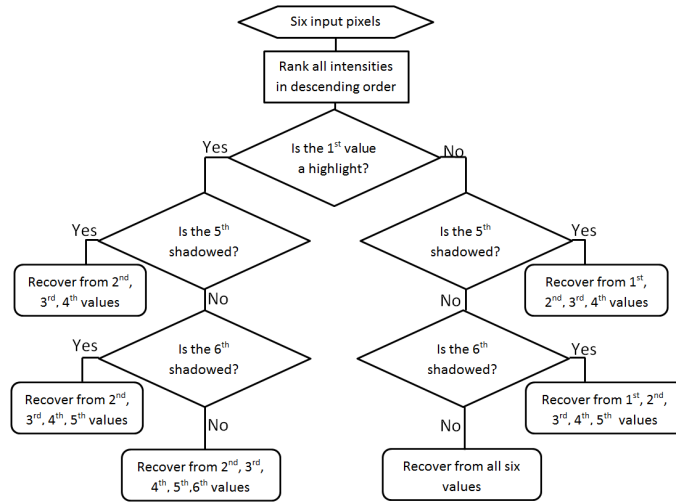


FIGURE 2.12: 6-source PS with specularities and shadows discarding from [71].

The presence of additional images permits a better detection of specular pixels as outliers. However, the choice of some pixel intensities and discarding others may provoke sharp passage between different surface regions.

Specularities correction before reconstruction

The authors of [72] propose to use a color space transformation for each initial image and realize specular removal as a PS pre-processing phase. The main principle of the algorithm is not to find specular and diffuse component of each pixel, but to find a transformation of the RGB color space which permits to avoid specularities. Due to the dichromatic reflection model, a set of color vectors of a dichromatic material lie in a dichromatic plane. This plane is spanned by two vectors: the vector of the interface reflection color (\mathbf{C}_i in Fig. 2.6) and the color of body reflection (\mathbf{C}_b). The authors propose to rotate the RGB color space so that one of the axis (R, G or B) becomes aligned with the direction of the effective light source. This transformation preserves general image shading information. The rotation permits to obtain a new 3-channel image (so-called SUV-image) with 2 channels (U and V) without specular component. Finally, the standard PS algorithm is applied to U and V channels, and thus the resulting surface does not suffer from the distortions related to specularities. This method requires the information on the light source color.

Methods of specularity pre-processing for PS require supplementary analysis of input images or these images transformation to exclude specular component. After such a treatment, the methods of this category permit to exploit the Lambertian PS model,

which significantly reduces the number of BRDF parameters to be known a priori. Nevertheless, the models used in the presented approaches still require information on light sources directions.

2.4 Proposed chain for uncalibrated non-Lambertian photometric stereo

In the previous section we presented several existing approaches for non-Lambertian surface reconstruction, but all of them need prior knowledge either on light sources or on surface properties. In the thesis work we aim at non-Lambertian photometric stereo reconstruction without this prior knowledge. We propose to exploit the Lambertian PS approach for non-Lambertian surfaces after images pre-processing. Lambertian surfaces can be reconstructed using the color channel-varying BRDF independent on the viewer (camera) position, Section 2.2.2. This reduces the number of unknown model parameters and gives us the possibility to adapt the reconstruction algorithms to the hypothesis of unknown lighting directions used during the input images acquisition: the so-called uncalibrated photometric stereo. In the hypothesis of Lambertian surface, the assumption of specularities is most violated. Thus, the specular effects should be specifically processed before application of the reconstruction algorithm. In our approach we do not aim at discarding the specular pixels, as it is done in the majority of algorithms of the same category, but we aim at correcting specular zones using the information on the pixel intensities present on all images of the input sequence. This enable to reconstruct surface using only three images of the initial sequence and to avoid image artifacts related to the borders of specularities.

Ideologically, our approach is close to the method proposed in [72] (correct specularities before applying a Lambertian model), but it needs no prior information on the light source(s) chromaticity.

The complete chain for our PS surface reconstruction is given in Fig. 2.13. It consists of two main stages:

- specularity correction;
- uncalibrated surface PS reconstruction;

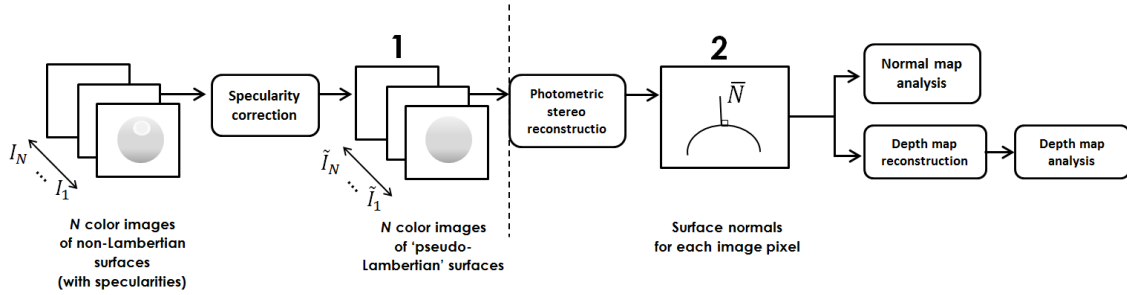


FIGURE 2.13: Proposed uncalibrated non-Lambertian surface reconstruction.

Each of these stages is very important for the complete processing, but the chain is modular: processing steps can be realized as presented in this thesis or replaced by any other existing algorithms developed for the same purpose and with the same inputs and outputs format as the presented stages.

2.4.1 Specularity correction

The first stage of the processing is the correction of specularities, in Fig. 2.13. The input data consist of N ($N \geq 3$) RGB color images of the non-Lambertian surface which were made from a fixed camera position and with varying illumination directions for each image. Such an acquisition concept is standard for the PS. Due to the fact that the photographed surfaces are non-Lambertian, optical effects such as specularities can occur for any input image. The location of this optical effect depends on the illumination direction and the surface orientation. If the photographed object remains the same and the illumination used for each acquisition is sufficiently different, the specularities will be located in different parts of the object on the input images. Thus, the information about the object surface hidden by the specularities on one image can be extracted using other images from the initial sequence. On the basis of this assumption, we propose an original specularity correction procedure which permits to obtain images of 'pseudo-Lambertian' surfaces. The most important property of this procedure, is the ability to correct specular zones while preserving shading particularities important for the further PS. Chapter 3 of the present work is dedicated to the proposed algorithm of specularity correction.

2.4.2 Uncalibrated photometric stereo reconstruction

The second phase of the processing is the surface normals reconstruction, in Fig. 2.13. Evaluation of the normals is proposed using a PS technique for the Lambertian surfaces. Representation of the PS problem as a matrix factorization task and the proposed algorithm of alternating optimization with structure, value and quality constraints permits to separate surface normals and albedo into the proper matrices. The main advantage of the proposed algorithm compared to other existing techniques consists in no prior knowledge usage. The demonstrated iterative algorithm uses only the pre-processed images of the 'pseudo-Lambertian' surface and no other information neither on the lighting directions used during acquisitions, nor on the camera or on the reconstructed object properties. This algorithm does not use any calibration object on the scene as well. The proposed algorithm is efficient and rather fast for microbial objects surfaces reconstruction as well as for any other images made for the PS purposes. We present details of the uncalibrated photometric reconstruction algorithm in Chapter 4.

In the following, we present the detailed description of the different stages of the proposed processing pipeline and their algorithmic implementation.

Chapter 3

Specularity correction

In this chapter we present an original method for specularities correction. Such a specular component elimination allows obtaining images of 'pseudo-Lambertian' surface with only diffuse reflectance. This diffuse component can be further used for Lambertian photometric stereo surface reconstruction.

Specularity is an existing optical phenomenon characteristic for some materials and absent for Lambertian surfaces. In an image, specularities appear as abnormally elevated intensity (often over-saturated) values which do not contain the properties of the photographed object. Presence of specularities in an image gives us ideas about the photographed object surface: smooth and rough objects have different repartition of specularities on the surface. If two different objects are made of the same material and photographed with the same direction of illumination and the same position of the camera, a more elevated object has a specular zone larger than a less elevated one. Specularities are also used as additional cues for techniques as 'shape-from-specularities' or 'shape-from-motion-and-specularities'. Useful properties of specular zones from perceptual and computational points of view have been already observed and demonstrated by Blake et al. in [73]. Moving stereoscopic surface observations are indispensable for specular motion model composition and estimation of surface normals. The authors in [74] propose to exploit specularities as additional constraints to compose a partial differential equation with additional smoothness term subject to known light sources, a camera position and a structure-from-motion approach implementation. Correctly detected specularities may be even directly used for object recognition. In [75] the authors render the object using a combination of Lambertian photometric stereo model and specular component. Then the results of such a rendering are compared with all objects in

the pre-defined database. The type of reference objects which matches the best the rendered image is considered as the result of the recognition.

On the other hand, specularities affect automated image processing. For example, the presence of specularities does not allow to realize a correct segmentation of objects, taking specular pixels for parts of another object due to their intensity values. In classical stereo reconstruction tasks, specular points can affect correspondence finding between several images taken from the different viewpoints. Photometric stereo suffers from specularities as well. Abnormally elevated intensity values hide information about truthful object surface shading and cause distortion of the reconstructed surface. Under certain conditions, specularities can occupy up to 20% of the object surface, making reconstruction using a photometric stereo technique impossible.

Information on specular zone locations on surface can be useful and exploitable, but all techniques which use this information suppose at least two views of an object surface with different positions of specularities as input data. It is not the case for the photometric stereo technique. Presence of only one fixed position of a camera severely limits specularities utilization for surfaces reconstruction. Thus, specular zones should be effectively detected and removed. In order to correct specular values and to obtain images suitable for the Lambertian photometric stereo, we propose a correction algorithm which aims at specular zones correction while preserving shading particularities important for the further photometric stereo reconstruction. In other words, for an efficient surface reconstruction the effects of shading and specularities (also called highlights) should be separated.

The algorithm of specularity correction proposed in this chapter can be used as a pre-processing for the Lambertian photometric stereo, or used alone as specularities correction for any images taken from a fixed camera position and varying illumination.

3.1 Specularity correction techniques

In this section we describe existing works on specularity correction. We refer the reader to [76] in order to obtain the most complete information on the topic. In our context of further photometric stereo reconstruction, methods of specularity

correction can be separated into three large categories depending on the number of initial images used for specularity correction and the number of resulting images.

3.1.1 ' $N - to - 1$ ' methods

The first category can be called ' $N - to - 1$ ' methods. These methods use information present on several images in order to obtain only one image without specularities. There exist many different approaches of this category in the literature. Each of these methods uses particular input data and processing concept.

One of the existing techniques of specularity correction supposes utilization of polarized images, as it is presented in [77]. The diffuse component of the image tends to be unpolarized, therefore a rotation of polarization filter in front of a camera will not lead to the image changing. On the contrary, rotations of the filter lead to image intensities variations under the cosine law for the specular component. The full image intensity is thus presented as a superposition of the diffuse image component, specular constant and specular varying components. The variation of the specular component is determined by the angle of the polarization filter. If there are several polarized images with known polarization angles, it becomes possible to evaluate constant and varying components of the intensity, and then eliminate the specular ones. The main disadvantage of the polarization usage is that it cannot be applied for the most conductive materials (metallic surfaces) for which the problem of the specular component correction is of current interest. It is related to the polarization properties of this type of materials.

Another example of the ' $N - to - 1$ ' methods is presented by Feris et al. in [78]. The authors use a special capturing setup with a multi-flash camera, therefore this camera takes several images with a fixed camera position and varying illuminations. Specular zones are shifted from one image of the sequence to another thanks to the different positions of flashes. Once the images are taken, the gradient of each image is calculated. After that, specular zone of the image is detected and the intensity values of this zone are replaced with respect to the calculated generalized gradient.

In [79] authors also propose to register and to combine two initial images in order to obtain the resulting one without abnormally elevated intensities. In their processing, they firstly use the Speeded Up Robust Features (SURF) matching to find correspondence between two images, detect abnormally saturated pixel values thanks to a thresholding, after that the authors utilize the minimum gray selection algorithm

for the detected specular pixels. The minimum gray selection algorithm is simple in implementation: if the first image \mathbf{I}^1 is to be corrected, then the detected specular pixels are replaced by the minimum value of these pixels in \mathbf{I}^1 and \mathbf{I}^2 , and vice versa for the correction of image \mathbf{I}^2 .

The ' $N - to - 1$ ' methods can be efficiently used for specular objects photography improvement but it is not possible to apply the photometric stereo algorithms after such treatment because for the photometric stereo we need at least three images with different lighting directions and shading. The resulting image of the ' $N - to - 1$ ' method does not preserve the necessary difference in shading. Moreover, the methods of this category in the most cases require specific acquisition devices.

3.1.2 ' $1 - to - 1$ ' methods

The second type of methods is the ' $1 - to - 1$ ' specularity correction approaches. Contrary to the ' $N - to - 1$ ' category, the methods of this category do not exploit information present on other images of the sequence and provide the correction on the basis of the image itself. Thus, if there are N initial images, we obtain N corrected images as well, but each image is processed independently.

The algorithm of [80] is stated as a single image highlights correction method which is based on the inpainting technique. The authors in [80] use the dichromatic reflection model. The color of the illuminant is supposed to be constant over the highlight region. This color can be found thanks to the total variation inpainting minimization, [81], and it is used for the final diffuse component finding. The main disadvantage of the presented technique mentioned by the authors is the significant time of calculations characteristic for all inpainting algorithms.

In [82] the authors use bilateral filtering in order to propagate non-specular values to specular zones. They consider that the maximum diffuse chromaticity in local patch should change smoothly. In order to apply this constraint to all image pixels, the authors of the article propose to evaluate maximum chromaticity of each pixel. After that, using bilateral filter the maximum diffuse chromaticity values are propagated from the diffuse pixels to specular zones.

Pixel chromaticity direct analysis and specular-free image concept implementation is another approach of specularity correction presented in [83] and [84]. These methods are based on the dichromatic reflection model. The diffuse component is adjusted

according to the criterion of smooth transition of color between diffuse and specular regions, and is evaluated by a nonlinear shifting of specular pixels intensity and chromaticity. Both methods require the information on the illuminant color.

The method of [72], already presented in Chapter 2, is also from the '1 - to - 1' category. The main principle of the algorithm is not to find specular and diffuse component of each pixel, but to find a transformation of the RGB color space which permits to avoid specularities. Information on the light source color, which is not always accessible, is necessary for this method.

The '1 - to - 1' methods may be very different in their conceptions but they all use the dichromatic reflectance model as a basis. If one of this method is applied to each initial image, the shading of this image is preserved. Thus, the images of the sequence all together can be further used for the photometric stereo reconstruction. The main disadvantage of these methods is that they do not consider information about specular zones present on other images of the initial sequence. The particularity of photometric stereo initial data is an availability of several images of the same object with specular zones in different object parts. Thus, information hidden by specularities may already be present in initial data and it could be used for efficient specularities correction. This can be useful for the correction of large zones or zones with a complex texture.

3.1.3 'N - to - N' methods

Our proposed method is of the 'N - to - N' category. It utilizes N initial images and results also in N images but with specular zones correction. Thus it is applicable for further photometric stereo reconstruction with correction of specular zones while preserving all shading information proper for each initial image of the sequence. We did not find in the literature other existing method of the 'N - to - N' category which aims at shading preserving and treating all input images jointly.

Classification of all the mentioned methods as well as their ability to preserve image shading and their usage of information present on other images of the input sequence is summarized in Table 3.1. It can be seen that only 'N - to - N' methods contain both properties, and thus can be efficiently used for as a pre-processing for the photometric stereo technique.

# Input images	N	1	N
# Output images	1	1	N
Existing approaches	Feris&al.[78] Xu&al.[79] Nayar&al.[77]	Tan P.&al.[80] Yang&al.[82] Tan R.T.&al.[83] Shen&al.[84] Mallick&al.[72]	Our method
Shading preservation	No	Yes	Yes
Use of information present on other images	Yes	No	Yes

TABLE 3.1: Specularity correction methods classification according to the number of input/output images.

The algorithms of photometric stereo are sensitive to the presence of shadows as well as to the presence of specularities. In [85] the authors propose a possible algorithm of shadows processing in order to obtain a qualitative surface of astronomic bodies. In this work we do not discuss the question of shadows specific pre-processing. Although, the proposed technique can be potentially adapted for this purpose. In our application, the microbial objects rarely have so elevated relief to create shadows and affect a reconstruction of neighbor objects, but for other real-world tasks the problem of shadows can be crucial and should be supplementary resolved for qualitative surface reconstruction.

3.1.4 Evaluation of specularity correction

There exist many different measures to evaluate the quality of image processing results. If both the ground truth and the processed images are available, according to [86], measures can be classified as: measures based on pixel differences, correlation-based measures, edge quality measures, spectral distance measures, context measures and human visual system based measures.

The mean square error (MSE) is a pixel-based measure, that has been one of the most common and intuitive measure for many years. The main reason for its popularity is its mathematical traceability and simplicity of implementation. For the algorithm of specularities correction, the MSE between the corrected image $\tilde{\mathbf{I}}$ and the diffuse component of the non-corrected image \mathbf{I}_{Diff} , considered as the ground truth, is

calculated as the following:

$$MSE = \frac{1}{X} \frac{1}{Y} \sum_{x=1}^X \sum_{y=1}^Y (\tilde{\mathbf{I}}(x, y) - \mathbf{I}_{Diff}(x, y))^2, \quad (3.1)$$

where X and Y are the number of image rows and columns respectively, $\tilde{\mathbf{I}}$ is the corrected gray-scale image and \mathbf{I}_{Diff} corresponds to the diffuse component of the initial image. However, this measure of quality alone is not sufficient. Indeed, we aim at specularities correction with object shading preservation which is not reflected by the MSE. With the following example, we show that for different correction results only visual analysis of the processed image can provide an idea of the correction quality for shading preserving.

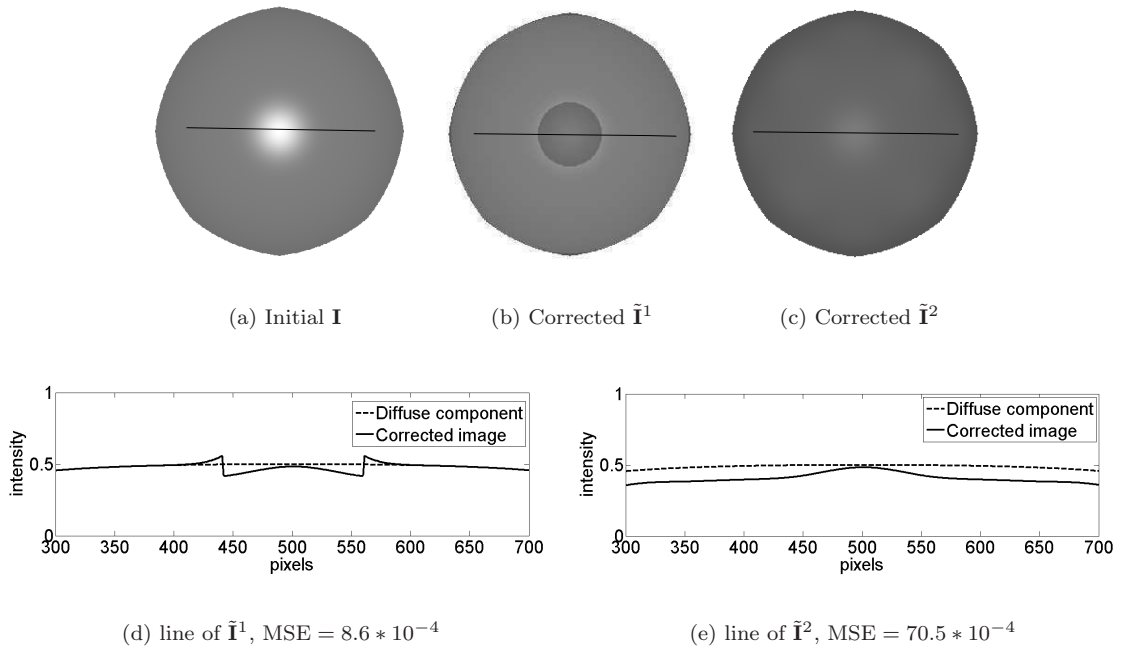


FIGURE 3.1: MSE results for specularity correction.

In Fig. 3.1, we propose to compare different corrected images, image $\tilde{\mathbf{I}}^1$, Fig. 3.1(b), and image $\tilde{\mathbf{I}}^2$, Fig. 3.1(c). The results of correction are well visible on the presented images, and even better on 1D extractions of the central image lines (Fig. 3.1(d) and Fig. 3.1(e)). For the first image we obtain a sharp transition between the corrected and non-corrected zones. The second image contains a general intensity degrading, but there are no sharp transitions. The MSE measure for the image $\tilde{\mathbf{I}}^1$ ($MSE = 8.6 * 10^{-4}$) denotes the better processed image quality, than for the image $\tilde{\mathbf{I}}^2$ ($MSE = 70.5 * 10^{-4}$). From the visual analysis, the quality of the second image is

more acceptable, than of the first one for the photometric stereo further usage. The second image shading is preserved and there is no artifact generation on the borders between corrected and non-corrected zones.

In this work, we will be able to evaluate the quality on the basis of the MSE only for artificially created images, when a reference diffuse component is known. For the real-world data the only means to evaluate the quality is a visual analysis.

In all the following experiments with artificial data, we use the MSE measure in combination with visual analysis of the resulting images. When the images after correction are similar, this measure permits to indicate the image with a better correction.

Nevertheless, for purely numerical results evaluation a specific quality measure should be developed. This point is in perspectives of this work. This measure should include a difference between images (which is already done by the MSE), sharpness of passage between corrected and non-corrected zones, as well as quality of the photometric stereo reconstruction (if such a reconstruction is intended). The measure of specularities correction should be based only on the resulting image without reference diffuse component, so that it can be applied to the processed real-world data.

3.2 Proposed specularity correction algorithm

In this section we propose an original method for specularity correction belonging to the ' $N - to - N$ ' category. In our context, specularity correction is a pre-processing before photometric stereo application, so the initial data for specular zones correction is suitable for the further photometric stereo. Images are taken from a fixed camera position and different illuminations.

If there is a sequence of N images, $\{\mathbf{I}\} = \{\mathbf{I}^1, \mathbf{I}^2, \dots, \mathbf{I}^N\}$, each image of this sequence should be corrected on the basis of references images, $\{\mathbf{I}_{\text{ref}}\}$, which can be the same or different for each \mathbf{I}_i . Images $\{\mathbf{I}_{\text{ref}}\}$ can correspond to all other images of the sequence $\{\mathbf{I}\}$ except the image \mathbf{I}_i to be corrected. At the same time, reference images can be any other images taken from the same camera position, which have no specularities or specularities located differently than in images to be corrected.

Due to the fact that images are taken from the same fixed camera, there is a total correspondence of pixel positions in all images to be corrected and reference images.

Thus, the pixel (x, y) in image \mathbf{I}^1 corresponds to the pixel (x, y) in image \mathbf{I}^2 , etc., corresponds to the pixel (x, y) in all reference images. The term 'pixel' refers to the position in the image.

To summarize, the problem which is addressed in the next sections consists in correction of specular zones in images from the set $\{\mathbf{I}\}$ using the sets of images $\{\mathbf{I}_{ref}\}$. The specular pixels of the input images should be corrected while preserving shading particularities appropriate to each image of $\{\mathbf{I}\}$.

3.2.1 Specularity evaluation

In order to correct specular pixels, we should first evaluate the degree of specular component in these pixels. This can be done on the basis of how the pixel intensity value in \mathbf{I}^i is different comparing to intensity values of reference images $\{\mathbf{I}_{ref}\}$. Due to the fact, that the image to be corrected and the reference images were taken with different illumination, specular zones should not have the same location. To compare pixels on image to be corrected and on reference images, we can use either difference, or ratio operations. We choose to use ratio, because it is easier to estimate specular zone as a percentage of abnormal intensity than as an absolute difference between specular and non-specular values. Indeed, even if the dichromatic model supposes the superposition of these two components, the difference strongly depends on the photographed object own intensity and the light source, and cannot be generalized for all possible cases without knowledge on a reference value for non-specular pixels.

Thus, we use a ratio measure in order to evaluate how much the current pixel (x, y) of the image \mathbf{I}^i is different from the same pixel on each reference image \mathbf{I}_{ref}^m , $m = 1, \dots, M$:

$$\mathbf{R}^m(x, y) = \frac{\mathbf{I}^i(x, y)}{\mathbf{I}_{ref}^m(x, y)}, \quad \forall(x, y) \mathbf{I}_{ref}^m(x, y) \neq 0. \quad (3.2)$$

If pixel intensity values are close on the two images, we obtain a ratio close to 1. Otherwise, if the ratio value is very high, it means that we detect a specular pixel in image \mathbf{I}^i .

If there are M reference images, we aim at combining information on specular pixels from all reference images. Combination (or aggregation) of the obtained ratio values

for \mathbf{I}^i and all reference images can be made as a weighted sum function:

$$\mathbf{W}^i(x, y) = \sum_{m=1}^M c_m \frac{\mathbf{I}^i(x, y)}{\mathbf{I}_{\text{ref}}^m(x, y)}, \quad (3.3)$$

where coefficients c_m denote weights of each reference image, $c_m \in [0, 1]$ and $c_1 + c_2 + \dots + c_M = 1$. M is the number of reference images in the set $\{\mathbf{I}_{\text{ref}}\}$ and N is the number of images to be corrected ($i = 1, \dots, N$) in $\{\mathbf{I}\}$. For a particular case when we use all other images of the set $\{\mathbf{I}\}$ to correct the current one, we have $M = N - 1$, but it is not a mandatory condition.

If we have supplementary information on light sources positions or are sure to have images with only diffuse reflectance, the weighted sum function permits to exploit this information and to give higher weights for images with less specularities, and vice versa to assign zero weights for specular ones. If there is no information on presence of specularities on reference images, which is the most frequent case, we can chose equal weights for all ratio values and realize a simple averaging operation: $c_1 = c_2 = \dots = c_M = \frac{1}{M}$. Another possibility is to define only one reference image as the median of all the initial images of the reference sequence.

As an example, suppose that we have an image \mathbf{I} with a specular zone, Fig. 3.2(a), which should be corrected using three other images generated with different illumination directions, $\mathbf{I}_{\text{ref}}^1$ (Fig. 3.2(b)), $\mathbf{I}_{\text{ref}}^2$ (Fig. 3.2(c)) and $\mathbf{I}_{\text{ref}}^3$ (Fig. 3.2(d)). Let us choose only the image line (in black) and look in details on this line diffuse and specular components, as well as on the ratio between pixels intensities of the image line and the reference images, Fig. 3.3. For this example, we propose to compare aggregation of the three ratios using mean and median operations in Fig. 3.3(h).

It can be seen that between the image \mathbf{I} and one of the reference images $\mathbf{I}_{\text{ref}}^3$ there exists a partial specular zones overlapping, Fig. 3.3(e) and 3.3(g). The overlapping complicates the process of specular zone detection, Fig. 3.3(f). For such a case, the combination of ratio values using averaging is not sufficient for the specular zone detection, Fig. 3.3(h), compared to using the median operation which is more robust to outliers.

But the median has also its own limitations. It can be efficiently used only when at least half of initial images have no specular zones overlapping. Potentially, the both methods of aggregation can be used for the values of ratio. However, if there is no certainty that at least half of images has no specular zones overlapping, it is

preferable to use the averaging operation due to its lower computational complexity comparing to the median.

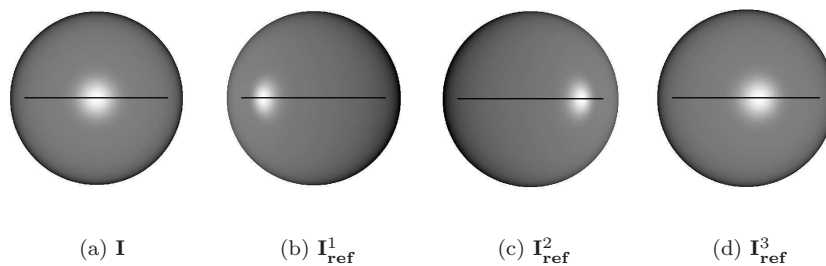


FIGURE 3.2: Example of the specular image to be corrected and reference images.

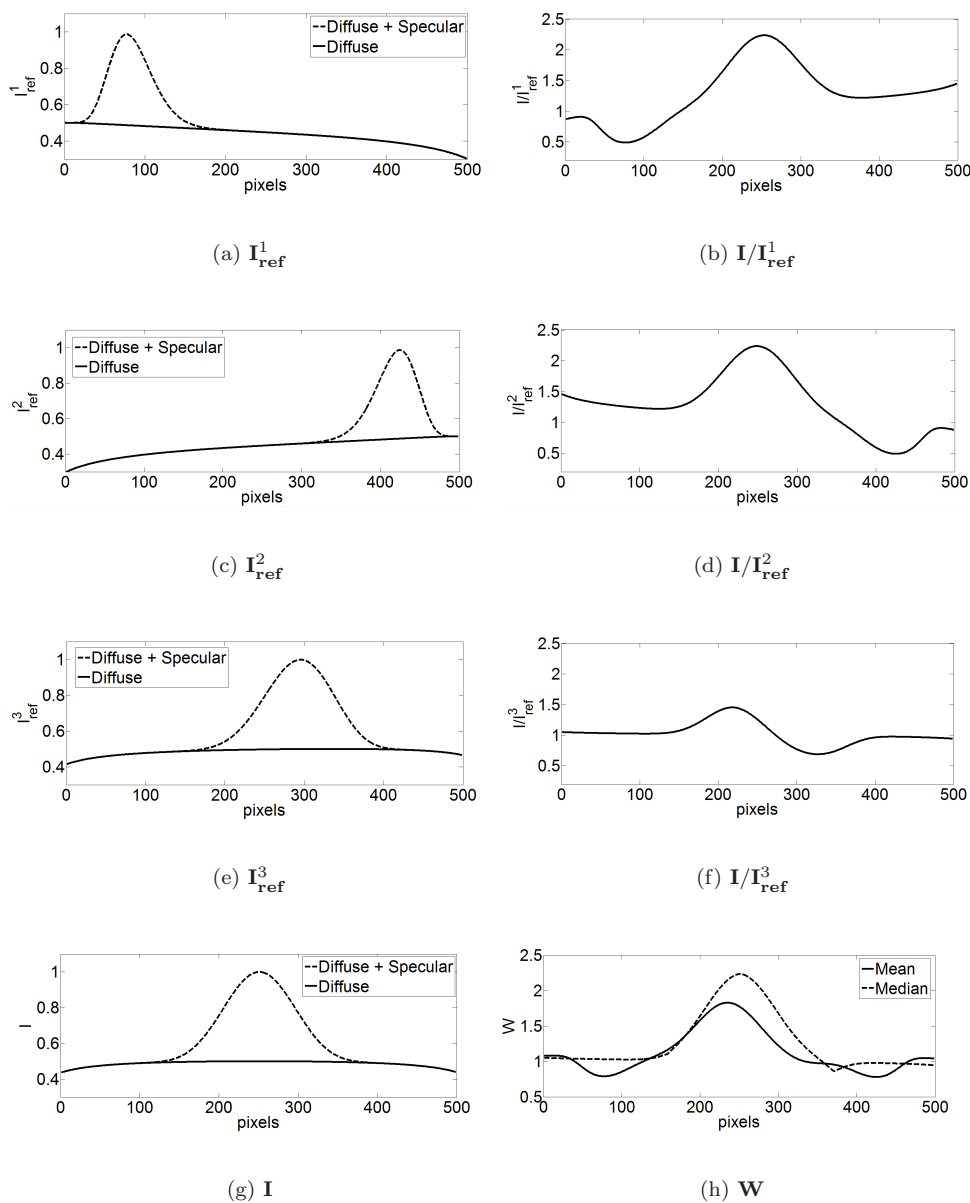


FIGURE 3.3: Ratio aggregation: mean vs median.

3.2.2 Strict specularity correction

The aggregated ratio values \mathbf{W}^i were used in the previous section to characterize the degree of the specularity. Once the specular pixels are characterized, Eq. 3.3, we should correct them to obtain the image $\tilde{\mathbf{I}}^i$. The objective is to correct specularities while still preserving shading particularities of each image.

Whatever the used aggregation function, values of $\mathbf{W}^i > 1$, Eq. 3.3, correspond to specular pixels in image \mathbf{I}^i . This can be defined by the means of a threshold value τ . For $\tau = 1$ we obtain the detection of all pixels which have a higher value than the generalized reference image intensity values. It is not what we aim to do. Then, choosing τ with a certain margin $\tau = 1 + \tau_{margin}$ allows to take into account the corrected image shading particularities and detect only the pixels with abnormally different values, thus, preserving shading which is important for further processing. The value of τ_{margin} depends on the initial images. Indeed, there is a large variation in intensity values for same pixels even for non-specular regions, τ_{margin} should be chosen more elevated.

On the basis of the aggregated ratio values \mathbf{W}^i , a correction can be made as:

$$\forall(x, y), \tilde{\mathbf{I}}^i(x, y) = \begin{cases} \frac{\mathbf{I}^i(x, y)}{\mathbf{W}^i(x, y)}, & \text{if } \mathbf{W}^i(x, y) > \tau, \mathbf{W}^i(x, y) \neq 0, \\ \mathbf{I}^i(x, y), & \text{otherwise.} \end{cases} \quad (3.4)$$

The described specularity correction algorithm is conventionally called strict and is summarized in Algorithm 1.

Now, we will show an example of specularities correction using Algorithm 1. Let us have an artificial image of the specular sphere, Fig. 3.4(c), which is a superposition of the diffuse, Fig. 3.4(a), and specular, Fig. 3.4(b), components. We aim at removing the specular component from the image of Fig. 3.4(c) using the proposed strict correction approach and obtaining the resulting image close to the diffuse image component from Fig. 3.4(a).

Image \mathbf{I} , Fig. 3.4(c), will be corrected using a set of reference images $\{\mathbf{I}_{ref}\}$ containing 7 images taken from the same viewing position and varying illuminations. A superposition of specular components of $\{\mathbf{I}_{ref}\}$ is presented in Fig. 3.5. There are 7 reference images made with light sources uniformly distributed around the sphere.

Algorithm 1 Strict specularity correction algorithm.

```

1: In:
    $\mathbf{I}^i$  - the  $i^{th}$  image of  $\{\mathbf{I}\}$  of size  $[X \times Y]$  to be corrected;
    $\{\mathbf{I}_{ref}\}$  - sequence  $M$  of reference images;
    $\tau$  - thresholding parameter.
2: Out:
    $\tilde{\mathbf{I}}^i$  - image with corrected specular zones.
3: for  $x = 1, \dots, X$  do
4:   for  $y = 1, \dots, Y$  do
5:     for  $m = 1, \dots, M$  do
6:        $\mathbf{R}^m(x, y) \leftarrow \mathbf{I}^i(x, y) / \mathbf{I}_{ref}^m(x, y)$ 
7:     end for
8:      $\mathbf{W}^i(x, y) \leftarrow \text{AggregRatio}(\mathbf{R}(x, y))$ 
9:     if  $\mathbf{W}^i(x, y) > \tau$  then
10:       $\tilde{\mathbf{I}}^i(x, y) \leftarrow \mathbf{I}^i(x, y) / \mathbf{W}^i(x, y)$ 
11:     else
12:       $\tilde{\mathbf{I}}^i(x, y) \leftarrow \mathbf{I}^i(x, y)$ 
13:     end if
14:   end for
15: end for

```

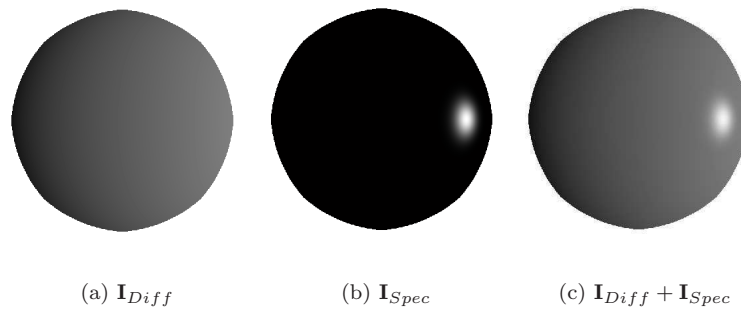


FIGURE 3.4: Diffuse and specular component of the artificial image.

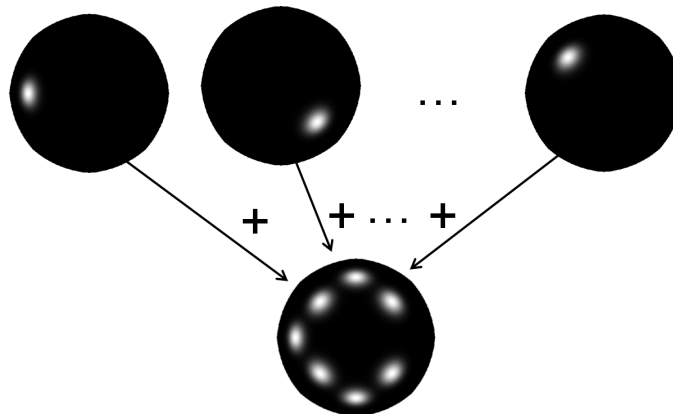


FIGURE 3.5: Superposition of the 7 reference images specular components.

Once the data are defined, we should choose a threshold value τ for the strict correction algorithm. In order to observe this example in more details, we propose to extract an image line (white line in the images of Fig. 3.4) and to define the threshold value for this line. In Fig. 3.6(a) there is the image line to be corrected and the diffuse component which we aim to obtain. In Fig. 3.6(b) we demonstrate the aggregation of the ratio values with mean function, Eq. 3.3. With the chosen threshold value $\tau = 1.2$ we are able to detect the specular zone, but not only this zone is considered as a zone with abnormally elevated intensities. A part of the image with too different shading is also considered as specularity and processed.

The results of the proposed strict correction is shown in Fig. 3.7(c). As it can be seen, the passage between the corrected and the non-corrected zone of the image shading is very sharp and visible on the resulting image $\tilde{\mathbf{I}}$. The main disadvantage of the strict specularity correction is that this method does not take into account the continuous nature of the specular component distribution. Instead of specularities, we get new intensity artifacts. To cope with these artifacts we propose a soft specularity correction algorithm.

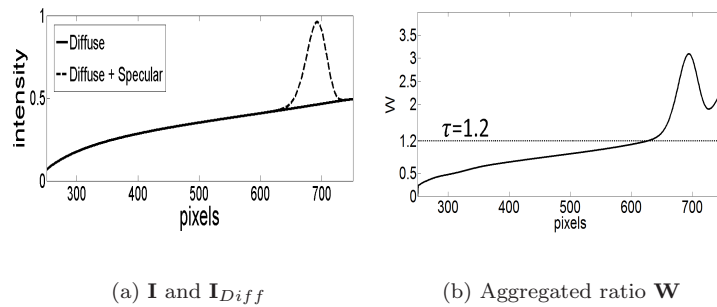


FIGURE 3.6: 1D example of the strict threshold.

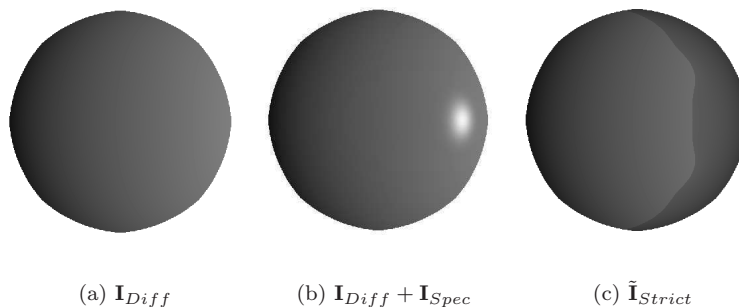


FIGURE 3.7: Example of the strict specular zone correction, $\tau = 1.2$.

3.2.3 Soft specularity correction

The presented strict specularity correction method copes with the task of the specular zone correction. However, the usage of this method leads to intensity artifacts related to the strict passage between corrected and non-corrected zones. The reason for this consists in possible intensity difference between the image to be corrected and reference images.

Indeed, we show in Chapter 2, that specularity boundaries are not sharp. Whatever the distribution of the specular zone intensities is (e.g., Blinn-Phong, Gaussian, Beckmann or Cook-Torrance), it always has a continuous nature, Fig. 2.5. Thus, a correcting function should also be continuous.

The correction of the specular zone should be maximum for the ideal reflecting direction and continuously degrading for other directions of the specular lobe. In our case, general shading preservation is very important, so the correction should be zero or close to zero for all parts of the image except the specular zone.

In order to preserve continuity we propose the soft correction of the image \mathbf{I}^i with the following equation:

$$\forall(x, y), \tilde{\mathbf{I}}^i(x, y) = \frac{\mathbf{I}^i(x, y)}{\mathbf{Q}^i(x, y)}, \mathbf{Q}^i(x, y) \neq 0 \quad (3.5)$$

where $\mathbf{Q}^i(x, y)$ is a correcting continuous function. It can be evaluated as:

$$\forall(x, y), \mathbf{Q}^i(x, y) = \mathbf{W}^i(x, y)^{\mathbf{F}_W^i(x, y)}. \quad (3.6)$$

As a continuous function, we propose to use the sigmoidal function:

$$\forall(x, y), \mathbf{F}_W^i(x, y) = \frac{k}{1 + e^{-\alpha(\mathbf{W}^i(x, y) - \tau)}}, \quad (3.7)$$

where k , α and τ are the sigmoidal function parameters, and $\mathbf{W}^i(x, y)$ denotes aggregated ratio values for the pixel (x, y) , as defined in Eq.3.3.

All the three sigmoid function parameters make sense for the correction: α defines the sharpness between corrected and non-corrected zones, τ defines the threshold for the ratio values as it is done for the strict method, and $k \in [0, 1]$ defines a degree of the specular zones correction. Influence of these parameters on the sigmoid \mathbf{F}_W function is shown in Fig. 3.8(a), 3.8(b) and 3.8(c) respectively.

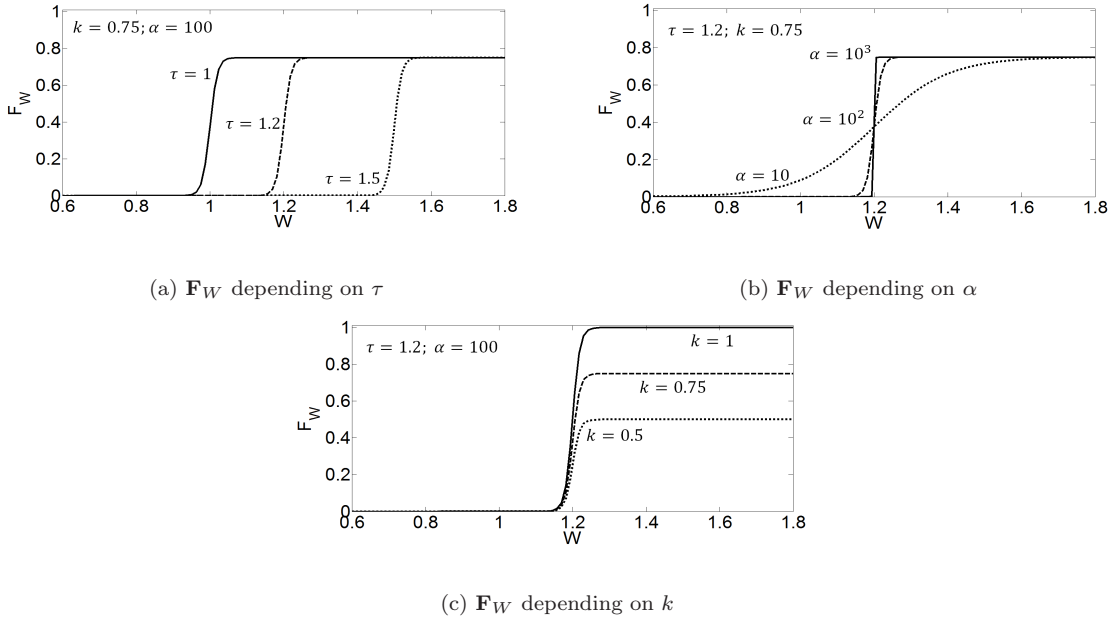


FIGURE 3.8: Sigmoid function F_W values depending on the τ , α and k parameters.

The parameter τ , as in the case of strict correction, permits to define if the pixel value is specular or not. For the sigmoid function, Fig. 3.8(a), it corresponds to the central position of passage between 0 and 1. As for the strict algorithm, this value should be pre-defined. We empirically choose $\tau = 1.2$. This value permits simultaneously to detect abnormal values of intensities and preserve particularities of shading of the image to be corrected.

The value of the parameter α influences the sharpness of passage between the corrected and non-corrected zones. This parameter more than all others takes into account the continuous nature of the specular zone to be corrected. Ideally, the value of α should be proportional to the standard deviation of the specularity intensities distribution. However, this information is never available for real-world tasks, so the value of α should be empirically chosen. In the artificial experiments with large specular zone, we can set a small value of $\alpha = 5$, however for the real-world data the specular zone is not usually large, moreover it is surrounded by other image part with complex texture. Thus, a small value of α for such images can cause a distortion of the adjoined image regions. The value $\alpha = 100$ permits to save the necessary continuity and not to involve the neighboring image pixels into the processing.

The last algorithm parameter k regulates the degree of correction. The value of k belongs to $[0, 1]$, and the maximum correction is achieved for $k = 1$. The maximum

Algorithm 2 Soft specularity correction algorithm.

```

1: In:
    $\mathbf{I}^i$  - the  $i^{\text{th}}$  image of  $\{\mathbf{I}\}$  of size  $[X \times Y]$  to be corrected;
    $\{\mathbf{I}_{\text{ref}}\}$  - sequence of  $M$  reference images;
    $k, \alpha, \tau$  - sigmoid function parameters.
2: Out:
    $\tilde{\mathbf{I}}^i$  - image with corrected specular zones.
3: for  $x = 1, \dots, X$  do
4:   for  $y = 1, \dots, Y$  do
5:     for  $m = 1, \dots, M$  do
6:        $\mathbf{R}^m(x, y) \leftarrow \mathbf{I}^i(x, y) / \mathbf{I}_{\text{ref}}^m(x, y)$ 
7:     end for
8:      $\mathbf{W}^i(x, y) \leftarrow \text{AggregRatio}(\mathbf{R}(x, y))$ 
9:      $\mathbf{F}_W^i(x, y) \leftarrow \frac{k}{1 + e^{-\alpha(\mathbf{W}^i(x, y) - \tau)}}$ 
10:     $\tilde{\mathbf{I}}^i(x, y) \leftarrow \mathbf{I}^i(x, y) / \mathbf{W}^i(x, y)^{\mathbf{F}_W^i(x, y)}$ 
11:   end for
12: end for

```

correction means that the detected specular region is totally replaced by the generalized reference image (evaluated on the basis of the weighted sum for $\mathbf{W}(x, y)$ in Eq. 3.3). The goal for some tasks, including photometric stereo, is not to replace the specular region but to reduce its intensity values while still preserving general shading. Even if the specular zone is not totally removed, its reduced value has not such a negative impact on the reconstructed surface as the initial specularity.

For the aggregated ratio values $\mathbf{W}^i(x, y) \gg \tau$, Eq. 3.3, the function of the soft thresholding $\mathbf{F}_W^i(x, y) \rightarrow k$, and for $\mathbf{W}^i(x, y) \ll \tau$ the function $\mathbf{F}_W^i(x, y) \rightarrow 0$. Thus, for the maximum correction ($k = 1$) and for the limit values of $\mathbf{W}^i(x, y)$ the soft correction tends to the strict one. The main difference is that an additional parameter α permits to regulate the sharpness of the transition of the sigmoid function for the evaluated $\mathbf{W}^i(x, y)$. The larger parameter α is chosen, the closer to the strict 'stair threshold' the function $\mathbf{F}_W^i(x, y)$ is. In other words, for $\frac{1}{\alpha} \rightarrow 0$ (or $\alpha \rightarrow \infty$) we obtain the strict correction.

Depending on the aggregated ratio value and on the values of the function $\mathbf{Q}^i(x, y)$, the correction is minimum for non-specular pixels and is maximum for the pixels which are considered to have abnormally elevated intensities. All the details of the described soft specularity correction algorithm are presented in Algorithm 2.

Let us compare results of the correction by the means of soft and strict thresholding for the same artificial example of the specular sphere, Fig. 3.7. For this example we

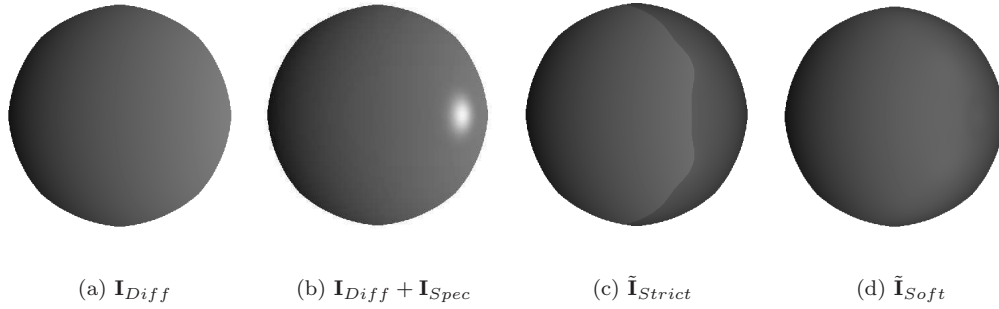


FIGURE 3.9: Example of soft specularity correction with $\tau = 1.2$, $\alpha = 5$, $k = 0.9$.

take the same value of the threshold $\tau = 1.2$. There are also two other parameters which we empirically choose as $\alpha = 5$ and $k = 0.9$. The results of this experiment are shown in Fig. 3.9. Comparing to the results of the strict correction there are no sharp passages. Even if the shading of the sphere is slightly changed from the side of the corrected specular zone, the most corrected pixels are close to the diffuse image component.

For the artificial hemispherical object which we used in the examples of this chapter we will never obtain a zero error between the diffuse component of the image to be corrected and the image without specularities. It is related not only to the imperfection of the proposed algorithm but also to the position of the corrected specular zone on the border of the object and the elevated hemispherical form of this object itself. Illumination from each of the side creates a shadow for the opposite border of the object. These shadows hide the information on the object border, and neither mean nor median operations used for the aggregation of the ratio values, permit to manage this effect.

For the color (RGB) images, each channel is treated independently by the algorithms of correction. Sometimes such an independent processing leads to the color artifacts for the corrected zone of the resulting image for the strict and soft algorithms. In order to avoid this, a multidimensional processing with image channels correlation could be proposed and is in perspectives of this study.

The soft correction permits partially to cope with the main problem of the strict correction, discontinuity between the corrected and non-corrected zones. Now, the main difficulty of the soft correction algorithm consists in parameters tuning. The parameters depend on the input images (e.g., if the image to be corrected and the reference images are close in shading or not) and on the desired correction. The parameter τ is the most sensitive to the input images and cannot be easily defined

a priori. The research on the automatic parameters tuning is also in perspectives of the thesis work.

3.3 Influence of the initial images

As for any other image processing algorithm, the efficiency of specularity correction by the proposed approach depends on the input images. The input images for the specularity correction algorithm can be characterized by the relative positions and the number of the light sources used during initial images acquisition.

3.3.1 Distribution of the illumination sources

The proposed specularity correction algorithms is from the ' $N - to - N$ ' category. Thus, we have N initial images and we aim at correcting specular zones in each of these images while still preserving shading. As for the ' $N - to - 1$ ' methods we exploit the information, hidden by specularities, from other images. However, there is a limitation of multi-images exploitation for specularities correction when there is specular zones overlapping.

The analysis of specular zones overlapping was first presented in [78]. There are three cases of overlapping, illustrated in Fig. 3.10. In the first case, Fig. 3.10(a), each image contains its own distinct region of specularity with no overlapping. For such a case the information about the image region hidden by the specular effect for one image, is present in the other $(N - 1)$ images. In the second case, Fig. 3.10(b), a partial overlapping is observed but the information about the specular zone is present in some of the other images. The last case, Fig. 3.10(c) is the most complex for specularity correction. In the case of total overlapping there are object regions for which there is no information about the surface intensity on any of the images. The ' $N - to - N$ ' method proposed in this work is efficient only for the cases shown in Fig. 3.10(a) and Fig. 3.10(b).

In Table 3.2 we show a comparison of specularity correction by the soft method depending on the spatial distribution of the light sources used for the images creation. In this table diffuse and specular components of the image to be corrected (\mathbf{I}_{Diff} and \mathbf{I}_{Spec}) are presented. For each configuration, we show a superposition of specular

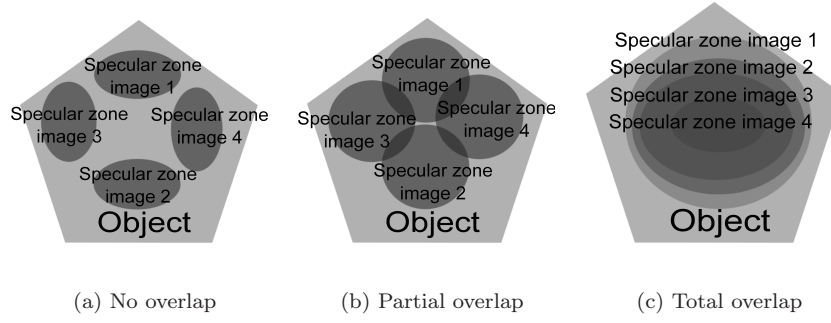


FIGURE 3.10: Specularity overlapping cases.

components of the reference images $\bigcup_{m=1}^M \{\mathbf{I}_{\text{refSpec}}^m\}$ used for the correction (7 reference images), as well as results of the correction by the soft algorithm ($\tau = 1.2$, $\alpha = 5$, $k = 0.9$, ratio values \mathbf{W}^i are obtained by averaging) and 1D extraction of the corrected central line. The results of correction are evaluated using the MSE between the corrected image and the diffuse component of the initial image. The number of reference images M remains the same for all the three examples of correction. From the 1D representation it can be seen, in Configuration 3 due to the overlapping of specular zones for reference images, the specularity is reduced in intensity but not completely treated. The information hidden by the specular zone of \mathbf{I} cannot be extracted from reference images. On the other hand, the best correction was obtained for the randomly distributed specular zones on reference images. For an efficient specularities correction by the proposed approach, there should be no total overlapping of specular regions between the initial and reference images. In order not to obtain specular zones overlapping, light sources used during initial images acquisition should be well distributed in space around the photographed objects.

3.3.2 The number of reference images

In this part we show, that not only the distribution of light sources is important, but also the number of reference images. The more reference images we have, the more information of different object parts without specularities is available. The number of the reference images in the initial sequence can be very different (at least one reference image). By the experiments presented in Table 3.3 it is shown, how the number of light sources influences the resulting specular zone correction.

As for the experiments from the previous section, the first column in Table 3.3 presents a superposition of specular components of the reference images $\bigcup_{m=1}^M \{\mathbf{I}_{\text{refSpec}}^m\}$

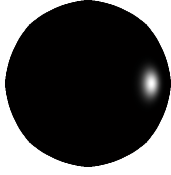
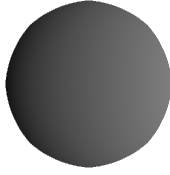
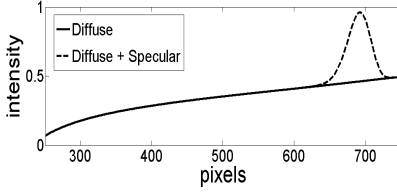
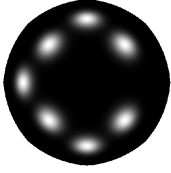
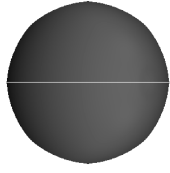
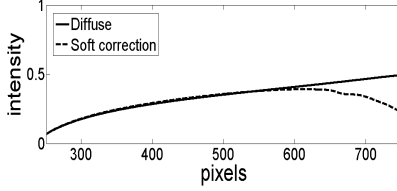
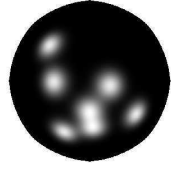

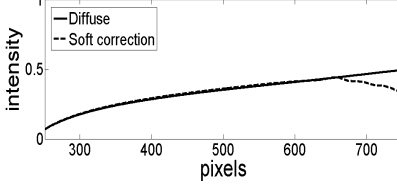
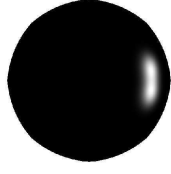
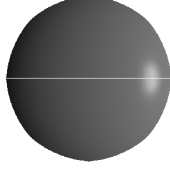
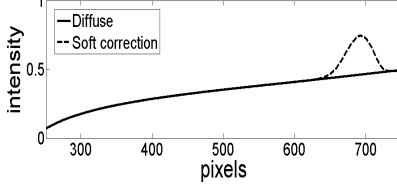
I_{Spec}	I_{Diff}	I line
		 MSE = $142.0 * 10^{-4}$
$\bigcup_{m=1}^M \{I_{\text{refSpec}}^m\}$	\tilde{I}	\tilde{I} line
Configuration 1  $M = 7$		 MSE = $51.7 * 10^{-4}$
Configuration 2  $M = 7$		 MSE = $15.1 * 10^{-4}$
Configuration 3  $M = 7$		 MSE = $52.2 * 10^{-4}$

TABLE 3.2: Specular zones overlapping and correction by the means of soft algorithm ($\tau = 1.2$, $\alpha = 5$, $k = 0.9$).

used for the correction. In the second column we present the results of the soft correction ($\alpha = 5$, $k = 0.9$, $\tau = 1.2$, ratio values \mathbf{W}^i are obtained by averaging), and the last column corresponds to the 1D image line extraction and the MSE to better evaluates results of the realized correction.

The conducted experiments show that the number of reference images is also crucial for the results of correction. For an efficient specularity correction by the proposed methods, it is preferable to have more images with well distributed light sources, so different location of specularities.

Due to Table 3.3 the best results of specularities correction were obtained for the

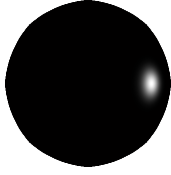
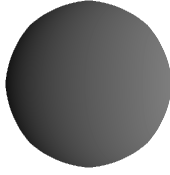
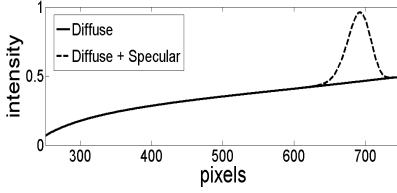
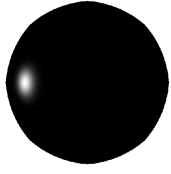
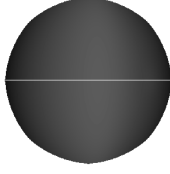
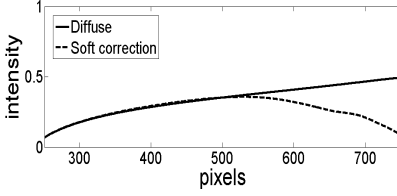
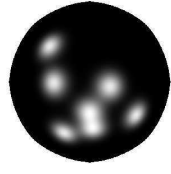
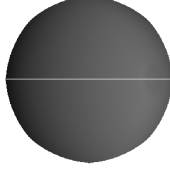
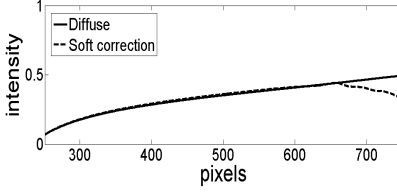
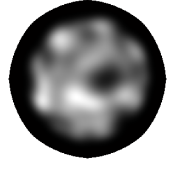
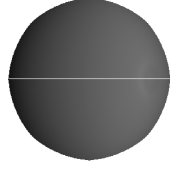
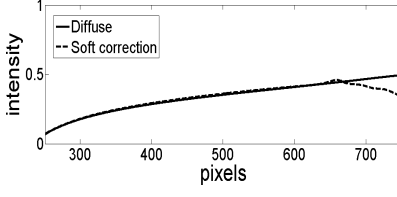
I_{Spec}	I_{Diff}	I line
		 <p>MSE = $142.0 * 10^{-4}$</p>
$\bigcup_{m=1}^M \{I_{\text{refSpec}}^m\}$	\tilde{I}	\tilde{I} line
Configuration 1  $M = 1$		 <p>MSE = $179.7 * 10^{-4}$</p>
Configuration 2  $M = 7$		 <p>MSE = $15.1 * 10^{-4}$</p>
Configuration 3  $M = 100$		 <p>MSE = $11.7 * 10^{-4}$</p>

TABLE 3.3: Specularity soft correction ($\tau = 1.2$, $\alpha = 5$, $k = 0.9$) depending on light sources distribution in space.

case of 100 reference images with randomly distributed light sources. The worst results were shown for the case of only one reference image. Using only one reference image, the specular zone is well corrected, however the shading particularities are also considered as specularities and are not preserved. Satisfactory results have been also obtained for 7 reference images with randomly distributed light sources positions.

Thus, the more initial images made with well spatially distributed light sources are used for the proposed approach of specularities correction, the more information about object region corresponding to specularity we have and the better correction results can be achieved.

In conclusion, we propose an original soft specularity removal method which allows specular zones correction while preserving shading particularities. The transition between the non-corrected and corrected image zones is not present in the resulting image thanks to the continuous correcting function containing three parameters: τ , α and k . We propose to choose τ slightly superior than 1 in order to less a margin to the image proper shading, $\tau = 1.3$. Values of two other parameters allows are not so crucial for the correction, but they regulate the sharpness of transition between corrected and non-corrected zones (α) and degree of correction (k). Thus, it is sufficient to define values of these parameters which provide the correction enough efficient for the further Lambertian reconstruction. It is the case for the chosen values $\alpha = 10$ and $k = 0.9$ used in the experimental part.

3.4 Experimental results for real-world images

In this section we demonstrate results of the proposed specularity correction algorithm application for real-world data from open databases of specular objects.

The disadvantage of specularity correction for real-world images consists in the inability to compare numerically the resulting images due to the absence of the initial specular-free images. There is no information on the diffuse component for image regions hidden by specularities as in the case of artificial images, such as the specular sphere. Thus, only visual evaluation and comparison with other images of the sequence is possible.

To demonstrate the viability of the proposed method we use images provided by the authors of [60]. In order to easily interpret the results of the correction we pass to the gray-scaled version of two sequences of images: images of the specular cat and the specular fish statues.

Specular cat statue

Let us have eight images of the specular cat statue taken from a fixed camera position and varying illumination, Fig. 3.11. We concentrate our attention on the zoomed region of the cat eye, Fig. 3.12. The image to be corrected \mathbf{I} is in Fig. 3.12(a). A set of reference images $\{\mathbf{I}_{\text{ref}}\}$ is presented by the other seven images, Fig. 3.12(b)

- Fig. 3.12(h). We apply the two discussed specularity correction algorithms (strict and soft) for this input data and compare the obtained results, Fig. 3.13.

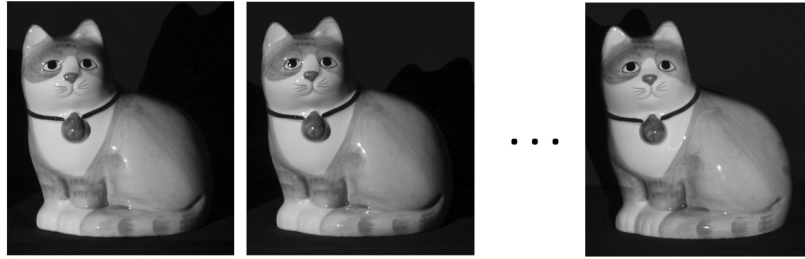


FIGURE 3.11: Example of the specular cat statue images taken from a fixed camera position and varying illumination.

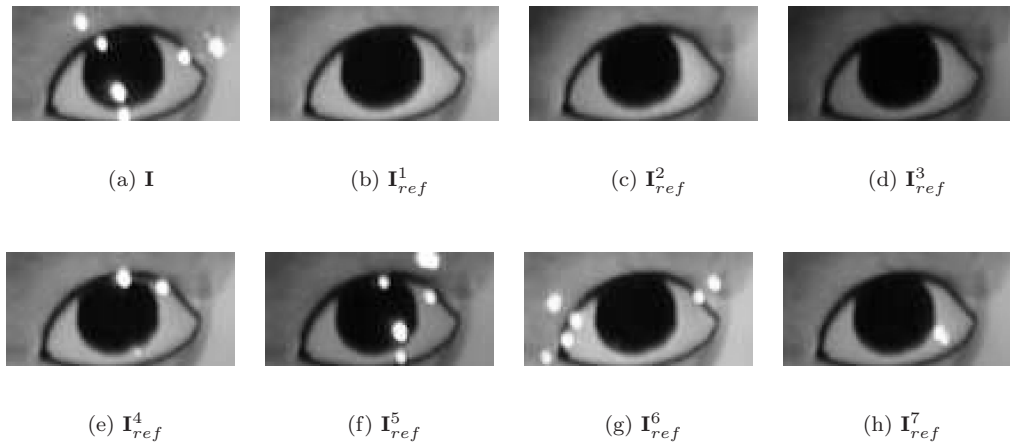


FIGURE 3.12: Initial eight images of the specular cat's eye.

The results of the strict algorithm, Fig. 3.13(b), contain sharp borders between the corrected and non-corrected image regions. The pre-defined threshold value $\tau = 1.3$ permits to detect not only specularities as abnormally high image intensities but also other image regions with a significant difference in shading. Thanks to the soft correction, Fig. 3.13(c), with chosen parameters $\tau = 1.3$, $\alpha = 10$ and $k = 0.9$ the artifacts related to the strict correction are eliminated, and there is no visible passage between the corrected and non-corrected zones. On the images of after strict correction, it can be seen that not only specular zones were detected and corrected, but also a region of shading which is very different on the image to be corrected and the reference images (left upper corner of the image). This is a disadvantage of the pre-defined unique value of τ for all the image pixels. The development of an adaptive thresholding value, which could be automatically evaluated separately for each image pixel, is a perspective of this work.

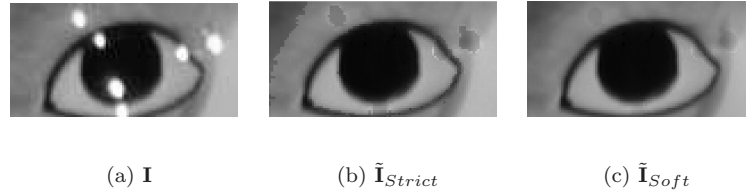


FIGURE 3.13: Example of the specular cat's eye correction with strict ($\tau = 1.3$) and soft ($\tau = 1.3$, $\alpha = 10$, $k = 0.9$) algorithms.

Specular fish statue

The performance of the proposed methods can also be shown for an image with complex texture such as specular fish, Fig. 3.14. The image to be corrected \mathbf{I} is in Fig. 3.14(a), the set of reference images consists of two other gray-scaled images of the fish statue \mathbf{I}_{ref}^1 , Fig. 3.14(b), and \mathbf{I}_{ref}^2 , Fig. 3.14(c). In order to better analyze the obtained results of the correction with strict ($\tau = 1.3$, Fig. 3.14(d)) and soft ($\tau = 1.3$, $\alpha = 10$, $k = 0.9$, Fig. 3.14(e)) algorithms we propose to pass to zoomed images of the texture. In Fig. 3.14(f) there is a texture sample of the initial image affected by specularities. The results of correction for this image region by the means of strict and soft methods are presented in Fig. 3.14(g) and Fig. 3.14(h), respectively. The both algorithms coped with the task of the specular zone correction remaining complex texture of the initial image. However, after strict correction there are still sharp boundaries between corrected and non-corrected zones. This default does not take place after soft correction.

3.5 Conclusions

In this chapter we presented an original approach of specular zones correction which preserves shading for images taken from a fixed camera position and varying illumination directions. Such data is typical for the photometric stereo reconstruction. We suppose that specular zones as abnormally high values of intensities can be detected by the means of a generalized ratio between the image to be corrected and all other images of the initial sequence, and then processed using the proposed correcting function.

In contrast to the existing methods, the proposed approach does not require any prior information on color of the illuminant used during the acquisition. It is from the

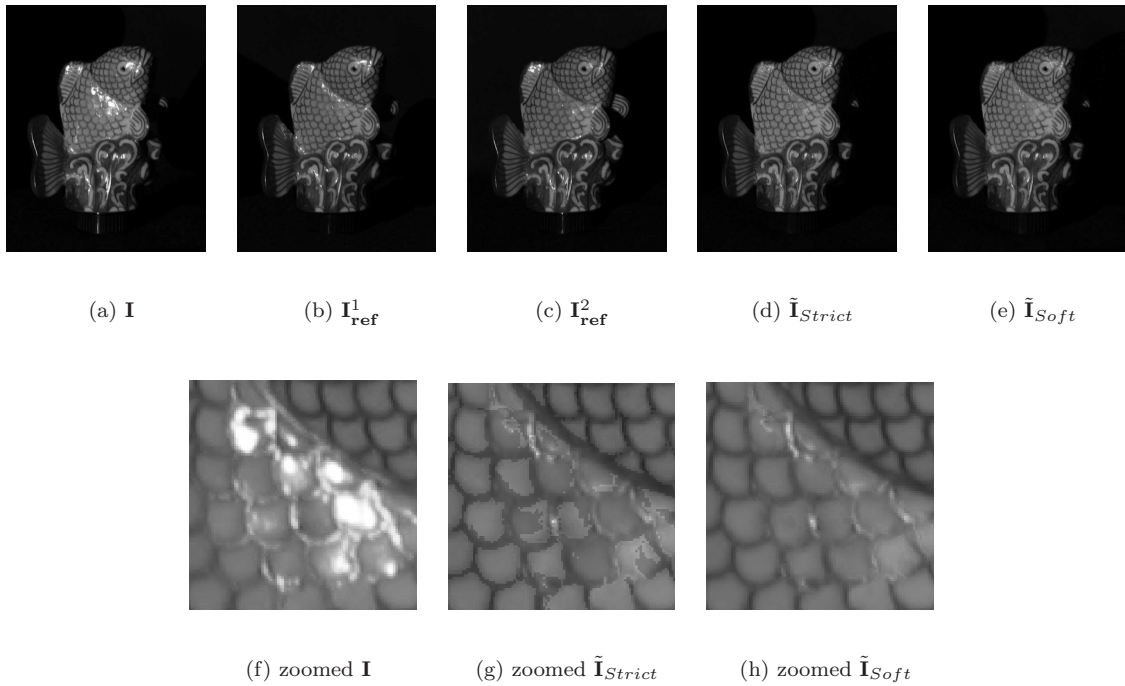


FIGURE 3.14: Example of the specular fish correction with strict ($\tau = 1.3$) and soft ($\tau = 1.3, \alpha = 10, k = 0.9$) algorithms.

category 'N-to-N' methods: on the basis of N input images with specularities we obtain the same number of images, N , without specular effect. The main advantage of this category consists in extraction of the information hidden by specularities from other images of the input sequence under condition of no total overlapping of specular zones for different images. The proposed algorithms can be applied to any set (at least two) of images made from a fixed camera position and different positions of light sources.

The images obtained after the correction meet the stated requirements: the specular zones are well corrected and the shading of image is preserved without additional artifacts on the borders between the corrected and non-corrected zones.

The viability of the proposed method was demonstrated for artificial and real-world images of specular objects.

The images obtained after the proposed correction are considered as images of 'pseudo-Lambertian' surfaces which could be directly used for the algorithm of Lambertian surface reconstruction as it is presented in Chapter 4.

Chapter 4

Uncalibrated photometric stereo

In this chapter we propose an original method of uncalibrated photometric stereo for Lambertian surface reconstruction. Uncalibrated photometric stereo refers to the reconstruction without any prior information on lighting directions. Due to a constantly growing interest to the 3D technologies, fields of photometric stereo application is vast, e.g., 3D face recognition devices used for security systems, [87], portable devices for 3D texture measurements in non-destructive testing, [88], etc. The structure of an industrial acquisition system can be slightly or considerably changed from one version of a device to another, as well as qualitative light sources calibration is not always made sufficiently often. Thus, the solution of the photometric stereo problem without information on illumination conditions can be useful for the variety of industrial tasks.

After the specular removal procedure presented in the previous chapter, the reconstructed surface is supposed to be Lambertian. The Lambertian nature of the surface allows not taking into account the viewer position. Even if there is no more dependency on the viewer position but only on the lighting directions, Lambertian photometric stereo surface reconstruction remains a difficult inverse task. In our approach, the problem of normals finding is presented as a matrix factorization task. The resolution becomes possible thanks to the alternating optimization scheme and implemented value and quality constraints. The presented iterative algorithm is completely 'data-driven'. It permits to separate information relating to the surface (normals and albedo) and to the light sources (illumination directions and intensities) using only initial (at least three) images made with a fixed camera position and varying unknown lighting directions.

4.1 Lambertian photometric stereo reconstruction

Photometric stereo (PS) reconstruction of Lambertian surfaces is a widely known and used technique since 1980s after has been invented and proposed by Woodham in [24]. The photometric stereo approach introduces a dependency between surface normals, illumination directions, and surface albedo values. The resolution of the linear system turns out to be simple when illumination directions are known and only surface properties should be found. If the information on illumination is not available, the reconstruction procedure becomes more complex and requires either supplementary hypotheses, or additional knowledge on acquisition process.

4.1.1 Basic photometric stereo model

The basic photometric stereo model of Woodham, [24], consists in three-component (albedo values, light directions and surface normals) modeling of the image formation process. This model is widely disclosed in the master thesis of Silver, [89], for more information in theoretical foundations we send the reader to this work.

If we consider N images which are presented in form of a vector with $P = XY$ elements in each image, we can refer to a pixel as (i, j) . The value of i corresponds to the number of the image \mathbf{I}^i , and the value of j denotes linear index of the pixel in the image. Due to the photometric stereo concept, an intensity of the image \mathbf{I}^i pixel $m_{i,j}$ is formed as a multiplication of the surface point albedo $a_{i,j}$ and the cosine of the angle between the lighting direction $(s_{x,i}, s_{y,i}, s_{z,i})^T$ and the surface point normal $(n_{x,j}, n_{y,j}, n_{z,j})$ as:

$$m_{i,j} = a_{i,j}(s_{x,i}, s_{y,i}, s_{z,i})^T(n_{x,j}, n_{y,j}, n_{z,j}). \quad (4.1)$$

In this notation we use x , y and z indices as representation of the 3D vector components in Euclidean space. The non-normalized vector of illumination directions $(s_{x,i}, s_{y,i}, s_{z,i})$ is supposed to be known from the acquisition device. The vector $(n_{x,j}, n_{y,j}, n_{z,j})$ is normalized, so $a_{i,j}$ is a norm of this vector, and the number of unknown variables in the equation is equal to three.

Due to the Woodham model, if initially there are at least three input images with different illuminations (which is one of indispensable conditions for photometric stereo techniques), then Eq. 4.1 can be expanded to a system of equations and rewritten

for the j^{th} image pixel as:

$$\begin{cases} m_{1j} = a_j(s_{x1}, s_{y1}, s_{z1})^T(n_{xj}, n_{yj}, n_{zj}), \\ m_{2j} = a_j(s_{x2}, s_{y2}, s_{z2})^T(n_{xj}, n_{yj}, n_{zj}), \\ m_{3j} = a_j(s_{x3}, s_{y3}, s_{z3})^T(n_{xj}, n_{yj}, n_{zj}). \end{cases} \quad (4.2)$$

In this system, the vector of illumination directions (s_{xi}, s_{yi}, s_{zi}) is supposed to be independent of the pixel j , moreover it is supposed to be known a priori.

To generalize the presented system for all pixels of all input images, let us have a set of N images taken from a fixed camera position and varying illumination directions. Let \mathbf{M} be a matrix of images intensities. It is of the size $[N \times P]$, where N is the number of input gray-scaled images, and $P = XY$ denotes the total number of pixels in each initial $[X \times Y]$ image. Thus, each row of the matrix \mathbf{M} represents a vectorized input image \mathbf{I}^i , $i = 1, \dots, N$. If we return to the notation of the previous chapter, and we aim at reconstructing images after specular zones correction, then each row in the matrix \mathbf{M} consists of vectorized corrected images $\tilde{\mathbf{I}}^i$. The general representation for all initial N images and P pixels in the matrix form is the following:

$$\mathbf{M} = \mathbf{A} \odot \mathbf{S}^T \mathbf{N}, \quad (4.3)$$

where \mathbf{M} is the matrix of initial images already described above; \mathbf{A} is the matrix $[N \times P]$ of albedo values, each column of this matrix corresponds to each pixel albedo value repeated N times; operation ' \odot ' signifies the point-to-point (Hadamard) product of matrices; \mathbf{S}^T is the matrix $[N \times 3]$ of illumination directions, each row of this matrix corresponds to each image illumination direction; and finally the matrix \mathbf{N} of size $[3 \times P]$ denotes the matrix of surface normals, each column of this matrix contains 3D normal of the image pixel. The matrix of illumination directions \mathbf{S} contains non-normalized vectors, it can be presented as $\mathbf{S} = \mathbf{V}\mathbf{L}_D$, where \mathbf{V} is $[3 \times N]$ matrix of normalized vectors of light source directions, and \mathbf{L}_D is a diagonal $[N \times N]$ matrix of norms for the vectors of \mathbf{V} . The matrix \mathbf{L}_D can also be seen as a matrix of light source intensities.

Thus, the basic Eq. 4.3 can also be presented in a four-component form:

$$\mathbf{M} = \mathbf{A} \odot (\mathbf{V}\mathbf{L}_D)^T \mathbf{N} = \mathbf{A} \odot \mathbf{L}_D \mathbf{V}^T \mathbf{N}, \quad (4.4)$$

The presented basic model turned out to be rather widespread in the literature.

With small extensions it can be applied to some specific more difficult reconstruction tasks. For example, the authors of [68] extend albedo matrix including information on reflecting power separately for each color channel. In our notation, the matrix model of this photometric stereo approach can be presented as:

$$\mathbf{M} = \begin{bmatrix} \mathbf{u}^T \mathbf{a}_R \\ \mathbf{u}^T \mathbf{a}_G \\ \mathbf{u}^T \mathbf{a}_B \end{bmatrix} \odot (\mathbf{S}^T \mathbf{N}), \quad (4.5)$$

where \mathbf{M} is the matrix of initial images of the size $[3N \times P]$ (first N rows of this matrix correspond to all red channels of the initial data, second N rows are for green channels, and the last N rows are filled by the values of blue channels), \mathbf{a}_R , \mathbf{a}_G and \mathbf{a}_B are vectors $1 \times P$ of each pixel albedo values for RGB channels, \mathbf{u} is a vector of ones, the length of this vector is equal to N , \mathbf{S} is the matrix $[3 \times N]$ of non-normalized light sources directions and \mathbf{N} of size $[3 \times P]$ corresponds to the matrix of the surface normals. The authors of the paper propose to use four input color images in order to treat cases of highlights and shadows as well. The spectral content of the images gives their algorithm additional cues for the detections of specular and shadowed zones. The prior information on lighting directions used during the input images acquisition is indispensable for the mentioned algorithm.

In [90] the authors apply the expectation maximization (EM) principle to the basic PS model in order to obtain a detailed (or dense) photometric stereo exploiting a huge set of initial images in presence of a mirror sphere in the scene to calibrate the model. 'To calibrate the model' in photometric stereo means to calculate lighting directions before surface reconstruction. Usually, for these purposes authors use an object of known geometry and/or known specular properties. However, not all acquisition devices or image scenes permit to insert such a reference object.

The concept of photometric stereo model calibration using reference object was also used by the authors in [60] and [59]. Image pixel intensity in this approach is presented as a superposition of photometric stereo models of several materials. For each reference material there exists information about its shading for different lighting directions and surface normals combinations. The reconstruction by the means of a look-up table permits to process non-Lambertian surfaces. Such a technique turned out to be efficient not only to reconstruct surface of the object but also to realize the object surface materials clustering. The disadvantage of this approach consists in having a base of reference materials close in properties to the material of the

reconstructed surface. Such an algorithm can be applied to a restricted real-world task when we know a priori all possible materials of the reconstructed surface or similar ones.

The authors of [91] propose to make illumination sources approximation using 3D information of the average face model. After such an approximation, the statistical errors of surface normals and lighting directions estimates are used to improve the albedo approximation. The found albedo is exploited to obtain the albedo-free images for the shape recovery. In the article, the resulting shape and albedo are efficiently used to produce 3D views of the object with illuminations different from the initial ones. Such approach is difficultly portable to a general case of reconstruction when no prior information on surface type is available.

From the evaluation point of view, the demonstrated basic photometric stereo model is rather simple. Its main disadvantage consists in requirement of prior knowledge on illumination directions. If this knowledge is not available, the mentioned algorithms aim at information extraction on the basis of a reference object (model calibration) or imposing constraints on the acquisition device.

In our approach we are not able to use a reference object, and we want to be independent from the images acquisition device. The goal is to reconstruct an object surface using the photometric stereo approach without supplementary knowledge neither on lighting directions nor on a capturing device or a reference object.

4.1.2 Photometric stereo via matrix factorization

In order to deal with the lack of prior information on illumination in photometric stereo model, matrix decomposition algorithms have been proposed. Theoretical argument about a photometric stereo reconstruction without prior knowledge on light sources were initialized by the authors of [92]. The basic model was extended by Hayakawa in [93] and presented as a four matrices product:

$$\mathbf{M} = \mathbf{A}_D \mathbf{N}^T \mathbf{V} \mathbf{L}_D, \quad (4.6)$$

where the matrices \mathbf{N} and \mathbf{V} denote surface normals and light sources directions and are remained in the same form as they were presented in Eq. 4.4; the albedo values matrix \mathbf{A}_D is presented as a diagonal matrix of size $[P \times P]$, each line nonzero

element represents a pixel albedo; and the matrix of light sources intensities \mathbf{L}_D have the same definition as in Eq. 4.4.

Albedo \mathbf{A}_D can be determined as the norm of the surface normals vectors in \mathbf{N} , and light source intensities \mathbf{L}_D correspond to the norms of the illumination directions in \mathbf{V} .

Thanks to the separation of the photometric stereo components into two groups (responsible for surface properties and illumination properties) provided by the four-component model, the methods of matrix factorization (decomposition into multipliers) can be used to solve the problem. This approach was applied originally in the article [93], and Eq. 4.6 can be rewritten as:

$$\mathbf{M} = \mathbf{K}\mathbf{S}, \text{ where } \mathbf{K} = \mathbf{A}_D\mathbf{N}^T \text{ and } \mathbf{S} = \mathbf{V}\mathbf{L}_D. \quad (4.7)$$

The magnitudes of the surface vectors in $\mathbf{K} = [\mathbf{k}_1 \dots \mathbf{k}_P]^T$ represents the albedo value for each pixel, and the magnitudes of the light source vectors in $\mathbf{S} = [\mathbf{s}_1 \dots \mathbf{s}_N]$ correspond to the light sources intensities. On the other hand, assuming $P \geq N$, the matrix of vectorized images \mathbf{M} can be decomposed by the means of the Singular Value Decomposition (SVD) into a $[P \times N]$ matrix \mathbf{U} , a $[N \times N]$ matrix $\mathbf{\Sigma}$, and a $[N \times N]$ matrix \mathbf{W} :

$$\mathbf{M} = \mathbf{U}\mathbf{\Sigma}\mathbf{W}^T. \quad (4.8)$$

Due to the properties of the SVD, $\mathbf{U}^T\mathbf{U} = \mathbf{W}^T\mathbf{W}$ are identity matrices. $\mathbf{\Sigma}$ is a non-negative diagonal matrix with sorted diagonal elements $\sigma_1 \geq \sigma_2 \geq \dots \sigma_N \geq 0$.

Initial data \mathbf{M} is always noisy, and the best approximation of \mathbf{M} can be presented as $\tilde{\mathbf{M}} = \tilde{\mathbf{U}}\tilde{\mathbf{\Sigma}}\tilde{\mathbf{W}}^T = \tilde{\mathbf{K}}\tilde{\mathbf{S}}$. In this case, only the most informative parts of the matrices \mathbf{U} and \mathbf{W} are chosen according to the largest singular values in $\mathbf{\Sigma}$. Due to the SVD, the pseudo-surface matrix is $\tilde{\mathbf{K}} = \tilde{\mathbf{U}}(\pm[\tilde{\mathbf{\Sigma}}]^{1/2})$ and the pseudo-light source matrix is $\tilde{\mathbf{S}} = (\pm[\tilde{\mathbf{\Sigma}}]^{1/2})\tilde{\mathbf{W}}$. An ambiguity in sign could be solved by the means of additional constraints if there is an information on surface reflectance of 6 point and if 3 normal vectors of these points do not lie in a plane. After $\tilde{\mathbf{K}}$ and $\tilde{\mathbf{S}}$ are found, they are located in an arbitrary 3D coordinates system. Only knowledge on three surface normals or three light sources directions can be used to align the resulting vectors to the viewer-oriented coordinates system. Thus, the matrix factorization could be an efficient solution for the photometric stereo model without prior information on light sources but it needs additional prior knowledge on several surface points orientation and reflectance.

In the article [94] the authors complement the solution proposed by Hayakawa in [93] replacing the classical SVD by the so-called hole-filling SVD, which permits to take into account the model outliers specific to the image specular and shadowed regions. The detection of the outliers is made on the basis of the graph-cut principle. This procedure permits to apply the proposed algorithm directly to images of non-Lambertian surfaces passing the phase of specularity correction. This algorithm also needs information on several surface points or lighting direction for model components finding.

The algorithm of the photometric stereo surface reconstruction presented in [95] is close in implementation to the basic one of [93]. Due to the concept of the authors, for Lambertian objects general unknown lighting conditions can be presented using low-order spherical harmonics, and thus the surface can be recovered using optimization in low-dimensional space. The order of space is 4D for the 1st order and 9D for the 2nd order approximations. The chosen order defines the size of the matrix Σ in the basic factorization, Eq. 4.8. The surface normals are found up to Lorentz transformations, [96]. The ambiguity of the transformations can be resolved if there is a prior information on the surface normals and albedo values in two points.

If we put ourselves to the conditions of total lack of information on acquisition process and surface properties, the existing algorithms of matrix factorization do not give the problem solution. However the photometric stereo model extension with the matrix of light sources intensities and the idea of matrix factorization are maintained in our approach.

4.1.3 Uncalibrated photometric stereo via alternating optimization

In order to separate information on surface and light sources contained in images with a fixed position of camera and different shading, one can use alternating optimization principle. In this part we present a general principle of alternating optimization and existing approaches of photometric stereo which exploit this principle. Alternating optimization is one of the existing heuristic algorithms which consists in minimization of the objective function successively fixing all but one variable, [97], [98], [99]. This group of algorithms is frequently used for the tasks of blind source separation when one aims at the observation matrix factorization into two composing matrices under certain conditions (e.g., non-negativity), [100].

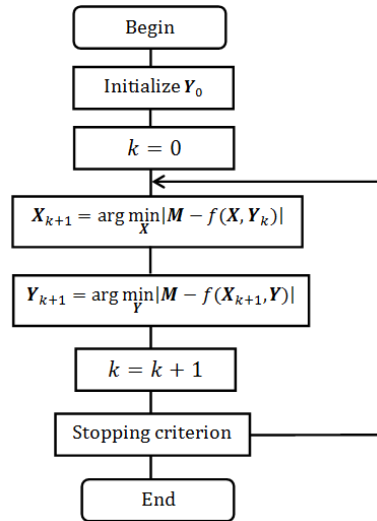


FIGURE 4.1: Alternating optimization general principle.

Alternating optimization general principle

Before explaining application of the alternating optimization principle for the photometric stereo model components estimation, we demonstrate the general principle of this approach. Let us have a matrix of observation \mathbf{M} which can be modeled by a two-variables function $f(\mathbf{X}, \mathbf{Y})$. Alternating optimization algorithm for this simple model components evaluation task is schematically presented in Fig. 4.1. First of all we should initialize values of one of the variables, for example \mathbf{Y}_0 . After that, we search the first approximation of another variable, \mathbf{X}_1 , minimizing the error between the observation matrix \mathbf{M} and the data model with \mathbf{Y}_0 values. Then we suppose variable \mathbf{X}_1 as a known one, and we evaluate the second variable \mathbf{Y}_1 minimizing the error between \mathbf{M} and the data model with \mathbf{X}_1 . These two minimizations are repeated at each iteration until a chosen stopping criterion is reached. The alternating optimization principle is widely used when the global problem solution is not jointly convex but can be decomposed into several consecutive convex sub-minimization tasks, [101].

Quality of the model variables approximation depends on the initialization, chosen stopping criterion and even on the order of model components evaluation.

Alternating optimization for the photometric stereo

Alternating optimization (AO) was already used as an approach for different photometric stereo aspects. Authors of [31] use alternating optimization approach to

estimate reflective parameters of the material for the data with known light sources directions. The known light source directions are also supposed in [102]. In this algorithm, the alternating optimization technique is used as a preliminary step of the general BRDF parameters tuning before shadows detection and surface reconstruction on the basis of the detected shadows boundaries.

For the first time, the iterative solution for a basic Lambertian model with no prior information on lighting was presented in [103]. The author propose to use iterative evaluation to minimize the functional:

$$\forall i, \forall j, \min_{\mathbf{n}_j, \mathbf{s}_i} (m_{i,j} - \mathbf{n}_j \mathbf{s}_i)^2 + \lambda(n_{x,j}^2 + n_{y,j}^2) + \mu_j(\mathbf{n}_j^2 - 1), \quad (4.9)$$

where \mathbf{n}_j corresponds to the j^{th} image pixel normal, \mathbf{s}_i is the i^{th} point light source used for the i^{th} image acquisition, $m_{i,j}$ is the image intensity, λ denotes relative importance of the regularization term and μ_j corresponds to a Lagrangian multiplier. The regularization term $\lambda(n_{x,j}^2 + n_{y,j}^2)$ enforces the surface normal orientation to vertical direction, and the term with a Lagrangian multiplier $\mu_j(\mathbf{n}_j^2 - 1)$ is destined for the normalization of the vector. Iteratively fixing \mathbf{n}_j and \mathbf{s}_i , the authors propose to find the global optimal solution. Adjustment of the solution at each iteration is proportional to the current difference between the predicted and observed intensities. Albedo and light source intensities are not included into the minimization procedure as independent components, moreover, the authors do not discuss any constraints for the light source vectors \mathbf{s}_i .

Our solution is close to [103], but we present differently minimization alternating stages, expanding the global solution to the finding four independent components and including supplementary constraints for each of these components.

For the matrix factorization problem we could use one of other existing algorithms (e.g., gradient descent method or projected gradient method), but in this work we chose to study in details alternating optimization principle in application to the photometric stereo.

4.1.4 Bas-relief ambiguity

Bas-relief ambiguity general principle

One of the existing problem in Lambertian photometric stereo is the bas-relief ambiguity, [26]. When an object is viewed orthographically, its visible surface can be reconstructed only up to a generalized bas-relief transformation. It means that for each image of the object illuminated by a light source there exist the same image for a transformed object taken with a transformed position of the light source.

In our following description we consider the depth map of an object surface presented by $\mathbf{z}(x, y)$ in 3D space. If we have the surface \mathbf{z} , the transformed surface $\check{\mathbf{z}}$ is obtained by applying a scaling and a translation to each pixel (x, y) such as:

$$\check{\mathbf{z}}(x, y) = \lambda \mathbf{z}(x, y) + \mu x + \nu y. \quad (4.10)$$

The bas-relief transformation can be written in matrix form:

$$G = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \mu & \nu & \lambda \end{bmatrix}. \quad (4.11)$$

For the reconstruction process it means that for any evaluated normal vector $\mathbf{n}(x, y)$ there exist the vector $\check{\mathbf{n}}(x, y)$ under a linear transformation

$$\check{\mathbf{n}}(x, y) = G^{-T} \mathbf{n}(x, y), \quad (4.12)$$

that produces the same image intensity with the transformed lighting direction vector $\check{\mathbf{v}}$:

$$\check{\mathbf{v}} = G^{-T} \mathbf{v}. \quad (4.13)$$

Thus, for a single viewpoint and unknown light sources directions, the surface relief can be reconstructed up to the three-parameter bas-relief transformation.

One of the solutions for this problem was recently proposed by authors in [28]. The proposed approach consists in establishing the integrability constraint for the objects under perspective projection. According to the hypothesis on the perspective projection, the authors prove the uniqueness of the reconstructive surface without a three-parameter ambiguity. The method is applicable to uncalibrated Lambertian photometric stereo without a priori on the reconstructed surface properties but it

needs the focal length of the camera in order to provide a transformation from the world to the image coordinates.

Certain constraints on lighting configurations can also be proposed as a solution for the photometric stereo, making the problem solution strongly dependent on the capturing device but resolving the problem of ambiguity. In the article [104] the authors propose a stratified algorithm for a so-called ring-light photometric stereo. The images used for the reconstruction are made by the device with five lights located on a view centered cone. Such a fixed and known lighting configuration reduces the general linear ambiguity. Used in combination with a multi-view stereo, the proposed approach is efficiently applied for the basic three-component photometric stereo.

On the other hand, for certain cases the bas-relief ambiguity can be not critical for the further surface analysis. If we are not specially interested in absolute object height or proportion of elevations for different object parts, which are potentially affected by the bas-relief transformation, the reconstruction obtained up to the three-component transformation may be already sufficient for the object analysis.

Depth map evaluation in presence of bas-relief ambiguity

It is important to note that due to the bas-relief ambiguity, the surface is reconstructed with a certain linear transformation, Eq. 4.10. Thus, before comparing the resulting depth map and the ground truth, the coefficients of scaling and translation should be found as a solution of the problem:

$$\operatorname{argmin}_{\lambda, \mu, \nu} \|\mathbf{z}(x, y) - (\lambda \hat{\mathbf{z}}(x, y) + \mu x + \nu y)\|_2^2. \quad (4.14)$$

The optimal coefficients λ , μ and ν for the surface transformation can be found using LSM. After the resulting surface $\hat{\mathbf{z}}$ is transformed into $\check{\mathbf{z}}$ with the optimal coefficients, it can be compared with the ground-truth surface \mathbf{z} by the means of any existing error measure.

One of the existing measures to estimate the quality of the reconstructed depth map, comparing to the ground truth, is the root-mean-square error:

$$\mathbf{z}_{RMSE} = \frac{1}{P} \sqrt{\sum_{x=1}^X \sum_{y=1}^Y (\check{\mathbf{z}}(x, y) - \mathbf{z}(x, y))^2}, \quad (4.15)$$

where $\mathbf{z}(x, y)$ and $\check{\mathbf{z}}(x, y)$ denote the ground-truth and the evaluated depth map after linear transformation for each pixel, P is the total number of pixels. This measure is not always applicable for reconstruction algorithm evaluation. First of all, the ground-truth surface is not always available, especially for real-world data. Secondly, \mathbf{z}_{RMSE} serves for general surface form evaluation, it does not evaluate small surface details which could also be important for certain surface analysis tasks. In order to evaluate surface details, a local measurement should be additionally proposed.

4.1.5 Depth map reconstruction from normals

Photometric stereo model resolution gives us a unit normal (surface orientation) for each pixel. In order to obtain a depth map from these normals one needs to apply an integration algorithm. Even if the surface raw normals could be directly used for surface recognition, the analysis of the depth map is easier for the human visual perception. In our experiments we are going to use both representations: the normal map and the depth map. For this 'normals-to-depth map' passage we tested several existing algorithms, that we present in details in this section. We also propose an implementation of the chosen algorithm which allows making it rather rapid and applicable for real-world data.

Coleman and Jain algorithm

Along with the photometric stereo algorithm for non-Lambertian surface reconstruction, the authors of [67] propose a propagation algorithm for the depth map reconstruction on the basis of the obtained normals map. One of the depth point (usually the central one) is supposed to be known or just set to some initial value. Finally, this value is not of a large importance, its changing leads only to the surface translation along axis Z . Next values of depth are evaluated at horizontal and vertical directions from this central point. After that, the values of the depth are evaluated for the remained quadrant points.

The authors of [67] propose to consider a bias between two neighbor points of the depth map as a tangent line between these points. The bias can be horizontal (changes of n_x value) or vertical (changes of n_y value). Let us investigate in details evaluation of tangent line for a patch given in Fig. 4.2. Tangent line between the

depth map points $\mathbf{z}(0, 0)$ and $\mathbf{z}(1, 0)$ can be presented by the following equation:

$$a_x x + b_x (\mathbf{z}(1, 0) - \mathbf{z}(0, 0)) = 0, \quad (4.16)$$

where $a_x = (n_{x0} + n_{x1})/2$ and $b_x = (n_{z0} + n_{z1})/2$. If the value of the point $z(0, 0)$ is supposed to be known, then:

$$\mathbf{z}(1, 0) = z(0, 0) - a_x/b_x. \quad (4.17)$$

On the basis of the same principle, vertical bias of surface is:

$$\mathbf{z}(0, 1) = \mathbf{z}(0, 0) - a_y/b_y, \quad (4.18)$$

where $a_y = (n_{y0} + n_{y2})/2$ and $b_y = (n_{z0} + n_{z2})/2$.

Diagonal bias $\mathbf{z}(1, 1)$ is proposed to be evaluated as average value between horizontal bias from $\mathbf{z}(0, 1)$ to $\mathbf{z}(1, 1)$ and vertical bias from $\mathbf{z}(1, 0)$ to $\mathbf{z}(1, 1)$.

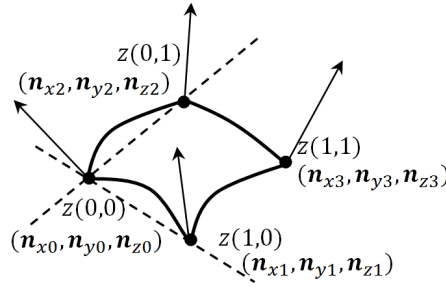


FIGURE 4.2: A patch of depth map with appropriate values of surface normals.

This method of depth evaluation is local, and with propagation of surface gradient values (normals) we obtain propagation of normals error to the neighbor pixels as well.

Frankot and Chellappa algorithm

The algorithm of Frankot & Chellappa, [105], is based on surface integrability constraint representation in terms of the Fourier transform coefficients. For the gradient surface space, from the definition of gradient we can write the following equations:

$$\begin{cases} p(x, y) = -\frac{n_x}{n_z} = \frac{\partial z}{\partial x}, \\ q(x, y) = -\frac{n_y}{n_z} = \frac{\partial z}{\partial y}. \end{cases} \quad (4.19)$$

Surface $z(x, y)$ can be evaluated from normals (n_x, n_y) by the means of integration. However integrating along different directions can provide different results. In order to avoid such a discrepancy, the authors of [105] propose to integrate supplementary integrability constraint which can be formally presented as:

$$\frac{\partial p(x, y)}{\partial y} = \frac{\partial q(x, y)}{\partial x} \Leftrightarrow \frac{\partial^2 \mathbf{z}(x, y)}{\partial x \partial y} = \frac{\partial^2 \mathbf{z}(x, y)}{\partial y \partial x}. \quad (4.20)$$

This constraint can be rewritten as a minimization problem:

$$\min_{\hat{p}(x, y), \hat{q}(x, y)} \int \int |p(x, y) - \hat{p}(x, y)|^2 + |q(x, y) - \hat{q}(x, y)|^2 dx dy. \quad (4.21)$$

Let us present surface $z(x, y)$ by functions $\phi(x, y, \omega)$ as:

$$\mathbf{z}(x, y) = \sum_{\omega \in \Omega} C(\omega) \phi(x, y, \omega), \quad (4.22)$$

where ω is two-dimensional indices from a finite set Ω . The coefficients $C(\omega)$ which minimize Eq. 4.21 can be given by:

$$C(\omega) = \frac{P_x(\omega)C_1(\omega) + P_y(\omega)C_2(\omega)}{P_x(\omega) + P_y(\omega)}, \quad (4.23)$$

where $P_x(\omega) = \int \int |\phi_x(x, y, \omega)|^2$ and $P_y(\omega) = \int \int |\phi_y(x, y, \omega)|^2$. The derivatives of the Fourier basis functions can be written as:

$$\begin{cases} \phi_x = j\omega_x \phi, \\ \phi_y = j\omega_y \phi, \end{cases} \quad (4.24)$$

and $P_x \propto \omega_x^2$, $P_y \propto \omega_y^2$, also it can be written $C_1(\omega) = C_x(\omega)/j\omega_x$ and $C_2(\omega) = C_y(\omega)/j\omega_y$. If $z(x, y)$ is expanded by the means of the Fourier basis functions, then Eq. 4.21 is minimized when:

$$C(\omega) = \frac{-j\omega_x C_x(\omega) - j\omega_y C_y(\omega)}{\omega_x^2 + \omega_y^2}, \quad (4.25)$$

where $C_x(\omega)$ and $C_y(\omega)$ are the Fourier coefficients of the reconstructed surface slope. If we present this relation as:

$$\begin{cases} C_x(\omega) = FFT(p(x, y)), \\ C_y(\omega) = FFT(q(x, y)), \end{cases} \quad (4.26)$$

then the integration results and the resulting depth map can be presented as the inverse Fourier transformation:

$$\mathbf{z}(x, y) = IFFT(C(\omega)). \quad (4.27)$$

This method has a strong theoretical approach, however the reconstruction algorithm is sensible to noise in normals and to sharp passage between neighbor normals directions. Despite these observations, Frankot & Chellappa algorithm is widely used in bibliographical sources as a final stage of the photometric stereo.

Kovesi algorithm

The authors of [106] propose to use shapelets (scaled basis function) and correlation between their gradient and the reconstructed surface gradient to obtain the final depth map.

Surface can be presented not only in Cartesian but also in spherical coordinates. This representation of surface is used in shapelet theory. Let the north pole of a sphere be in the $-\mathbf{z}(x, y)$ direction. Then the latitude of the sphere is called slant angle, σ , and the longitude of the sphere corresponds to the tilt angle, τ . Usually, for the surface representation it is supposed that $\sigma \in (0, \pi/2)$ and $\tau \in (0, 2\pi)$. The depth gradient

$$\nabla \mathbf{z}(x, y) = \left(\frac{\partial \mathbf{z}(x, y)}{\partial x}, \frac{\partial \mathbf{z}(x, y)}{\partial y} \right) = (p(x, y), q(x, y)) \quad (4.28)$$

in terms of the slant and tilt angles is presented as:

$$\nabla \mathbf{z}(x, y) = |\nabla \mathbf{z}(x, y)| (\cos(\tau), \sin(\tau)), \quad (4.29)$$

where $|\nabla \mathbf{z}(x, y)| = \sqrt{p^2 + q^2} = \tan(\sigma)$, and the depth gradient becomes:

$$\nabla \mathbf{z}(x, y) = \tan(\sigma) (\cos(\tau), \sin(\tau)) \quad (4.30)$$

The gradient discrete correlation between the surface and the shapelet is:

$$C_{\nabla} = \sum_x \sum_y |\nabla_z| |\nabla_s|, \quad (4.31)$$

where ∇_z and ∇_s are the gradient of the surface depth and the shapelet. This formulation of correlation does not include dependency on the tilt angle. In order to introduce this information, cosine of the difference between the tilt of the reconstructed surface, τ_z , and the tilt of the shapelet, τ_s , can be used:

$$\cos(\tau_z - \tau_s) = \cos(\tau_z)\cos(\tau_s) + \sin(\tau_z)\sin(\tau_s). \quad (4.32)$$

Thus the tilt alone correlation can be written as:

$$C_\tau = \sum_x \sum_y \cos(\tau_z)\cos(\tau_s) + \sum_x \sum_y \sin(\tau_z)\sin(\tau_s). \quad (4.33)$$

The final correlation, \mathbf{C} , can be obtained as a multiplication of the gradient correlation, C_∇ , and the tilt correlation, C_τ :

$$\mathbf{C} = C_\nabla C_\tau = \sum_x \sum_y |\nabla_z| |\nabla_s| \left(\sum_x \sum_y \cos(\tau_z)\cos(\tau_s) + \sum_x \sum_y \sin(\tau_z)\sin(\tau_s) \right). \quad (4.34)$$

This correlation is evaluated for multiple scales of the shapelet function. The correlation results summation produce the reconstructed surface depth map.

As shapelet functions, the author of [106] proposes to use Butterworth shapelet and 'rounded' 2D Haar wavelet which show better results comparing to the cosine Gabor wavelet usage.

The author states that this algorithm is very robust to noise in normals. Results of reconstruction obviously depend on amount of used scaled shapelets. There is no universal approach for the shapelet functions choice depending on properties of the reconstructed surface.

Basri, Jacobs and Kemelmacher algorithm

We consider the approach described in [95] as the most convenient algorithm for a passage between the surface normals to the depth map. The problem is presented as a least-square means task with fast and direct solution. The surface gradient values can be written by the means of the obtained surface as:

$$\begin{cases} p(x, y) = -\frac{n_x(x, y)}{n_z(x, y)}, \\ q(x, y) = -\frac{n_y(x, y)}{n_z(x, y)}. \end{cases} \quad (4.35)$$

On the other hand, the surface gradient can be presented as a discrete partial derivative of the depth map $\mathbf{z}(x, y)$ as:

$$\begin{cases} p(x, y) = \mathbf{z}(x + 1, y) - \mathbf{z}(x, y), \\ q(x, y) = \mathbf{z}(x, y + 1) - \mathbf{z}(x, y). \end{cases} \quad (4.36)$$

It means that for each depth map pixel (x, y) the following system of equations can be built combining the first and the second surface gradient representations:

$$\begin{cases} n_z(x, y)\mathbf{z}(x + 1, y) - n_z(x, y)\mathbf{z}(x, y) = -n_x(x, y), \\ n_z(x, y)\mathbf{z}(x, y + 1) - n_z(x, y)\mathbf{z}(x, y) = -n_y(x, y). \end{cases} \quad (4.37)$$

We propose to increase connectivity of the pixel gradient proposed in [95] by adding pixels $(x - 1, y)$ and $(x, y - 1)$ in the system of equations, which involves a symmetry into the final solution, so for each depth map pixel it can be written as:

$$\begin{cases} n_z(x, y)\mathbf{z}(x + 1, y) - n_z(x, y)\mathbf{z}(x, y) = -n_x(x, y), \\ n_z(x, y)\mathbf{z}(x, y + 1) - n_z(x, y)\mathbf{z}(x, y) = -n_y(x, y), \\ n_z(x, y)\mathbf{z}(x - 1, y) - n_z(x, y)\mathbf{z}(x, y) = n_x(x, y), \\ n_z(x, y)\mathbf{z}(x, y - 1) - n_z(x, y)\mathbf{z}(x, y) = n_y(x, y). \end{cases} \quad (4.38)$$

The number of pixels is P , then this global system consists of $4P$ equations. In order to find the depth map $\mathbf{z}(x, y)$ we should resolve this overdetermined linear system of equations. It can be done, for example, by the least squares method (LSM), [25]. Our implementation of this algorithm is given in Appendix A.

There are also other less spread surface from normals reconstructing algorithms which could be found in literature (e.g., [107], [108] or [109]).

Comparison of depth map reconstruction algorithms

In order to choose which of the algorithms can cope the best with a sophisticated normal map, we present an experiment of the depth map from normals reconstruction for the artificial surface presented in Fig. 4.3. The ground-truth surface is demonstrated in Fig. 4.3(a). First, we extract the surface normals, and after we integrate the obtained normals map to the depth map using the four described approaches. The depth maps are presented in false colors according to the color-map in Fig. 4.3.

The example surface has smooth and sharp transitions, which permits to better evaluate the different algorithms of the depth map estimation. Visual comparison of the obtained surface with the ground-truth (GT) as well as a root-mean-square errors (\mathbf{z}_{RMSE}) allow judging the quality of the obtained results. It is worth noting, that the algorithms which use a decomposition into basis functions, Frankot and Chellappa (FC) in Fig. 4.3(c) and Kovesi (Kov) in Fig. 4.3(d), do not cope with a non-uniform background. The best results are obtained for the algorithm of Coleman and Jain (CJ), Fig. 4.3(b), and Basri, Jacobs, Kemelmacher (BJK), Fig. 4.3(e), algorithms. Both of them can be successfully used for the experiment results demonstration.

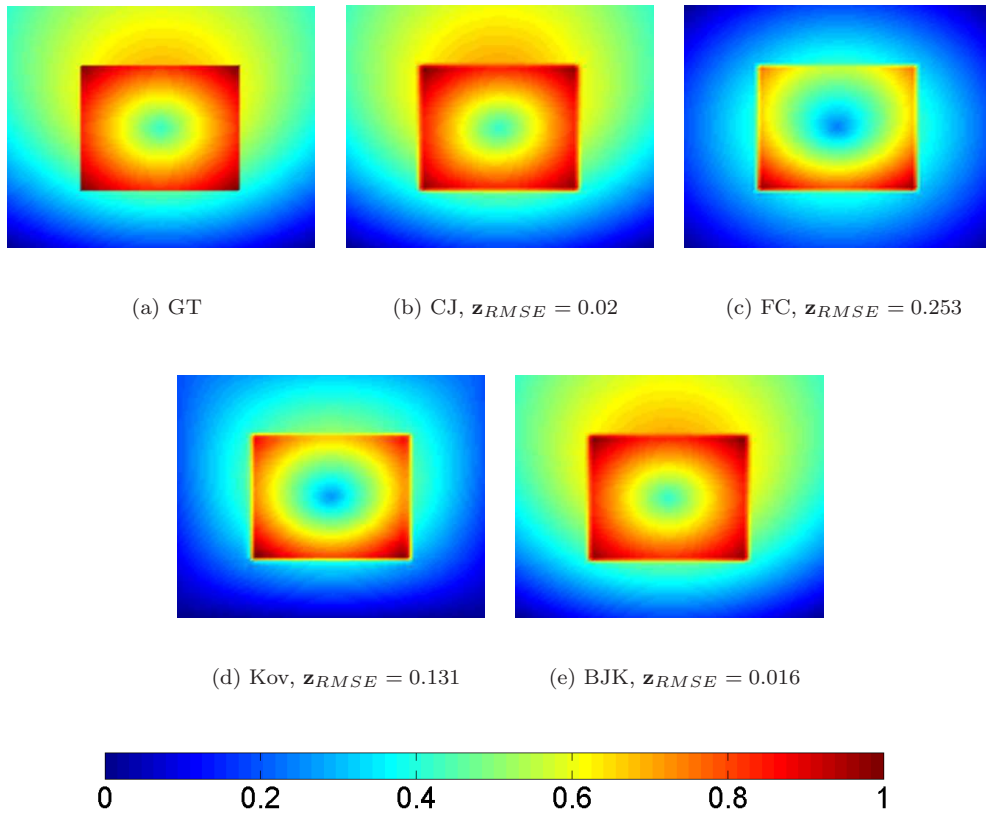


FIGURE 4.3: Comparison of depth map reconstruction algorithms.

We select the algorithm of Basri, Jacobs and Kemelmacher according to the smallest error obtained in this experiment of artificial surface reconstruction. In all the following experimentations we use this algorithm, [95], with our implementation presented in Appendix A.

4.2 Proposed method of uncalibrated PS using alternating optimization

In this section we present the proposed photometric stereo method for Lambertian surfaces. On the basis of images, this method permits to separate information related to surface and illumination into different matrices using alternating optimization principle and without any prior information neither on light sources nor on the reconstructed surface.

4.2.1 General model description

As it has already been presented in the previous sections, let us have N color RGB images of an object made from a fixed camera position. Each $[X \times Y]$ image consists of $P = XY$ pixels. These images are vectorized to the matrix \mathbf{M} of size $[3N \times P]$. Then the proposed model can be presented as:

$$\mathbf{M} = \mathbf{A} \odot \mathbf{L} \odot (\mathbf{V}^T \mathbf{N}), \quad (4.39)$$

where ' \odot ' is the Hadamard product; \mathbf{A} is the matrix $[3N \times P]$ of albedo values; \mathbf{L} is the matrix $[3N \times P]$ of light sources intensities values; \mathbf{V} is the matrix $[3 \times 3N]$ of light sources positions used during acquisition of each initial image; \mathbf{N} is the matrix $[3 \times P]$ of each surface point (pixel) normals. Schematically this representation with all matrices sizes is given in Fig. 4.4. We supposed to have initial color images but each color channel is captured independently. For our multiplicative model it means that a pixel in different channels of the same image has a unique albedo value, but the same surface orientation and lighting direction.

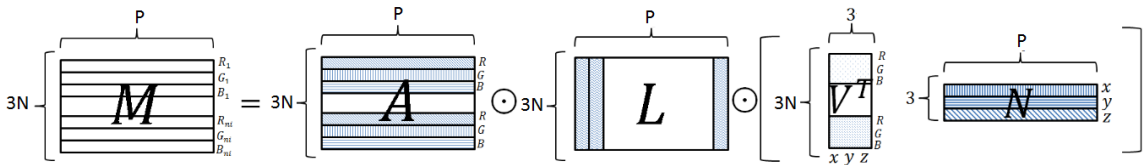


FIGURE 4.4: The proposed PS model matrix representation.

In the presented model Eq. 4.39 there is a lot of information which is repeated in matrices \mathbf{A} , \mathbf{L} , and \mathbf{V} because of the physics of the image formation process. Albedo value is a property of surface point (presented by the images pixel) which depends on

the image color channel but is the same for the same color channel of different images. It can be interpreted as $\mathbf{A} = \mathbf{W}_A \tilde{\mathbf{A}}$, where $\tilde{\mathbf{A}}$ is a $[3 \times P]$ matrix of albedo values different for each image channel and each pixel, and \mathbf{W}_A is an auxiliary structuring matrix. If \mathbf{I}_3 is the $[3 \times 3]$ identity matrix, $\mathbf{W}_A = [\mathbf{I}_3 \mathbf{I}_3 \dots \mathbf{I}_3]^T$, where \mathbf{I}_3 is repeated N times.

The matrix \mathbf{L} should be composed of a repeated column of each channel light sources intensities for all image pixels. To implement such structure light sources intensities matrix is presented as $\mathbf{L} = \tilde{\mathbf{L}} \mathbf{W}_L$, where $\tilde{\mathbf{L}}$ is of the size $[3N \times 1]$ and \mathbf{W}_L is an auxiliary vector of ones with the length equal to P .

It is supposed that each initial image was taken with a single light source, or with several light sources which can be roughly presented by a main illumination direction with three intensities (separately for R, G and B channels). The matrix of light sources positions is presented as $\mathbf{V}^T = \mathbf{W}_V \tilde{\mathbf{V}}^T$, where $\tilde{\mathbf{V}}$ is the $[3 \times N]$ matrix of lighting directions and \mathbf{W}_V is an auxiliary structuring matrix, $\mathbf{W}_V = [\mathbf{D}_1 \mathbf{D}_2 \dots \mathbf{D}_i]^T$, where \mathbf{D}_i is $[N \times 3]$ matrix with i^{th} row of ones and all other elements equal to zero.

The matrix of normals \mathbf{N} is of size $[3 \times P]$ and presents each pixel normals in 3D space. It contains no intentional repetitions.

Taking into account all these explanations, the model can be rewritten as:

$$\mathbf{M} = (\mathbf{W}_A \tilde{\mathbf{A}}) \odot (\tilde{\mathbf{L}} \mathbf{W}_L) \odot ((\mathbf{W}_V \tilde{\mathbf{V}}^T) \mathbf{N}), \quad (4.40)$$

where the tilde symbol in $\tilde{\mathbf{A}}$, $\tilde{\mathbf{L}}$ and $\tilde{\mathbf{V}}$ denotes squeezed version of matrices \mathbf{A} , \mathbf{L} and \mathbf{V} respectively. Schematically, representation with all the matrices sizes is given in Fig. 4.5. The initial images intensities (matrix \mathbf{M}) and structuring matrices (\mathbf{W}_A , \mathbf{W}_L and \mathbf{W}_V) are known, and all other components ($\tilde{\mathbf{A}}$, $\tilde{\mathbf{L}}$, $\tilde{\mathbf{V}}$ and \mathbf{N}) should be evaluated.

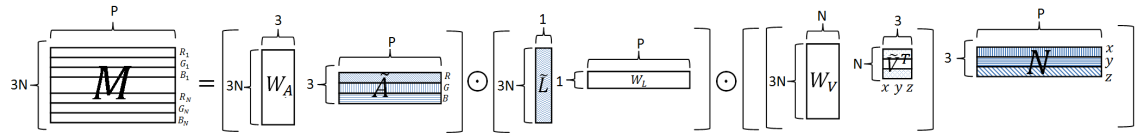


FIGURE 4.5: The proposed PS model matrix representation with structure constraints.

4.2.2 Alternating optimization scheme for photometric stereo

In our proposed method the calculation of the photometric stereo (PS) model components will be done using alternating optimization (AO) principle. In this approach, the order of these components calculation as well as initial values for the alternating optimization are important for the final results.

Initialization

If there is no prior information, initialization of model components, necessary for alternating optimization algorithm, should be made on the basis of the input images. In order to compare the influence of the initialization of model variables, let us have the following experiments. There are four artificially created images with a fixed viewpoint and varying illuminations, Fig. 4.6(a) - 4.6(d). On the basis of these images we are going to reconstruct the surface using alternating optimization and compare the results of the reconstruction with the ground truth (GT) surface, Fig. 4.6(e). The surfaces are presented in false colors with red corresponding to the most elevated and dark blue corresponding to the lowest surface points, as for the colormap of Fig. 4.3.

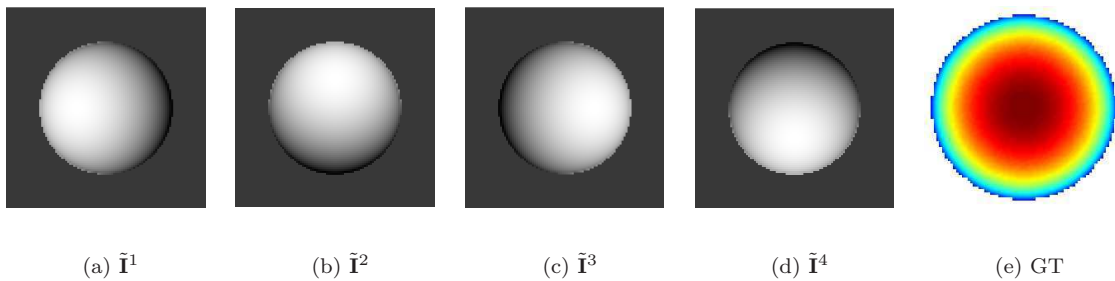


FIGURE 4.6: Input images for initialization tests.

For these experiments on initialization impact, we choose to use the order $\tilde{\mathbf{V}}$, \mathbf{N} , $\tilde{\mathbf{L}}$ and $\tilde{\mathbf{A}}$ for AO. This order will be described in more details in the next section. We choose to first evaluate $\tilde{\mathbf{V}}$, because directions of light sources are difficult to initialize using only input images.

According to the pre-defined order of the components evaluation, let us first test different initial values for the matrix \mathbf{N} while fixing $\tilde{\mathbf{L}}$ and $\tilde{\mathbf{A}}$ as the matrices of ones. In Fig. 4.7, we provide the resulting surface for the following values of \mathbf{N} :

- unit matrix, Fig. 4.7(a);
- normally distributed random values ($\mathcal{G}(0, 1)$), Fig. 4.7(b);
- uniformly distributed random values ($\mathcal{U}(0, 1)$), Fig. 4.7(c);
- 3D surface normals to the median image intensity values (3Dnorm), Fig. 4.7(d).

Only with visual analysis of the obtained surface, we can conclude that the reconstruction process is very sensible to the initialization of the component \mathbf{N} . The initialization with 3D surface normals to the median image intensity values gives the best results of reconstruction. If the image is color, the median can be taken for $3N$ channels as for independent images. The choice of such specific initial values of \mathbf{N} was made due to the fact, that the object shading in images already contains certain information on the object surface. Median operation permits to aggregate intensity values of several input images. After that, the map of intensities is seen as the initial depth map, and normals to this initial depth map are taken as initial normal for the surface evaluation. We use this initialization of \mathbf{N} for all the following experiments.

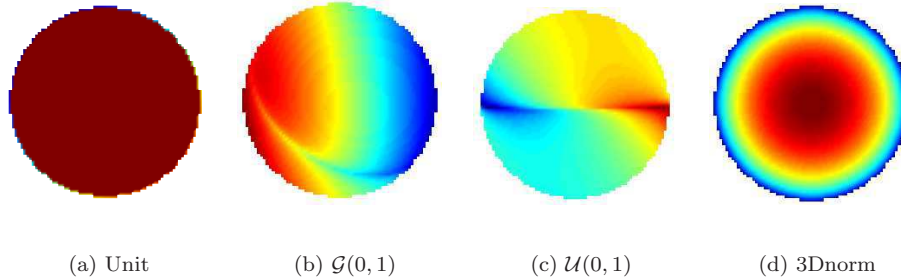


FIGURE 4.7: Results of the reconstruction for different initializations of the normal map \mathbf{N} .

The vector of light source intensities $\tilde{\mathbf{L}}$ can also be initialized in different ways. We test four various initializations for the vector $\tilde{\mathbf{L}}$ while fixing \mathbf{N} as 3D normals to the median intensities and the unit matrix for $\tilde{\mathbf{A}}$. The results of reconstruction for different $\tilde{\mathbf{L}}$ is presented in Fig. 4.8. The vector of light source intensities may be chosen as:

- unit vector, Fig. 4.8(a);
- normally distributed random values ($\mathcal{G}(0, 1)$), Fig. 4.8(b);
- uniformly distributed random values ($\mathcal{U}(0, 1)$), Fig. 4.8(c);

- maximum image intensity values for each image channel (MaxInt), Fig. 4.8(d).

Visual analysis of the obtained surfaces allows us to conclude that the initialization of this vector is not crucial for the reconstruction. The value of light source intensities are strongly related to the values of light source directions. Indeed, illuminating sources intensities can be presented as norms of illuminating sources directions. In all our following experiments we choose initial $\tilde{\mathbf{L}}$ as a vector of maximum image intensity values for each image channel.

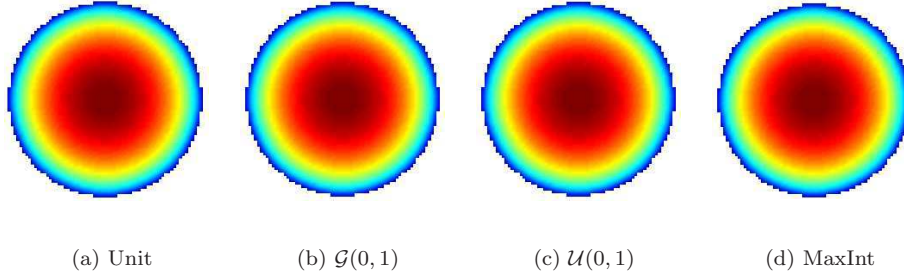


FIGURE 4.8: Initialization of the matrix of light source intensities.

We tested four different initializations for the matrix of albedo values $\tilde{\mathbf{A}}$:

- unit matrix, Fig. 4.9(a);
- median values for each color channel (MedRGB), Fig. 4.9(b);
- normally distributed random values ($\mathcal{G}(0, 1)$), Fig. 4.9(c);
- uniformly distributed random values ($\mathcal{U}(0, 1)$), Fig. 4.9(d).

The obtained results show the importance of the matrix $\tilde{\mathbf{A}}$ initialization, and randomly chosen values are not suitable for the qualitative reconstruction. Initialization with unit matrix and median values for each channel gave the same results of reconstruction. For the reason of computation simplicity, in our following experiments we choose $\tilde{\mathbf{A}}$ as a unit matrix.

To conclude, the initialization of the model components is important for the alternating optimization procedure. The initialization of $\tilde{\mathbf{A}}$ and \mathbf{N} matrices is crucial for the reconstruction results. At the same time, the initialization of the matrix $\tilde{\mathbf{L}}$ does not influence the final reconstruction. According to the hold experiments, we initialize values in the matrix \mathbf{N} as 3D surface normals to the grayscale median image intensity values, vector $\tilde{\mathbf{L}}$ contains maximum intensity values of each input image, and the matrix of albedo $\tilde{\mathbf{A}}$ is initialized as the unit matrix.

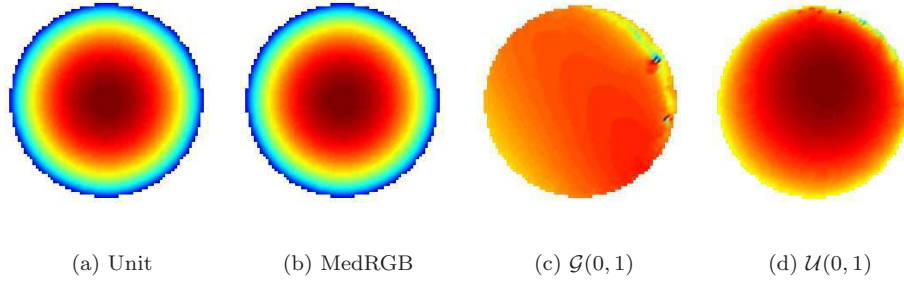


FIGURE 4.9: Initialization of the matrix of albedo values $\tilde{\mathbf{A}}$.

Order of components evaluation

Since the order of calculation is also crucial for the AO algorithms, we tested different orders of the model components calculation. Let us have a hemisphere, which can be described by the normals map and the depth map in false colors, Fig.4.10(a). The normal map is encoded by the RGB colors: portions of red, green and blue colors corresponds to the component $n_{x,j}$, $n_{y,j}$ and $n_{z,j}$ respectively, $j = 1, \dots, P$. This hemisphere is photographed from above with 4 opposite lighting sources, as in Fig. 4.6. On the basis of the obtained images we reconstruct the normal and the depth for each pixel and compare the resulting normals and depth maps with the ground truth. In all the tested configurations we firstly evaluate the values of light sources directions $\tilde{\mathbf{V}}$. As described in previous section, there is no prior information on these variables and they are difficult for an empirical initialization from images.

As it can be seen from Fig. 4.10, for the orders $(\tilde{\mathbf{V}}, \mathbf{N}, \tilde{\mathbf{L}}, \tilde{\mathbf{A}})$, $(\tilde{\mathbf{V}}, \mathbf{N}, \tilde{\mathbf{A}}, \tilde{\mathbf{L}})$ and $(\tilde{\mathbf{V}}, \tilde{\mathbf{L}}, \mathbf{N}, \tilde{\mathbf{A}})$ we obtain identical results of reconstruction for both, normals and depth maps, and these results are close to the ground truth. At the same time, for the other orders $(\tilde{\mathbf{V}}, \tilde{\mathbf{L}}, \tilde{\mathbf{A}}, \mathbf{N})$, $(\tilde{\mathbf{V}}, \tilde{\mathbf{A}}, \mathbf{N}, \tilde{\mathbf{L}})$ and $(\tilde{\mathbf{V}}, \tilde{\mathbf{A}}, \tilde{\mathbf{L}}, \mathbf{N})$ we obtain normals maps with normal outlier artifacts, but the reconstructed depth map is still close to the ground truth. It means that the normals are well evaluated except for several outlier points, and negative effect of these points is smoothed during the integration of the normals map to obtain the depth map. On the basis of the obtained results, it can be concluded that the matrix $\tilde{\mathbf{A}}$ should be evaluated after the matrix \mathbf{N} in order to avoid outlier values in the resulting normals map.

In our algorithm, we propose to estimate values of $\tilde{\mathbf{V}}$ and \mathbf{N} matrices in the first place. These two variables represent the interaction between the light sources and surface normals, which results in the shading. The matrices of albedo values $\tilde{\mathbf{A}}$ and light source intensities $\tilde{\mathbf{L}}$ can be seen as weights for the principal matrices $\tilde{\mathbf{V}}$ and

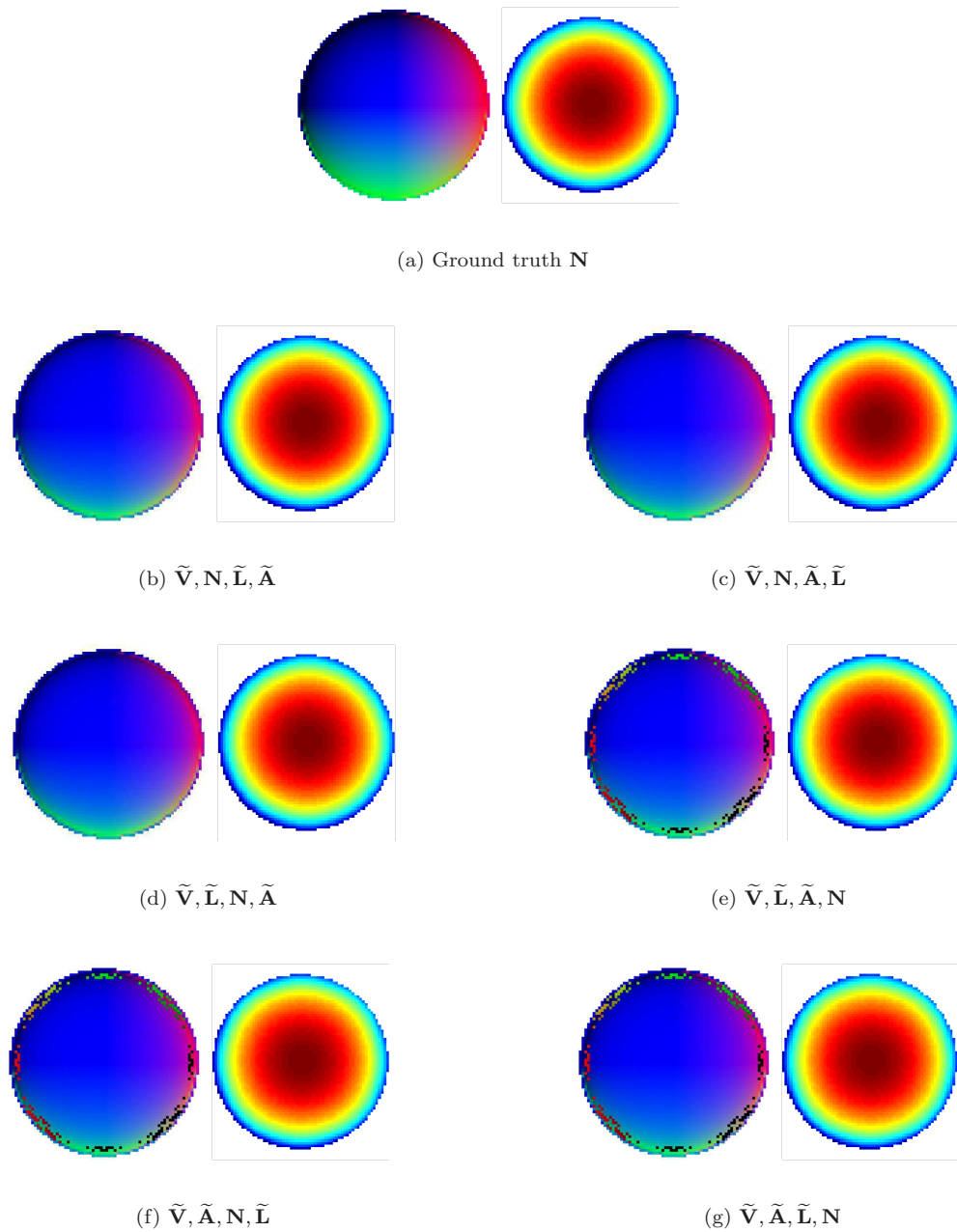


FIGURE 4.10: Testing of different alternating optimization algorithm configurations for simple artificial object reconstruction.

\mathbf{N} , respectively, and order of their evaluation is not so important for the resulting surface, Fig. 4.10(b) and Fig. 4.10(c).

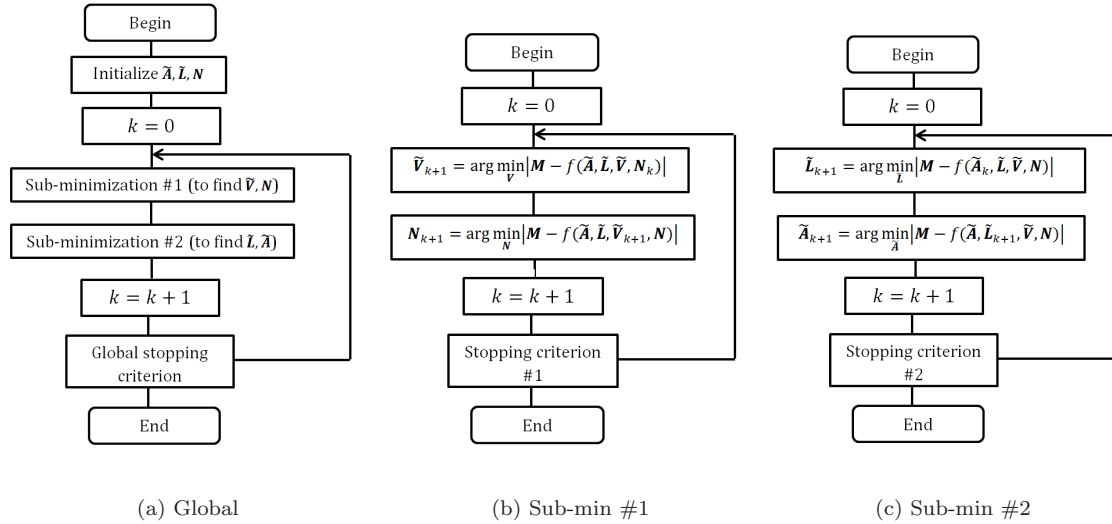


FIGURE 4.11: Alternating optimization for the PS model evaluation.

Alternating optimization with sub-minimizations

According to the fact, that $\tilde{\mathbf{V}}$ and \mathbf{N} components are more crucial for the alternating procedure and their directional nature is different from the other two components $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{L}}$, we propose to divide the global minimization procedure into two sub-minimizations. One of these sub-minimizations contains evaluation of the two matrices of directions, $\tilde{\mathbf{V}}$ and \mathbf{N} , and the second sub-minimization contains the evaluation of the weighting matrices $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{L}}$ also as independent model components.

Global algorithm for the model components evaluation on the basis of the optimization is shown in Fig. 4.11(a). It consists in alternating two sub-minimization procedures ('Sub-minimization #1' and 'Sub-minimization #2'), which are presented in details in Fig. 4.11(b) and Fig. 4.11(c), respectively.

The matrix of observations \mathbf{M} in Fig. 4.11 is of size $[3N \times P]$ and corresponds to the matrix of the vectorized input color images. The model of image formation process corresponds to the multiplicative model also presented in the previous section, $f(\tilde{\mathbf{A}}, \tilde{\mathbf{L}}, \tilde{\mathbf{V}}, \mathbf{N}) = (\mathbf{W}_A \tilde{\mathbf{A}}) \odot (\tilde{\mathbf{L}} \mathbf{W}_L) \odot ((\mathbf{W}_V \tilde{\mathbf{V}}^T) \mathbf{N})$, Eq. 4.40. Loss function, or error between the observation matrix and the model can be written as $e = \left\| \mathbf{M} - f(\tilde{\mathbf{A}}, \tilde{\mathbf{L}}, \tilde{\mathbf{V}}, \mathbf{N}) \right\|_2^2$. For searching each model component any existing minimization algorithm can be used. For example, the least squares method (LSM), [25], is efficient to find the model components on each minimization stage. The LSM employment for all stages of the alternating optimization is entitled as the alternating

Algorithm 3 The model components evaluation by the AO without constraints.

```

1: In:
    $\mathbf{M}$  - the matrix  $[3N \times P]$  of vectorized color (at least three) images;
2: Out:
    $\mathbf{A}$  - the matrix  $[3 \times P]$  of the albedo values;
    $\mathbf{L}$  - the vector  $[3N \times 1]$  of the light sources intensities;
    $\mathbf{V}$  - the matrix  $[3 \times N]$  of the light sources directions;
    $\mathbf{N}$  - the matrix  $[3 \times P]$  of the normals values.
3:  $\mathbf{N} \leftarrow \text{InitializeN}(\mathbf{M})$ 
4:  $\mathbf{A} \leftarrow \text{InitializeA}(P, N)$ 
5:  $\mathbf{L} \leftarrow \text{InitializeL}(\mathbf{M})$ 
6: repeat
7:   repeat
8:      $\mathbf{V} \leftarrow \text{EvalVNoConstr}(\mathbf{M}, \mathbf{A}, \mathbf{L}, \mathbf{N})$ 
9:      $\mathbf{N} \leftarrow \text{EvalNNoConstr}(\mathbf{M}, \mathbf{A}, \mathbf{L}, \mathbf{V})$ 
10:    until LocStopCriter1( $\mathbf{M}, \mathbf{A}, \mathbf{L}, \mathbf{V}, \mathbf{N}$ )
11:   repeat
12:      $\mathbf{L} \leftarrow \text{EvalLNoConstr}(\mathbf{M}, \mathbf{A}, \mathbf{V}, \mathbf{N})$ 
13:      $\mathbf{A} \leftarrow \text{EvalANoConstr}(\mathbf{M}, \mathbf{L}, \mathbf{V}, \mathbf{N})$ 
14:   until LocStopCriter2( $\mathbf{M}, \mathbf{A}, \mathbf{L}, \mathbf{V}, \mathbf{N}$ )
15: until GlobStopCriter( $\mathbf{M}, \mathbf{A}, \mathbf{L}, \mathbf{V}, \mathbf{N}$ )

```

least squares method (ALS), [110]. The presented algorithm is deterministic, which means that for the same initialization we always obtain the same results.

We do not detail here the stopping criteria choice, but we return to this important question in Section 4.2.4.

All the mathematical transformations used to find the final equations separately for each model component by the means of the LSM are detailed in Appendix B.

Results of surface reconstruction

The chosen order of the model components evaluation by the alternating optimization approach is presented in the final Algorithm 3. The mathematical interpretation of the algorithm functions `EvalANoConstr()`, `EvalLNoConstr()`, `EvalVNoConstr()` and `EvalNNoConstr()` correspond to the appropriate Appendix sections B.1, B.2, B.3, and B.4.

Before the reconstruction results demonstration, let us define the measure which is used to visualize the algorithm convergence. In all the following examples, for each algorithm iteration k we calculate the following relative deviation of the estimated

components and initial data matrix \mathbf{M} :

$$e_k = \log_{10} \left(\frac{\|\mathbf{M} - f(\tilde{\mathbf{A}}_k, \tilde{\mathbf{L}}_k, \tilde{\mathbf{V}}_k, \mathbf{N}_k)\|_2^2}{\|\mathbf{M}\|_2^2} \right). \quad (4.41)$$

Logarithmic function in the equation allows to stretch too small values of error, which is useful for visualization. Relative error calculation (division by the data matrix) makes it comparable for different ranges of image intensity values in \mathbf{M} .

We will now present results of reconstruction by the proposed approach. It is demonstrated a comparison of the reconstruction results for a simple hemispherical object and for the object from the MIT Intrinsic Images dataset, [111]. For the presented examples, the alternating scheme from Fig. 4.11 was used with 10 global iterations and 5 internal iterations for each sub-minimization. We propose to visualize the depth maps in a gray-scale format. It allows observing more object surface details which are hidden by a false color representation. Initial images for this example are given in Fig. 4.12(a) - 4.12(d), there are four color images illuminated from opposite sides. The obtained results of reconstruction for the simple hemispherical object, Fig. 4.12, are with small error measure $\mathbf{z}_{RMSE} = 0.04$. According to visual surface analysis, the obtained surface is close to the ground truth. The algorithm convergence for this object, Fig. 4.12(g), indicates the qualitative reconstruction for the first iteration and further non-significant fluctuations of the error. This rapid convergence can be explained by the fact that the artificial hemisphere was created exactly with the same multiplicative model which is used for the surface reconstruction. Moreover, this surface does not have sharp passages or small details, and used four algorithm input images explicitly describe it.

On the other hand, for more complex object. We demonstrate an example of reconstruction for 10 initial images, Fig. 4.13(a) - 4.13(j). In Fig. 4.14 we obtain non-satisfactory results of reconstruction. The algorithm for this surface does not converge, Fig. 4.14(c), for the presented 10 global iterations.

Even if for the simple object as the hemisphere, the alternating optimization is correct, more complicated surfaces are badly reconstructed. We then propose to introduce constraints on model variables to cope with this problem.

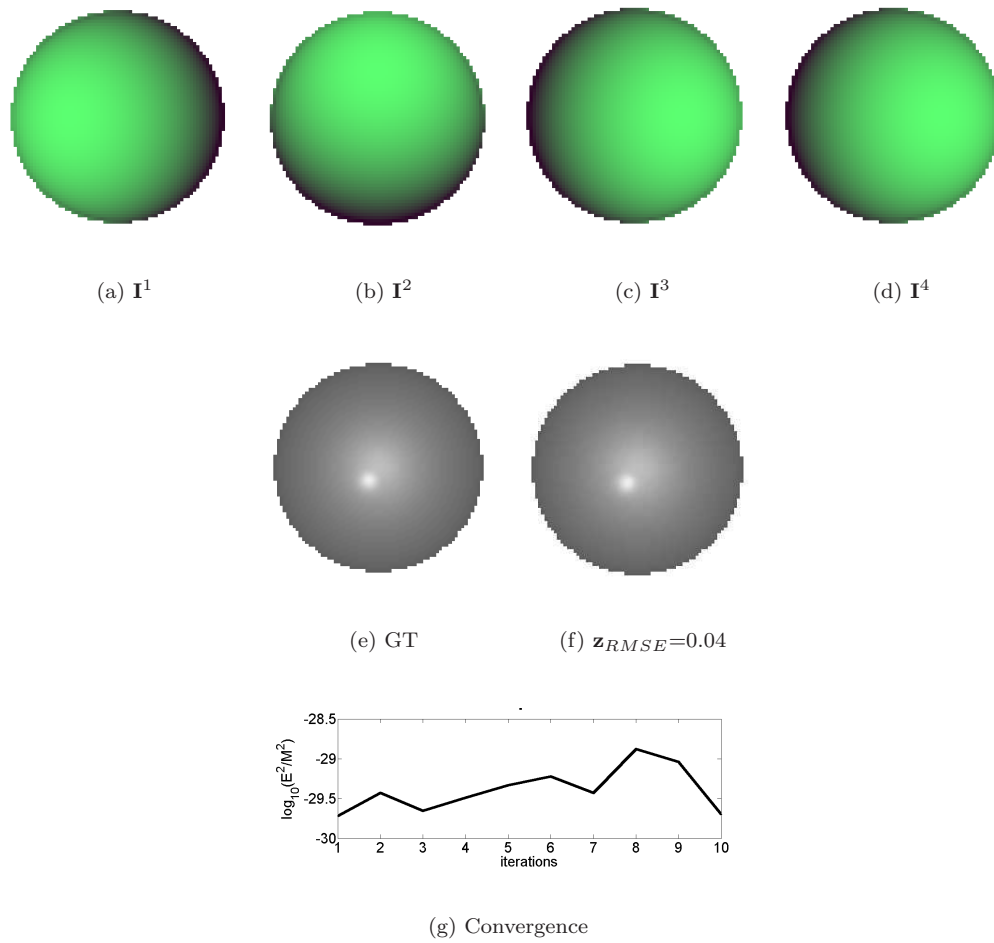


FIGURE 4.12: Reconstruction of an artificial hemisphere using AO without constraints.

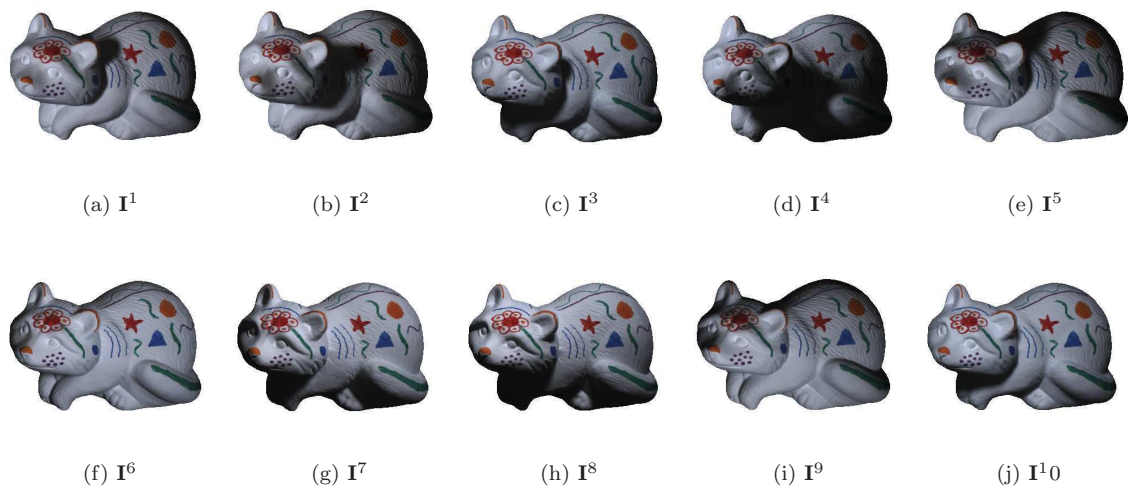


FIGURE 4.13: Initial images of a complex surface.

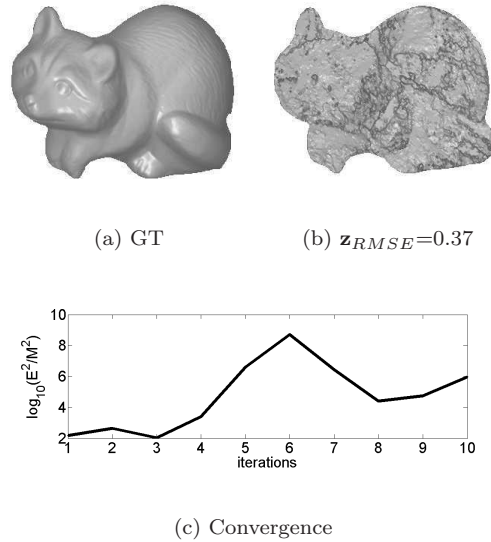


FIGURE 4.14: Reconstruction of a complex object using AO without constraints.

4.2.3 Constrained alternating optimization

Constraints on values for the PS model components are widely used in order to compensate a lack of information and overcome an ill-posed problem. The shape-from-shading approach can be seen as a special case of the photometric stereo when there is only one input image, and the number of unknown variables is several times greater than the number of input values. That is why different types of constraints are popular in shape-from-shading techniques (widely applied in [23]) and less presented in photometric stereo (e.g., as in [103] - quality constraint for normals, or in [112] - bilateral symmetry constraint for isotropic BRDFs). We provide our model with constraints in order to better present physical nature of each component. We define two groups of constraints: value and quality constraints. Each of these groups has its own influence on the resolution process and is explained in details below.

Value constraints

This category contains interval constraints for albedo and light source intensity values as well as norm constraints for normals and lighting directions vectors.

Albedo denotes an ability of a surface to reflect and absorb incident light. In fact, it is a ratio between reflected and incident quantity of light. The more light is absorbed by the surface material, the less light is reflected, and the smaller value of the albedo is defined for the point. The albedo value in the matrix $\tilde{\mathbf{A}}$ of Eq. 4.14(c) is in the

interval between zero and one:

$$\tilde{a}_{r,j} \in [0, 1]. \quad (4.42)$$

The values of light source intensities are non-negative. If the light source intensity is zero, this signifies that the point is not visible on an image. Even if the value of the light source intensity is larger than the maximal possible value of the image intensity, it would be registered with error by the matrix of camera sensors would be still equal to the maximal possible intensity value. Thus, we intend always to have light sources intensities in the vector $\tilde{\mathbf{L}}$ of Eq. 4.40 greater than zero and less than maximal possible value of intensity. For a general case, images intensities can be in interval between 0 and H :

$$\tilde{l}_i \in (0, H). \quad (4.43)$$

The maximal possible value $H = (2^b - 1)$, where b corresponds to the number of bits allocated for each image pixel. If the image is presented in the interval $[0, 1]$, then $H = 1$, on the other hand, for images which are defined by the intensity values $[0, 255]$, $H = 255$.

Norm constraint for the lighting directions matrix $\tilde{\mathbf{V}}$ in Eq. 4.40 signifies that norm of each line of this matrix, which corresponds to a particular lighting source, equals to one. It can be formally written as:

$$\tilde{v}_{i,x}^2 + \tilde{v}_{i,y}^2 + \tilde{v}_{i,z}^2 = 1. \quad (4.44)$$

Norm constraint for the surface normal matrix \mathbf{N} in Eq. 4.40 signifies that norm of each column of this matrix, which corresponds to the image pixel, equals to one:

$$n_{x,j}^2 + n_{y,j}^2 + n_{z,j}^2 = 1. \quad (4.45)$$

All value constraints are intuitive and empirical but only their explicit usage gives an adequate and correctly interpretable problem solution.

Value constraints post-implementation

One of the existing numerical solutions is to integrate value constraints into each iterative step by projection of the obtained values to the space of constraints. Each component can be evaluated as in Algorithm 3 using least squares method with no

constraints. Then, in the iterative procedure after each sub-minimization, we can project the found values to the space of the pre-defined constraints. Lighting directions $\tilde{\mathbf{V}}$ and surface normals \mathbf{N} should be simply normalized after 'Sub-minimization #1' evaluation, Fig. 4.11(a). For albedo component $\tilde{\mathbf{A}}$, it means to equate all values less than zero to zero, and to set to one all values greater than one inside of 'Sub-minimization #2'. The similar operation can be made for lighting sources intensities. All values in found $\tilde{\mathbf{L}}$ matrix under zero should be set to zero or to a very small value. Such a post-evaluation constraints application is possible thanks to the iterative nature of the algorithm.

The example of surface reconstruction by the means of value constraints post-implementation is given in Fig. 4.15. This is the surface which we were unable to reconstruct by the alternating scheme without constraints implementation, Fig. 4.14. With presented value constraints, the resulting surface is of an acceptable quality with $\mathbf{z}_{RMSE} = 0.1461$. The convergence of the reconstruction algorithm for these data in logarithmic scale is shown in Fig. 4.15(c). Already after the 6th iteration there is no significant changes of the model error.

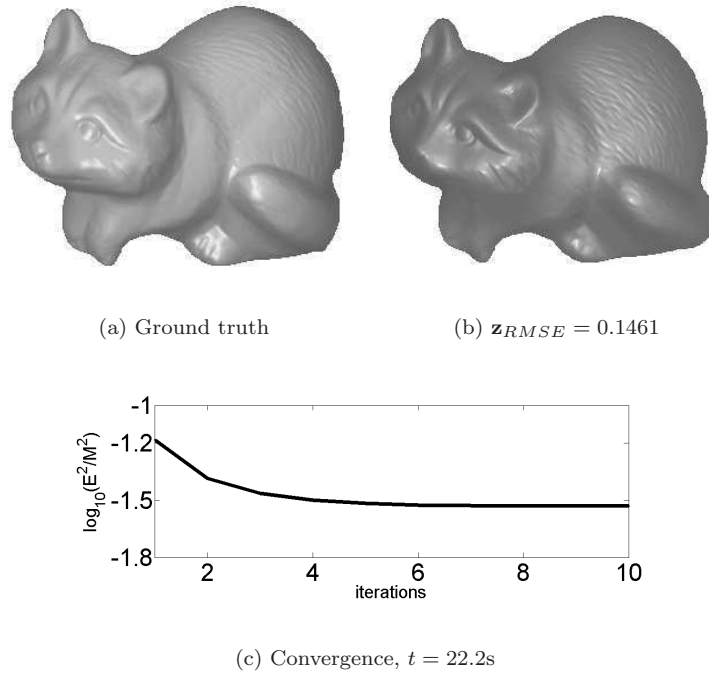


FIGURE 4.15: Post-implementation of value constraints.

This approach of post-evaluation constraints is very fast and can be used as a primary photometric stereo model components approximation. However, post-implementation

of constraints does not guarantee optimal solution at each algorithm iteration, and as a result, the obtained final surface can be not sufficiently exact.

Integrated value constraints

In order to integrate value restrictions for the photometric stereo components, the minimization functions for finding albedo values, light sources intensities, light sources directions and surface normals matrices should be expanded including supplementary equality and inequality constraints. Let us rewrite the minimization function for each of these matrices.

In Fig. 4.11(b) light sources directions matrix including norm constraints should be found using the following expression for the $(k^{th} + 1)$ iteration of the inner loop of minimization:

$$\begin{aligned} \tilde{\mathbf{V}}_{k+1} &= \arg \min_{\tilde{\mathbf{V}}} \left\| \mathbf{M} - f(\tilde{\mathbf{A}}, \tilde{\mathbf{L}}, \tilde{\mathbf{V}}, \mathbf{N}_k) \right\|_2^2, \\ &\text{subject to } \tilde{v}_{i,x}^2 + \tilde{v}_{i,y}^2 + \tilde{v}_{i,z}^2 = 1, \end{aligned} \quad (4.46)$$

where $i = 1, \dots, N$, and vector $(\tilde{v}_{i,x}, \tilde{v}_{i,y}, \tilde{v}_{i,z})$ corresponds to the lighting direction used during the acquisition of the i^{th} image. All the mathematical transformations as well as the final expression for the matrix $\tilde{\mathbf{V}}$ evaluation are given in Appendix C.3.

In order to find surface normals matrix with value norm constraint, the minimization procedure for matrix \mathbf{N}_{k+1} in Fig. 4.11(b) should be replaced by the following one:

$$\begin{aligned} \mathbf{N}_{k+1} &= \arg \min_{\mathbf{N}} \left\| \mathbf{M} - f(\tilde{\mathbf{A}}, \tilde{\mathbf{L}}, \tilde{\mathbf{V}}_{k+1}, \mathbf{N}) \right\|_2^2, \\ &\text{subject to } n_{x,j}^2 + n_{y,j}^2 + n_{z,j}^2 = 1, \end{aligned} \quad (4.47)$$

where $j = 1, \dots, P$, $(n_{x,j}, n_{y,j}, n_{z,j})$ corresponds to 3D normal of the j^{th} pixel.

Light sources intensity vector should contain only values greater than 0 and less than the maximal possible value of the image, H . In order to realize this constraint, one should rewritten the formula for $\tilde{\mathbf{L}}_{k+1}$ finding in Fig. 4.11(c) as:

$$\tilde{\mathbf{L}}_{k+1} = \arg \min_{\tilde{\mathbf{L}}} \left\| \mathbf{M} - f(\tilde{\mathbf{A}}_k, \tilde{\mathbf{L}}, \tilde{\mathbf{V}}, \mathbf{N}) \right\|_2^2, \quad (4.48)$$

$$\text{subject to } \begin{cases} \tilde{l}_i > 0, \\ \tilde{l}_i < H, \end{cases}$$

where $i = 1, \dots, 3N$, N is the number of input color images, and \tilde{l}_i corresponds to the intensity of light source used during the acquisition of the i^{th} input image. The resolution for the component $\tilde{\mathbf{L}}$ is presented in Appendix C.2.

And finally, albedo matrix should contain only values from the interval $[0, 1]$. To implement this inequality constraint the minimization function for albedo approximation in Fig. 4.11(c) should be replaced by the following one:

$$\tilde{\mathbf{A}}_{k+1} = \arg \min_{\tilde{\mathbf{A}}} \left\| \mathbf{M} - f(\tilde{\mathbf{A}}, \tilde{\mathbf{L}}_{k+1}, \tilde{\mathbf{V}}, \mathbf{N}) \right\|_2^2, \quad (4.49)$$

$$\text{subject to } \begin{cases} \tilde{a}_{r,j} \geq 0, \\ \tilde{a}_{r,j} \leq 1, \end{cases}$$

where $r = 1, \dots, N$, $j = 1, \dots, P$, N is the number of input color images and P is the number of pixel in each input image. Mathematical transformations and final expression for $\tilde{\mathbf{A}}$ can be found in Appendix C.1.

Resolution of the minimization problems with inequality constraints (albedo, Eq. 4.49, and light sources intensities, Eq. 4.48) can be found by analyzing Karush-Kuhn-Tucker (KKT) conditions, [113], for the optimization error function including Lagrange multipliers, [114]. This analysis is done in Appendices C.1 and C.2. Due to a nonlinear nature of the equality constraints for lighting directions, Eq. 4.46, and surface normals, Eq. 4.47, the best solution for sub-minimizations can be found using Levenberg-Marquardt algorithm for the loss function also including Lagrange multipliers, [115]. This algorithm general representation is given in Appendix D.

The obtained results of reconstruction by the alternating optimization with integrated constraints are demonstrated in Fig. 4.16. The quality of the surface reconstruction for 10 iterations of the global algorithm is slightly better than for the value-constraints post-implementation (for this algorithm $\mathbf{z}_{RMSE} = 0.1329$ comparing to $\mathbf{z}_{RMSE} = 0.1461$ of the post-implementation). The final error in the graphics of the algorithm convergence, Fig. 4.16(c), is also lower than for the case of the constraints post-implementation. The main weakness of the algorithm of integrated constraints implementation consists in long computation time ($t = 931.97\text{s}$ comparing to $t = 22.7\text{s}$ for the post-implementation).

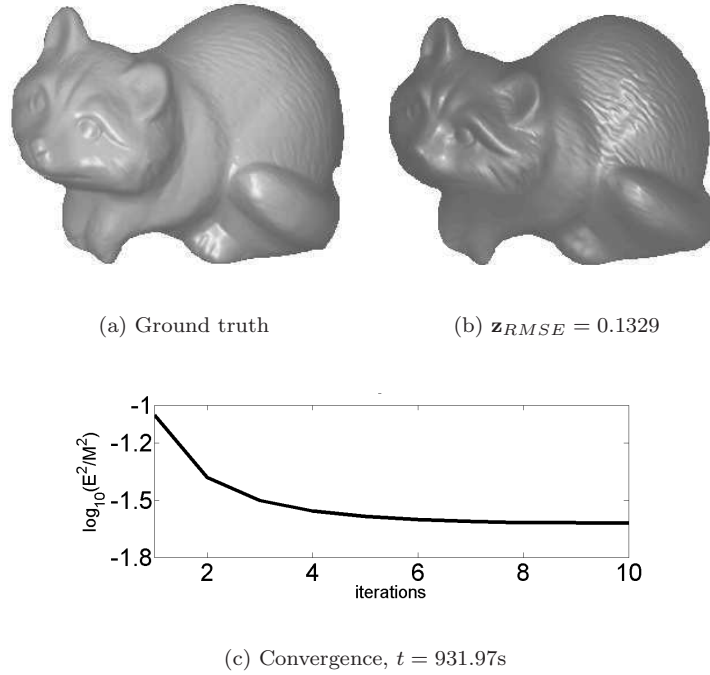


FIGURE 4.16: Embedded value constraints.

Quality constraint

Additionally to value constraints, a quality constraint for the normals can also be used. This constraint influences on the visual quality of the resulting surface. It is not mandatory and can be complemented by others if there are special requirements for the final normals.

This is a constraint of smoothness. We suppose that the resulting surface has no sharp transitions. For this constraint post-implementation, found surface normals should be filtered by an averaging low-pass filter with $[3 \times 3]$ sliding window in each iteration of the global algorithm. The algorithm of AO with post-implementation of the value and quality constraints is summarized in Algorithm 4. The results with such a filter application after value constraints post-implementation are given in Fig. 4.17. With a small improvement of the surface quality for 10 global iterations ($z_{RMSE} = 0.1446$ comparing to $z_{RMSE} = 0.1461$ obtained for post-implementation of only value constraints) we obtain the algorithm convergence graphics given in Fig. 4.17(c). Time of evaluation remain the same, the low-pass filter application is not time-consuming compared to the whole AO procedure.

For the integration of this constraint directly into the minimization procedure, the cosine between the current surface normal \mathbf{n} and each of its 8 neighbors should tend

Algorithm 4 The model components evaluation by the AO with constraints post-implementation

```

1: In:
    $\mathbf{M}$  - the matrix  $[3N \times P]$  of vectorized color (at least three) images;
2: Out:
    $\mathbf{A}$  - the matrix  $[3 \times P]$  of the albedo values;
    $\mathbf{L}$  - the vector  $[3N \times 1]$  of the light sources intensities;
    $\mathbf{V}$  - the matrix  $[3 \times N]$  of the light sources directions;
    $\mathbf{N}$  - the matrix  $[3 \times P]$  of the normals values.
3:  $\mathbf{N} \leftarrow \text{InitializeN}(\mathbf{M})$ 
4:  $\mathbf{A} \leftarrow \text{InitializeA}(P, N)$ 
5:  $\mathbf{L} \leftarrow \text{InitializeL}(\mathbf{M})$ 
6: repeat
7:   repeat
8:      $\mathbf{V} \leftarrow \text{EvalVNoConstr}(\mathbf{M}, \mathbf{A}, \mathbf{L}, \mathbf{N})$ 
9:      $\mathbf{N} \leftarrow \text{EvalNNoConstr}(\mathbf{M}, \mathbf{A}, \mathbf{L}, \mathbf{V})$ 
10:    until LocStopCriter1( $\mathbf{M}, \mathbf{A}, \mathbf{L}, \mathbf{V}, \mathbf{N}$ )
11:     $\mathbf{N} \leftarrow \text{LowPass2DFilter}(\mathbf{N})$ 
12:     $\mathbf{N} \leftarrow \text{Normalize}(\mathbf{N})$ 
13:     $\mathbf{V} \leftarrow \text{Normalize}(\mathbf{V})$ 
14:    repeat
15:       $\mathbf{L} \leftarrow \text{EvalLNoConstr}(\mathbf{M}, \mathbf{A}, \mathbf{V}, \mathbf{N})$ 
16:       $\mathbf{L} \leftarrow \text{LimitL}(\mathbf{L})$ 
17:       $\mathbf{A} \leftarrow \text{EvalANoConstr}(\mathbf{M}, \mathbf{L}, \mathbf{V}, \mathbf{N})$ 
18:       $\mathbf{A} \leftarrow \text{LimitA}(\mathbf{A})$ 
19:    until LocStopCriter2( $\mathbf{M}, \mathbf{A}, \mathbf{L}, \mathbf{V}, \mathbf{N}$ )
20: until GlobStopCriter( $\mathbf{M}, \mathbf{A}, \mathbf{L}, \mathbf{V}, \mathbf{N}$ )

```

to one. It can be implemented by minimization of the expression:

$$\min_{\mathbf{n}} \sum_{\substack{p=-1 \\ p,l \neq 0}}^1 \sum_{l=-1}^1 (\mathbf{n}^T \mathbf{n}^{(p,l)} - 8)^2, \quad (4.50)$$

where $\mathbf{n}^{(p,l)}$ defines normals of 8 neighbor pixels of the current pixel in 2D image space.

The minimization procedure for matrix \mathbf{N}_{k+1} in Fig. 4.11(b) should be then replaced by the following one:

$$\mathbf{N}_{k+1} = \arg \min_{\mathbf{N}} \left\| \mathbf{M} - f(\tilde{\mathbf{A}}, \tilde{\mathbf{L}}, \tilde{\mathbf{V}}_{k+1}, \mathbf{N}) \right\|_2^2, \quad (4.51)$$

$$\text{subject to } \begin{cases} n_{x,j}^2 + n_{y,j}^2 + n_{z,j}^2 = 1, \\ \sum_{\substack{p=-1 \\ p \neq 0, l \neq 0}}^1 \sum_{l=-1}^1 \mathbf{n}_j^T \mathbf{n}_j^{(p,l)} = 8, \end{cases}$$

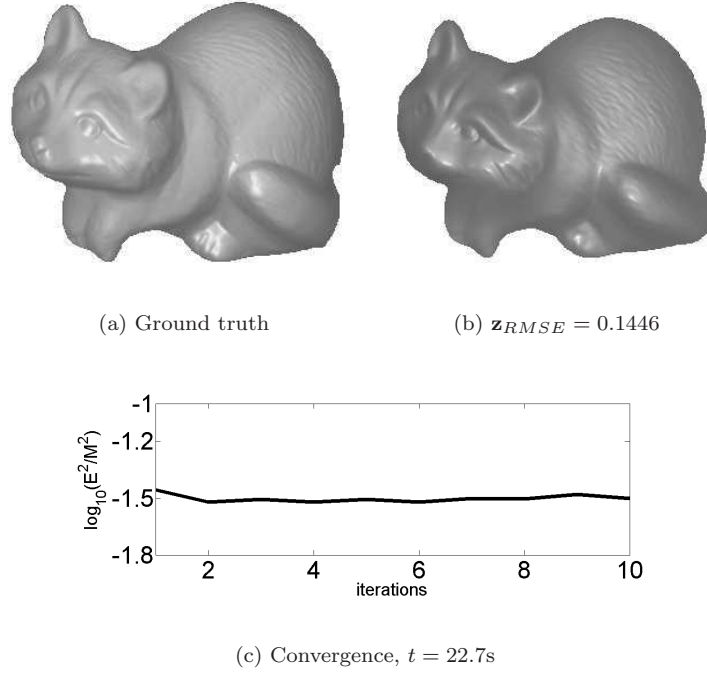


FIGURE 4.17: Quality and value constraints post-implementation.

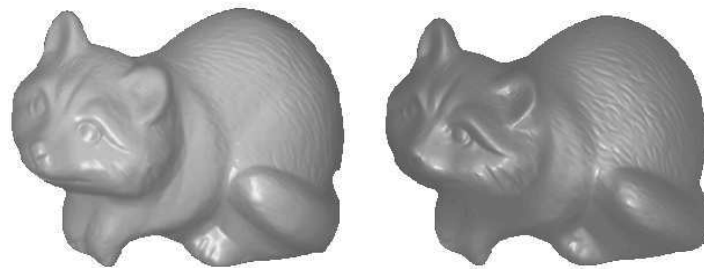
where $j = 1, \dots, P$, \mathbf{n}_j corresponds to 3D normal of the j^{th} pixel, and $\mathbf{n}_j^{(p,l)}$ denotes normal of the j^{th} pixel neighbors in the initial not vectorized image. The resulting system of equations for the matrix \mathbf{N} calculation with embedded value and quality constraints is presented in Appendix C.4.

The results of reconstruction for the value and quality constraints integrated into minimizations are given in Fig. 4.18. The final algorithm of the model components evaluation with embedded constraints, Algorithm 5, is similar in structure with the algorithm which exploits non-constrained ALS, Algorithm 3. However, the constrained versions of the algorithm functions EvalAWithConstr(), EvalLWithConstr(), EvalVWithConstr() and EvalNWithConstr() are computationally more complex, and the resulting matrices contain values compatible with their physical significations.

As for the case of only value constraints integration, the obtained computation time (934.94s) is considerably elevated compared to the post-implementation version of the same constraints (22.7s). The measure of $\mathbf{z}_{RMSE} = 0.1461$ is slightly higher than $\mathbf{z}_{RMSE} = 0.1446$ obtained for the post-implementation.

Visually, the surfaces evaluated with the additional surface quality constraint is smoother but the obtained errors do not show significant improvements comparing to

the usage of value constraints. On the other hand, the quality smoothness constraint is not mandatory and should be used with caution. Its usage can provoke a loss of small details important for further surface analysis.



(a) Ground truth

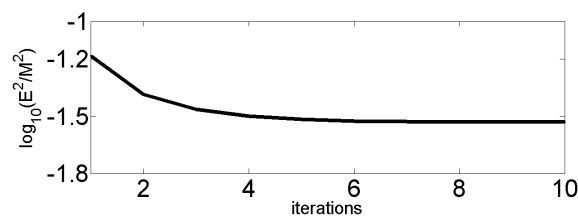
(b) $z_{RMSE} = 0.1461$ (c) Convergence, $t = 934.94s$

FIGURE 4.18: Embedded quality and value constraints.

The results of reconstruction obtained by algorithm of the constraints post-implementation, Algorithm 4, are comparable with the results obtained by the means of algorithm of the embedded constraints implementations, Algorithm 5, in terms of z_{RMSE} . However, Algorithm 4 gains in the computation time. This is the main reason why we choose it for all our following experiments.

4.2.4 Stopping criteria

Stopping criteria is an important parameter of an iterative algorithm. There exist three main types of stopping criteria:

- limited number of iterations;
- error between the initial data and the evaluated model;
- variation of variables (or just one, the most important variable) from one iteration to another.

Algorithm 5 The model components evaluation by the AO with embedded constraints

```

1: In:
    $\mathbf{M}$  - the matrix  $[3N \times P]$  of vectorized color (at least three) images;
2: Out:
    $\mathbf{A}$  - the matrix  $[3 \times P]$  of the albedo values;
    $\mathbf{L}$  - the vector  $[3N \times 1]$  of the light sources intensities;
    $\mathbf{V}$  - the matrix  $[3 \times N]$  of the light sources directions;
    $\mathbf{N}$  - the matrix  $[3 \times P]$  of the normals values.
3:  $\mathbf{N} \leftarrow \text{InitializeN}(\mathbf{M})$ 
4:  $\mathbf{A} \leftarrow \text{InitializeA}(P, N)$ 
5:  $\mathbf{L} \leftarrow \text{InitializeL}(\mathbf{M})$ 
6: repeat
7:   repeat
8:      $\mathbf{V} \leftarrow \text{EvalVWithConstr}(\mathbf{M}, \mathbf{A}, \mathbf{L}, \mathbf{N})$ 
9:      $\mathbf{N} \leftarrow \text{EvalNWithConstr}(\mathbf{M}, \mathbf{A}, \mathbf{L}, \mathbf{V})$ 
10:  until LocStopCriter1( $\mathbf{M}, \mathbf{A}, \mathbf{L}, \mathbf{V}, \mathbf{N}$ )
11:  repeat
12:     $\mathbf{L} \leftarrow \text{EvalLWithConstr}(\mathbf{M}, \mathbf{A}, \mathbf{V}, \mathbf{N})$ 
13:     $\mathbf{A} \leftarrow \text{EvalAWithConstr}(\mathbf{M}, \mathbf{L}, \mathbf{V}, \mathbf{N})$ 
14:  until LocStopCriter2( $\mathbf{M}, \mathbf{A}, \mathbf{L}, \mathbf{V}, \mathbf{N}$ )
15: until GlobStopCriter( $\mathbf{M}, \mathbf{A}, \mathbf{L}, \mathbf{V}, \mathbf{N}$ )

```

There are three stopping criteria in the proposed algorithm, Fig. 4.11: the global stopping criterion to stop the alternating optimization procedure, Fig. 4.11(a), and two internal criteria to stop each sub-minimization, Fig. 4.11(b) and Fig. 4.11(c).

Let us have only one global iteration for the surface reconstruction, Fig. 4.17. Inside of this global iteration we choose two internal stopping criteria as limited numbers of sub-loops equals to 5. It means that inside the global minimization procedure we execute the first and the second sub-minimizations 5 times each. The first sub-minimization is dedicated to the evaluation of $\tilde{\mathbf{V}}$ and \mathbf{N} matrices, and the second one is for $\tilde{\mathbf{L}}$ and $\tilde{\mathbf{A}}$ matrices. In Fig. 4.19 we present the convergence of the error after each component evaluation. It means, that the first point of the graphics corresponds to the model error calculated with the newly evaluated $\tilde{\mathbf{V}}_1$, the second point corresponds to the model error evaluated with \mathbf{N}_1 , the third point is for $\tilde{\mathbf{V}}_2$, the fourth point is for \mathbf{N}_2 , and so on till the 5th iteration. After that, we pass to the sub-cycle of $\tilde{\mathbf{L}}$ and $\tilde{\mathbf{A}}$ components. The error evaluation for these components also alternates between $\tilde{\mathbf{L}}_1$, $\tilde{\mathbf{A}}_1$, $\tilde{\mathbf{L}}_2$, $\tilde{\mathbf{A}}_2$, etc. In Fig. 4.19 the passage between two sub-minimizations is well visible and corresponds to the jump of the error. In the algorithm of constraints post-implementation, Algorithm 5, this jump is related to the normalization of vectors in $\tilde{\mathbf{V}}$ and \mathbf{N} matrices. However, inside of each of

sub-minimizations we can observe rapid error convergence. On the basis of these observations, we choose the limited number of internal iterations equals to 5, but it is possible to search another internal stopping criteria which potentially could be efficient for more rapid global algorithm convergence.

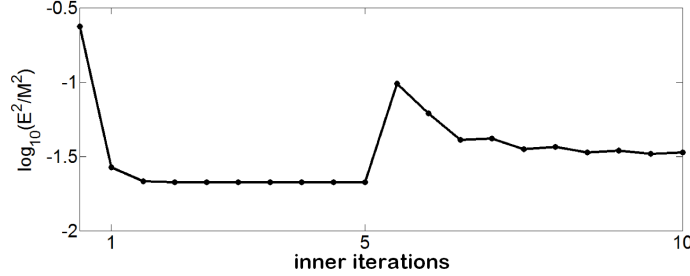


FIGURE 4.19: Investigation of internal stopping criteria.

Let us compare applicability of these criteria for the task of surface reconstruction by the proposed alternating scheme, Table 4.1. We will test only different global stopping criteria, Fig. 4.11(a), remaining a number of internal cycles equals to 5 for each sub-minimization.

Results of reconstruction of two different surfaces are presented in Table 4.1 for three stopping criteria: 10 algorithm global iterations, analysis of the model error for the k^{th} iteration calculated as

$$Err = \frac{\left\| \mathbf{M} - f(\tilde{\mathbf{A}}_k, \tilde{\mathbf{L}}_k, \tilde{\mathbf{V}}_k, \mathbf{N}_k) \right\|_2^2}{\left\| \mathbf{M} \right\|_2^2} \quad (4.52)$$

and the normals variation also for the k^{th} iteration calculated as

$$Var = \frac{1}{3P} \left\| \mathbf{N}_k - \mathbf{N}_{k-1} \right\|_2^2, \quad (4.53)$$

All the stopping criteria turn out to be efficient and the resulting surfaces for all of them are comparable, it is indicated by the evaluated \mathbf{z}_{RMSE} in Table 4.1, which does not significantly change for different criteria. In Fig. 4.20 there are evolutions of Err and Var variables for 10 iterations of the global loop. The results presented in Table 4.1 correspond to the errors thresholds $Err = 0.03$ and $Var = 0.01$ in Fig. 4.20(a) and Fig. 4.20(b), respectively. The pre-defined errors for model (Err) and normals (Var) have been achieved before the 10th algorithm iteration. The normals variations Var are similar for two presented examples. This measure converges after the first iteration and does not significantly change after that. At the same





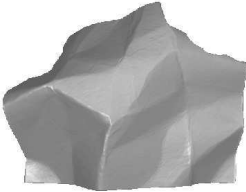
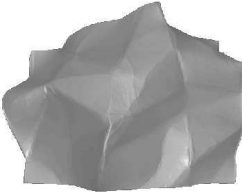
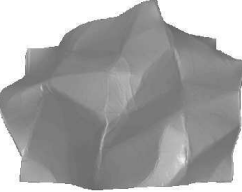
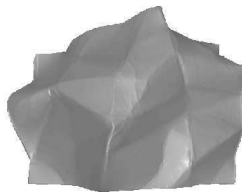
GT surface	Criterion #1 $it_{max} = 10$	Criterion #2 $Err < 0.03$	Criterion #3 $Var < 0.01$
 $334 \times 467 \times (3 \times 10)$	 $t = 22.2s$ $\mathbf{z}_{RMSE} = 0.1461$	 $t = 10.6s$ $\mathbf{z}_{RMSE} = 0.1477$ $it = 5$	 $t = 9.74s$ $\mathbf{z}_{RMSE} = 0.1477$ $it = 1$
 $467 \times 600 \times (3 \times 10)$	 $t = 50.2s$ $\mathbf{z}_{RMSE} = 0.1536$	 $t = 22.9s$ $\mathbf{z}_{RMSE} = 0.1536$ $it = 1$	 $t = 22.5s$ $\mathbf{z}_{RMSE} = 0.1536$ $it = 2$

TABLE 4.1: Global stopping criteria comparison.

time, the model errors Err converge to different values, which creates a problem of the universal Err threshold.

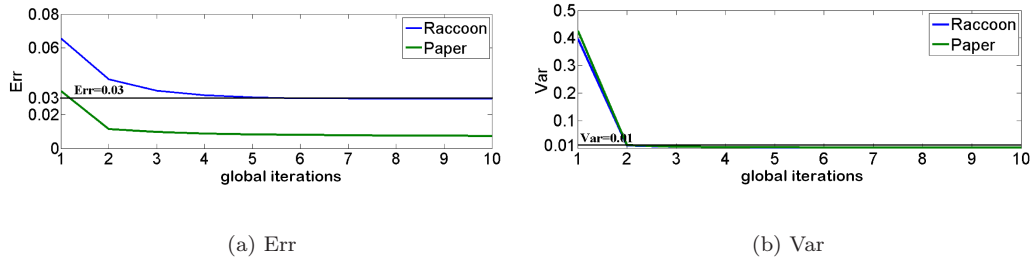


FIGURE 4.20: Evolution of model and normals error for 10 iterations.

The simplest and the most universal criterion which can be chosen is the limited number of iterations. Exactly this choice was made for all the following experiments. A number of iterations in global loop is equal to 10, and the number of iterations in each of sub-minimizations is 5. For these values we obtain an algorithm convergence, and the quality of the resulting surfaces satisfactory for further analysis. However, in perspectives of this work an automatic choice for each of three stopping criteria is envisaged.

4.2.5 Final reconstruction algorithm summary

To conclude this section, let us present the details of the algorithm which we choose to use in all the following experimentations:

- initializations: \mathbf{N} is initialized with 3D normals to the median image intensity values, $\tilde{\mathbf{L}}$ is chosen as maximum intensity value for each image, $\tilde{\mathbf{A}}$ is a unit matrix;
- value and quality constraints are included to the evaluation using the algorithm of post-implementation, Algorithm 4;
- number of internal iterations for $(\tilde{\mathbf{V}}, \mathbf{N})$ and $(\tilde{\mathbf{L}}, \tilde{\mathbf{A}})$ evaluations equals to 5;
- global stopping criterion is the number of global iterations equals to 10;
- resulting depth map is calculated using Basri, Jacobs and Kemelmacher algorithm with the implementation presented in Algorithm 6 of Appendix A.

The number of initial images will change from one experiment to another depending on the purpose of the experiment and available data.

4.3 Lambertian objects reconstruction

This section is dedicated to the presentation of the reconstruction results for different types of Lambertian objects. We also demonstrate the ability of the proposed algorithm to separate information on the surface and light sources, and investigate the influence of the input images on the reconstruction results. Our experiments are made on objects from the MIT Intrinsic Images dataset, [111]. This database contains artificially generated images and the ground truth surfaces used for these images creation as well. Thus, we can compare the results of surface reconstruction and the ground truth surfaces. For all the experiments we use the algorithm configuration presented in the end of Section 4.2. The reconstructed surface is presented in false colors in order to illustrate general object shape (according to the colormap of Fig. 4.3 used in all the examples of this chapter), and in gray-scale mode in order to illustrate the surface details.

4.3.1 Separation of information

Matrix factorization procedure consists in separating information present in data matrix (matrix of vectorized images \mathbf{M}) into several other matrices ($\tilde{\mathbf{A}}$, $\tilde{\mathbf{L}}$, $\tilde{\mathbf{V}}$ and \mathbf{N} for our model). In order to verify how this information is distributed between different variables and how this distribution evolves from one iteration to another let us have the following experiment. There are three input color images of paper, Fig. 4.21(a), 4.21(b) and 4.21(c).

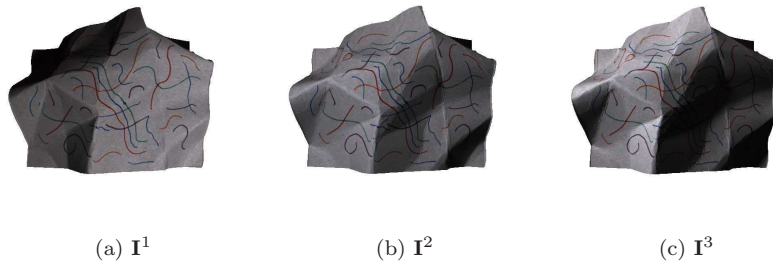


FIGURE 4.21: Three input RGB images of folded paper used for the surface reconstruction.

These three images are used as the input set for the proposed photometric stereo approach. The only difference of the configuration from Section 4.2 consists in increasing of the number of global iterations to 20 in order to observe the algorithm behavior after the 10th. All other details of the reconstruction algorithm configuration and parameters are preserved.

The evolution of surface \mathbf{Z} for the 0th, 1st, 10th and 20th iterations is presented in Fig. 4.22(a) - 4.22(d). In Fig. 4.22(f) - 4.22(i) we present the evolution of albedo values for the the same iterations. It can be seen that for the chosen initialization of the matrix \mathbf{N} (3D normals to median image intensities values) all albedo information is initially contained in the matrix of normals. After that, from one iteration to another we obtain more and more separation. Information about surface color migrates to the matrix $\tilde{\mathbf{A}}$ and information corresponding to the object surface is contained in the matrix \mathbf{N} . The difference of surfaces and albedo values matrix between the 10th and the 20th iterations is not significant. This fact is confirmed by a small model error changing between these two iterations, Fig. 4.23(a). There is also the ground truth surface and albedo for this images sequence, Fig. 4.22(e) and Fig. 4.22(j) respectively, thus we can make a visual comparison of the resulting and the ground truth surfaces, as well as the resulting and the ground truth albedo

values. The evolution of the z_{RMSE} measure for 20 iterations of the reconstruction algorithm is shown in Fig. 4.23(b). It should be mentioned, that the found surface is not exactly similar to the ground truth one. Mostly it is related to the fact, that the shading of three input images does not contain complete information on the object surface. In the next section we discuss, how the information present in the input images influences the reconstruction results.

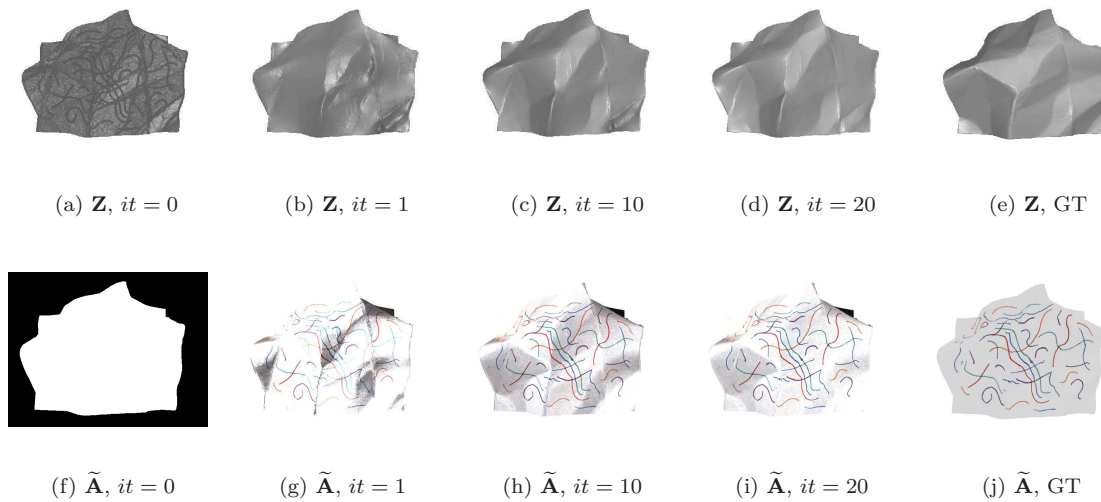


FIGURE 4.22: Demonstration of information separation into albedo and normals matrices for images from Fig. 4.21 for the 0^{th} , 1^{st} , 10^{th} and 20^{th} iterations.

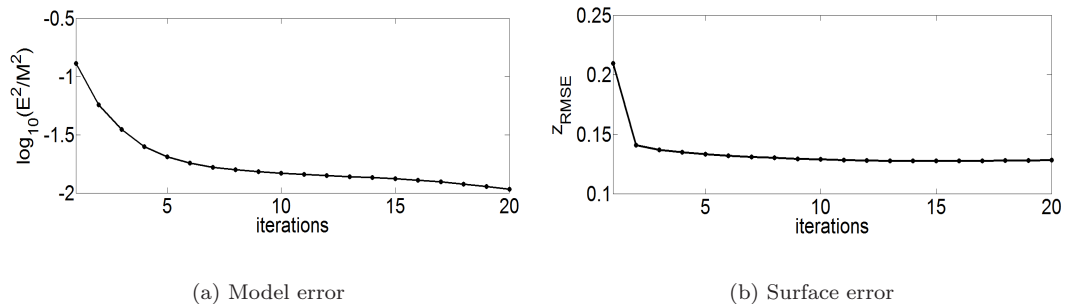


FIGURE 4.23: Convergence of errors for the separation of information experiment.

4.3.2 Influence of the spatial light sources distribution on reconstruction

The objective of the two following sections is to show the influence of the input images on the resulting surface. There are two characteristics of the input images

which can be important for the reconstruction: spatial distribution of light sources used during acquisitions and the number of input images.

First of all, let us investigate the influence of the light sources spatial distribution on the quality of surface reconstruction. The minimum number of images with a fixed camera and varying illumination which can be used for the photometric stereo equals to three. This minimum allows to solve the problem mathematically. However, depending on the distribution of the light sources used for the images acquisition, we obtain images containing different information on the object surface. If all the three necessary images are taken with spatially close light sources, the information on the different object parts can be unbalanced. Thus, the more illuminated part of the object will be better reconstructed than others.

In order to show this effect, we propose to use the example of a square box whose 10 initial images are presented in Fig. 4.24. From this set of images we form two sub-sets with poorly and well distributed light sources: $\{\mathbf{I}_{poorD}\} = \{\mathbf{I}^2, \mathbf{I}^3, \mathbf{I}^4, \mathbf{I}^5\}$ and $\{\mathbf{I}_{wellD}\} = \{\mathbf{I}^2, \mathbf{I}^6, \mathbf{I}^8, \mathbf{I}^{10}\}$. The set $\{\mathbf{I}_{poorD}\}$ contains images with the object illuminated only from one side, and in $\{\mathbf{I}_{wellD}\}$ there are images with the object illuminated from all sides.

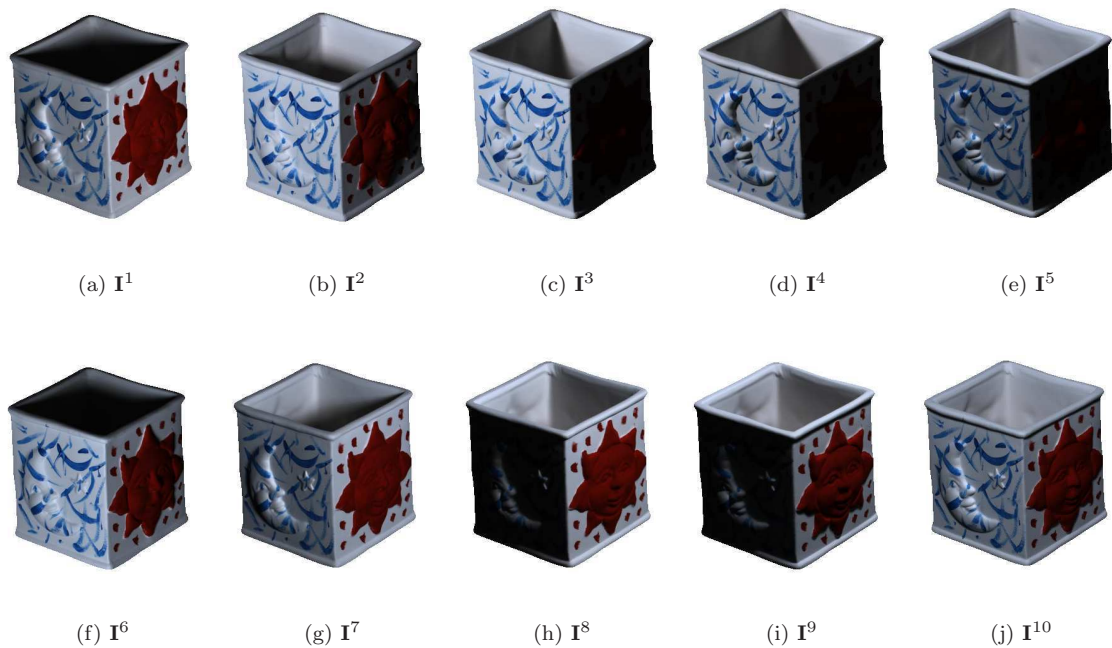


FIGURE 4.24: Images for the demonstration of the input data influence

After that, we reconstruct the object surface separately using $\{\mathbf{I}_{poorD}\}$ and $\{\mathbf{I}_{wellD}\}$. For these reconstruction we also exploit algorithm resumed at the of Section 4.2.4.

The results of reconstruction as well as the ground truth surface are shown in Fig. 4.25. For the images with the one-side illumination, the reconstruction of non-illuminated side details is not realized efficiently, Fig. 4.25(b) and Fig. 4.25(e). For the images with well distributed light sources the surface details are reconstructed for all object parts, Fig. 4.25(c) and Fig. 4.25(f), this is especially well visible in the gray-scale representation. However, only four images even with well distributed light sources is not sufficient to describe the presented object general shape which is visible from the false color surface representation (Fig. 4.25(c) comparing to Fig. 4.25(a)). Not only the distribution of the light source is important but also their total number.

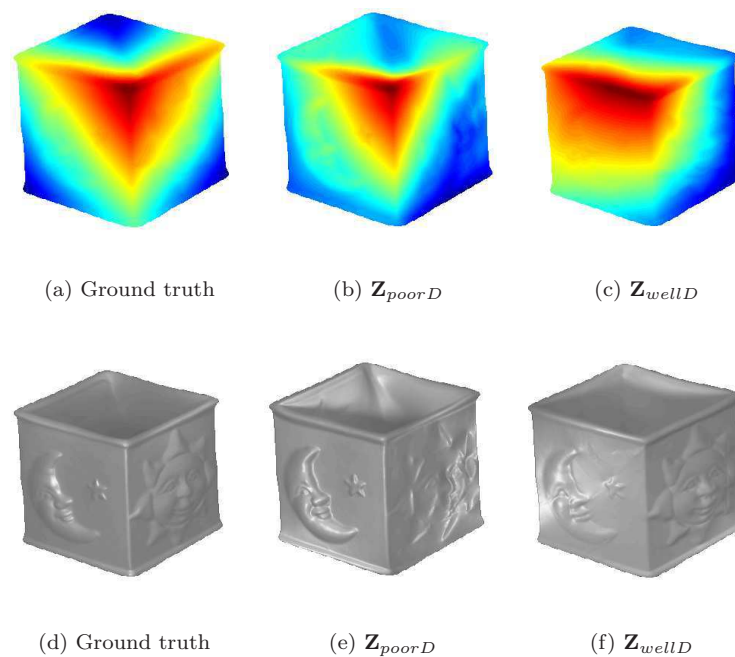


FIGURE 4.25: Influence of the light sources distribution.

4.3.3 Influence of the light sources number on reconstruction

Here we aim at showing reconstruction results for the different numbers of images with well distributed light sources. Let us have two sub-sets of the input images taken with different number of spatially well distributed light sources, $\{\mathbf{I}_4\}$ and $\{\mathbf{I}_{10}\}$, formed from the images of Fig. 4.24. $\{\mathbf{I}_4\}$ is chosen the same as $\{\mathbf{I}_{wellD}\}$, and $\{\mathbf{I}_{10}\}$ contains all the images of the input sequence. The results of reconstruction for the both sub-sets are demonstrated in Fig. 4.26. The surface obtained for 10

well distributed light sources is of higher quality than the surface obtained for 4, also well distributed, light sources. It can be concluded that the more images with well distributed light sources we have in the initial sequence, the more information on the object surface is contained in the input data, and thus, the more qualitative surface can be obtained using the proposed reconstruction approach.

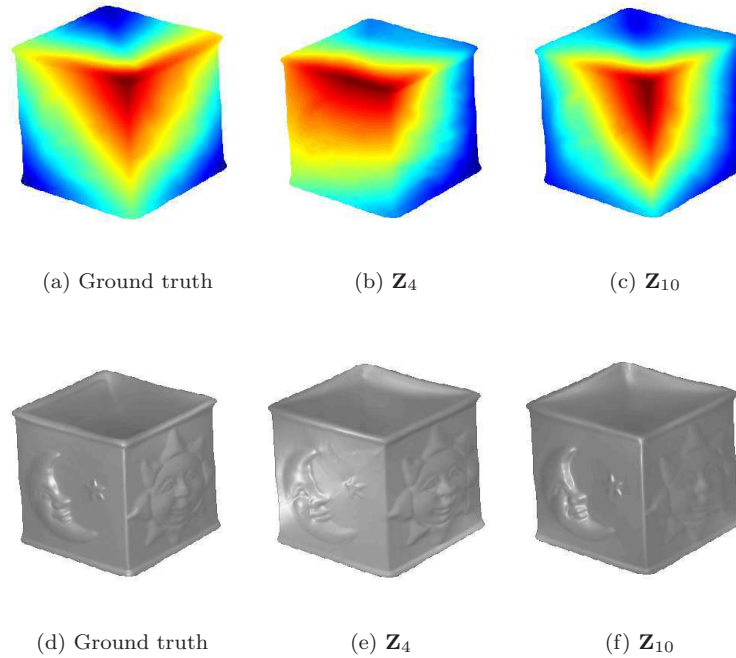


FIGURE 4.26: Influence of the light sources number.

Another interesting observation that can be extracted from the conducted experiment concerns the algorithm convergence. In Fig. 4.27(a) there are three curves of the evaluated model error for 10 iterations for different initial images: taken with 4 poorly, 4 well and 10 well distributed light sources. The algorithm of reconstruction converges for all the three input sequences. Due to the fact, that for 10 images the evaluated model components should take into account more diverse information from all 10 images (system for resolving consists of more equations and contains more constraints), the final model error is more elevated than for the cases of 4 input images. At the same time, the surface component extracted from 10 images with well distributes sources is closer to the ground truth and the error of the reconstruction is the smallest, Fig. 4.27(b). Contrary to the expectations, the surface error \mathbf{z}_{RMSE} is more elevated for 4 well, than for 4 poorly distributed sources. This is related to particularities of the error measure. General surface distortion is more taken into account than the small reconstructed details.

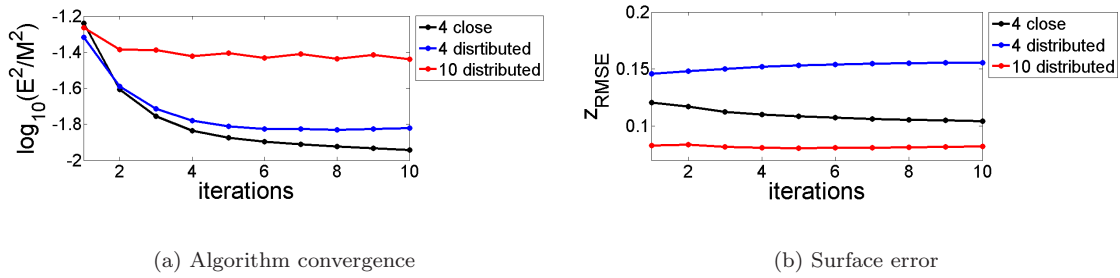


FIGURE 4.27: Algorithm convergence and surface error depending on the input images.

4.4 Non-Lambertian objects reconstruction

The objective of this section is to show the importance of specularities pre-processing for the reconstruction. We apply the reconstruction photometric stereo approach of this chapter in combination with specularities correction proposed in Chapter 3 in order to obtain resulting surface of a specular object.

We can take a familiar example of the specular artificial sphere from Chapter 3, presented in Fig. 3.7. There are eight artificial images created with eight light sources with different spatial positions, uniformly distributed around the photographed object, Fig. 4.28. For these images we know diffuse components, Fig. 4.29. In Fig. 4.30 we present results of correction by the algorithm of soft correction, Algorithm 2, of Chapter 3 with the following parameters: $\tau = 1.2$, $\alpha = 5$, $k = 0.9$, and using mean for ratio values aggregation.

First of all, let us analyze the results of soft specularity correction. For the chosen set of the input images, we do not obtain a total correction, however the specular zones are significantly reduced in intensity values. Moreover, thanks to the continuous correcting function, there is no sharp passages between the corrected and non-corrected zones. Images particularities of shading are preserved, because looking at the corrected images we are still able to define, from which side they were initially illuminated.

The influence of the specularities correction on the surface reconstruction results by the Algorithm 4 of this chapter is demonstrated in the following experiment. The surfaces in this experiment are presented by level contours of the evaluated depth map. With such a visualization we can best assess general surface deformation. In Fig. 4.31(b) there are contours of the surface reconstructed on the basis of images

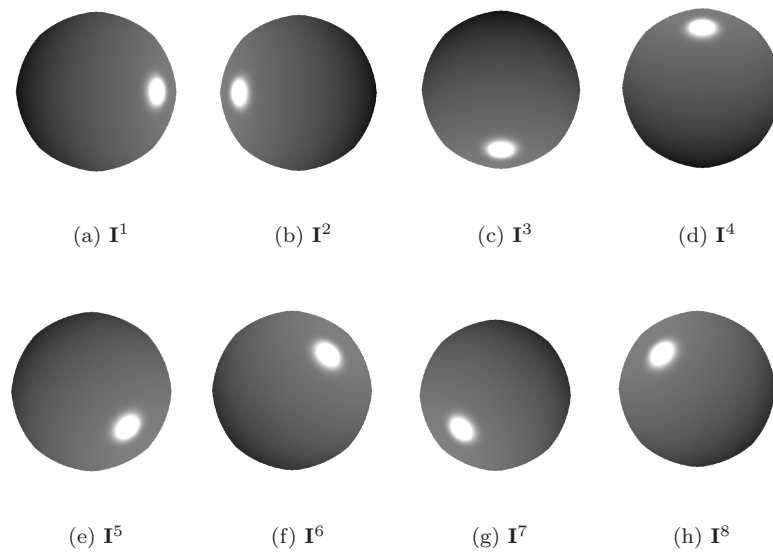


FIGURE 4.28: Artificial specular sphere images.

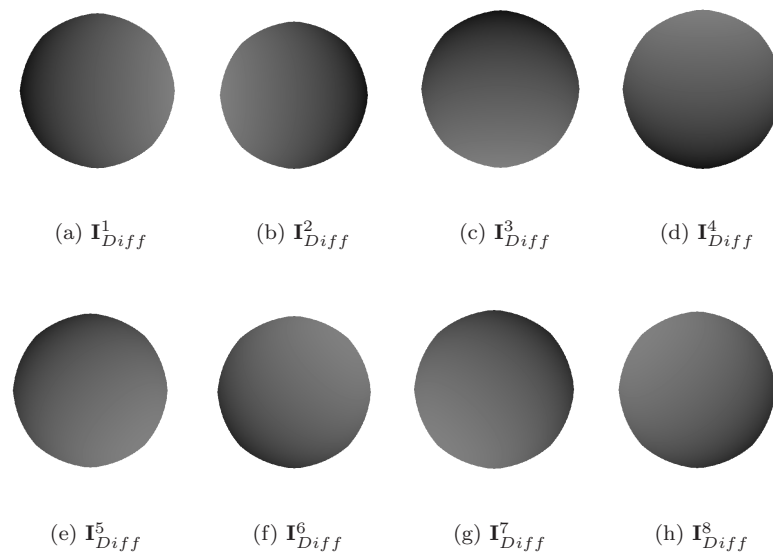


FIGURE 4.29: Diffuse component of the artificial specular sphere images.

without any pre-processing. The surface is deformed and we cannot recognize circular shape appropriate to the initial sphere, Fig. 4.31(a), which can be seen for the reconstruction obtained on the basis of diffuse image components, Fig. 4.31(c). After the proposed correction for images of Fig. 4.30, we obtain contours presented in Fig. 4.31(d), which are close to the sought contours of Fig. 4.31(a).

The phase of specularities correction is important and required for images of non-Lambertian objects. Thanks to this phase, the obtained surfaces do not suffer from

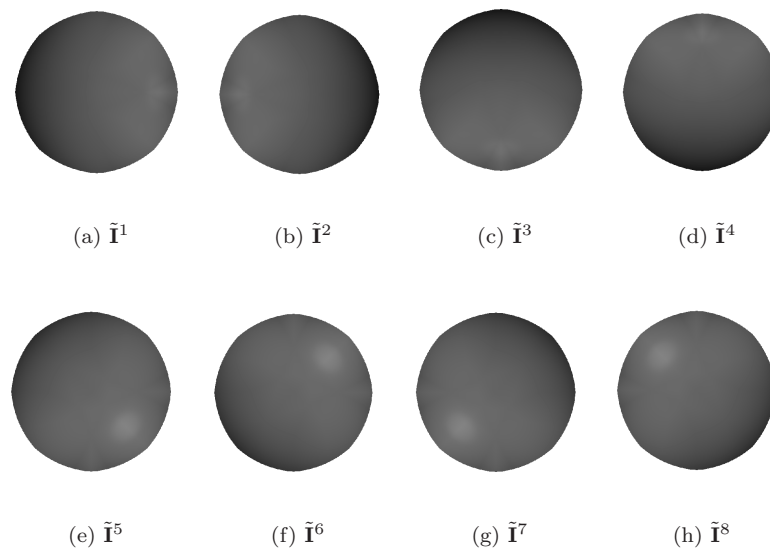


FIGURE 4.30: Corrected artificial specular sphere images.

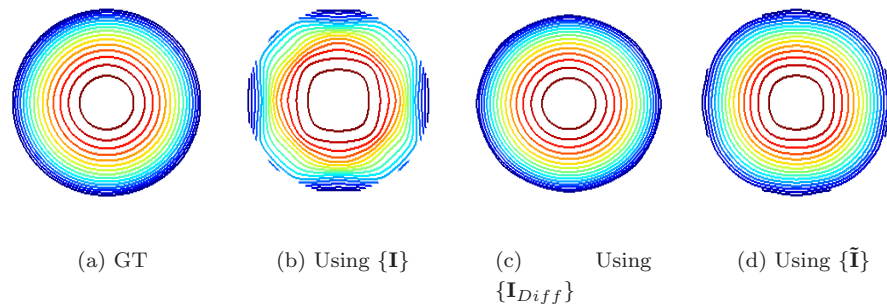


FIGURE 4.31: Reconstructed surface of sphere.

distortions and can be more efficiently analyzed visually and used automatic recognition algorithms.

4.5 Conclusions

A novel iterative approach of the Lambertian uncalibrated photometric stereo reconstruction is proposed. This technique consists in separation of the lighting sources (intensities and directions) and the object surface (albedo and normals) information by the means of the matrix factorization and the alternating optimization principle. All the model components may be divided into two groups: directional (normals and light sources directions) and weighting (albedo and light sources intensities). The resolution is presented as an iterative finding of model components from each

of these groups. Value and quality constraints implementation gives the possibility to include model variables particularities and regularization component to the minimization procedure. The proposed algorithm is completely 'data-driven', it does not need any supplementary knowledge a priori. Even model components initialization is based on the input images. The efficiency of the proposed algorithm (resumed in Section 4.2.5) is demonstrated for artificial images.

We also demonstrate the efficiency of the proposed algorithm for non-Lambertian artificial object reconstruction in combination with the algorithm of specularities correction of Chapter 3. The specularities removal permits to avoid the problem of surface deformation which is typical for the specular objects. Examples of the complete chain application (specularities removal and surface reconstruction) for more sophisticated real-world surfaces are presented in the next applicative chapter.

In this work we do not provide any proves of convergence of the presented algorithm. Such proves are important for alternating optimization algorithms and are in perspectives of our research. Also we did not touch the question of the bas-relief ambiguity, [26], which is very important and can be crucial for some recognition tasks.

Chapter 5

Application to microorganisms surface reconstruction from Petri dish images

This thesis is realized as a part of a collaborative project between CEA LIST and bioMérieux company, [1]. The main purpose of this project consists in automated processing of Petri dish images to provide useful support for decision making in microbiological analysis.

3D surface completes information present in images and is useful for the object characterization. Microorganisms surface reconstruction can also provide microbiologist with necessary knowledge. In this chapter we propose to apply the approach proposed in the previous chapters to the images of Petri dish. Indeed, images are taken using a special imaging device developed by bioMérieux company, [2]. During images acquisition, complex illuminating sources which are supposed to have unknown directions and intensities are used. Thus, the proposed algorithm of uncalibrated non-Lambertian surface reconstruction is applicable to this type of data.

5.1 Work motivations

Microbiology is a science which studies microorganisms of different nature: from small unicellular organisms invisible with the naked eye, such as viruses, bacteria, archaei or protozoa, to multicellular ones which can be seen without special optical

devices, such as algae or fungi. Bacteriology is a microbiology discipline, which main goal is characterization and classification of bacterial species. The first direct demonstration of the role of bacteria in provoking diseases came from the study of German physician Robert Koch (1843-1910), [116]. He related the presence of *Bacillus anthracis* with the emergence of anthrax.

One possibility to observe and study bacteria is to make them grow inside a special glass or plastic dish, called a Petri dish from its inventor's name, Julius Richard Petri (1852-1921). Soon after Petri dishes were presented to the scientific world, they became (and they still remain) a popular instrument to study microorganisms.

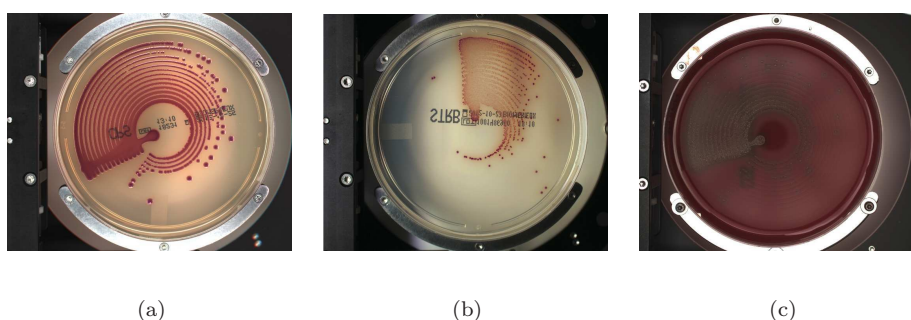


FIGURE 5.1: Examples of Petri dishes with different media and grown colonies: (a) *Escherichia coli* in CPS medium, (b) *Streptococcus agalactiae* in STRB medium, (c) *Streptococcus pneumoniae* in COS medium.

Usually, a Petri dish is filled by a nutritious solidified medium which allows, facilitates or controls the growth of bacteria. Depending on the sample to be inoculated into the dish and on expected nature of microorganisms, a given medium will be used with a suitable composition and properties (e.g., PH, nutrients, antibiotics, temperature, viscosity, etc.). After that, inoculated dishes are placed into special physical conditions depending on the microorganisms growing inside, in order to achieve desired development. When conditions of growth are favorable, there can be thousands of bacteria inside a Petri dish after a few hours of incubation. These bacteria can form colonies which are well visible without any optical support. In Fig. 5.1 there are several examples of Petri dishes with different media and microbial colonies. The diversity of possible nutritious media and colonies types is restricted but yet it remains very vast. From year to year this domain evolves: new media are invented by biochemists and new colonies are revealed by microbiologists.

In laboratories, microbial colonies are analyzed by specialists in order to diagnose different diseases of the respiratory and nervous systems or skin infections (e.g., peritonitis, coleitis, prostatitis provoked by *E.coli* depending on the studied samples).

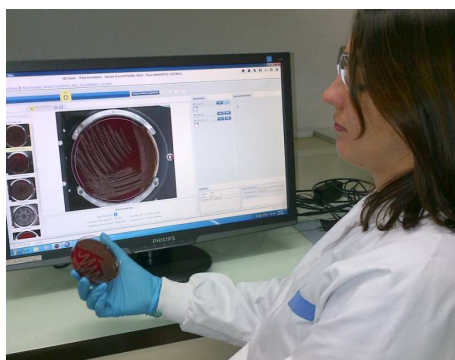


FIGURE 5.2: Visual analysis of a Petri dish with microbial colonies.

The earlier these diseases are detected the more efficiently they can be treated. Colonies analysis is also widely used for industrial purposes to control air quality or any other product containing bacteria (e.g., dairy product). Analysis of colonies inside a Petri dish is realized visually by a laboratory specialized technician. Naturally, a technician may turn the Petri dish in order to see the colony under different incident light directions, Fig. 5.2. Such a manual procedure enables colony characterization in terms of color, shape, type of contours and even 3D surface. Accurately characterized colonies can be further recognized by means of existing protocols, biological bibliography and technician's own experience.

BioMérieux company, [1], has developed a dedicated imaging device, [2], aiming at facilitating microbiological analysis. This device is capable of taking a sequence of images of Petri dishes under varying lighting conditions. The structure of this device will be further detailed in this chapter. A camera and different illuminations allow imitating the Petri dish rotation instigated by a human technician. Moreover, algorithmic solutions are investigated so as to provide useful support for decision making.

By increasing the device functionality, automation of a series of essential microbiological tasks becomes possible and applying suitable image processing and machine learning techniques can enable to achieve the desired results. One of the main goals of bacteriology is growth characterization and identification. Colony characterization is made on the basis of different growth features such as color, size, shape, contour line, shining, surface elevation type, etc. Some of these properties can be directly extracted from the image, while other need special pre-processing procedures and algorithmic solutions. Depending on applicative microbiological task, some features will be of higher interest than others.

In this thesis part we are interested in reconstruction of 3D surface of microbial colonies. Indeed, microbial colony surface information appears to be relevant in many colonies analysis issues. An object surface is a property which can be hardly extracted directly from an image. However, a presence of several images enables the extraction of surface using photometric stereo technique. Surface of microbial colonies can be of Lambertian or non-Lambertian nature. We can obtain them by the proposed method for uncalibrated photometric stereo surface reconstruction of non-Lambertian objects.

5.2 Imaging device

In order to facilitate Petri dishes analysis performed by microbiologists, bioMérieux company is proposing to use a specific imaging device, [2], which is able to capture high-resolution (1 pixel \approx 0.05 mm) Petri dish images with a fixed position of the Charge-Coupled Device (CCD) camera and varying lighting directions. Images taken under the described conditions are ideally suited for photometric stereo reconstruction. In this section we will present the acquisition device so as to better understand input images used for the applicative task of microbial colonies surface reconstruction.

The imaging device is a closed cupboard, Fig. 5.3(a), with a special pullout stand for a Petri dish. Petri dishes of different diameters (90 cm, 60 cm, 55 cm, etc.) can be placed into the stand. Once the Petri dish is placed, the stand is pushed inside of the cupboard and images can be taken. The closed space created by the cupboard does not allow the external light in. This structure hence allows obtaining images which are not affected by ambient light.

Objects inside the cupboard are illuminated by lighting sources which are positioned on four opposite sides of the photographed object (left, right, front and back) and at two different height levels above the object (bottom annular and top annular). The CCD camera is placed in a special hole above the bars of diodes and perpendicular to the plane of the photographed dish. The internal device structure can be seen in Fig. 5.3(b) and a 3D representation is provided in Fig. 5.3(c).

The details of the lighting sources dispositions inside of the device are presented in Fig. 5.4(a). As already mentioned, two height levels of lighting sources are available with four opposite directions of lighting for each level. The illumination of the

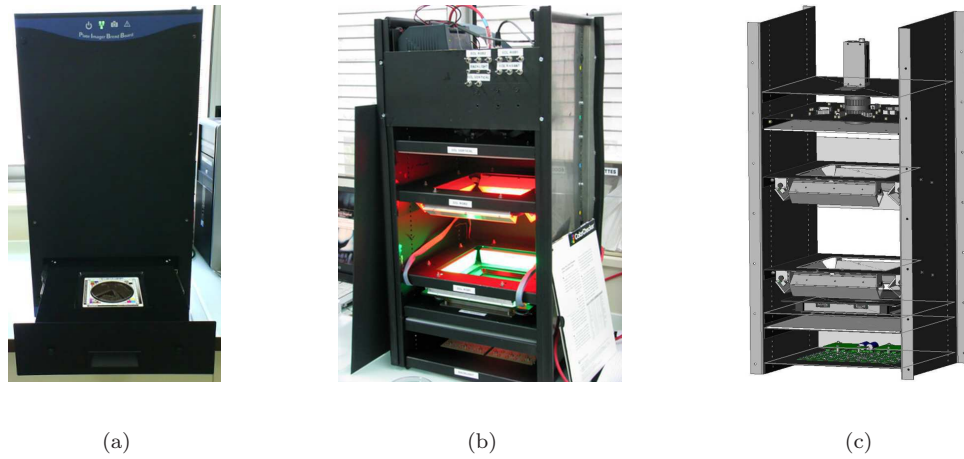


FIGURE 5.3: The bioMérieux imaging device for the Petri dishes: (a) external view, (b) internal view, (c) schematic 3D representation.

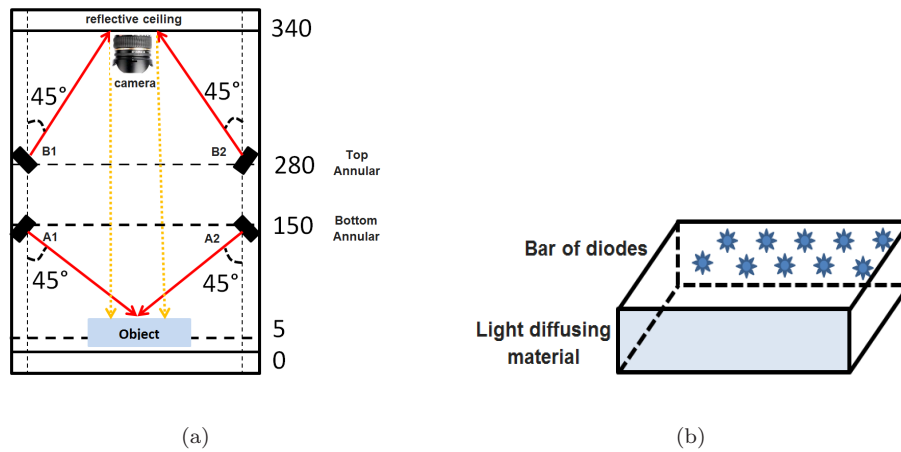


FIGURE 5.4: The bioMérieux imaging device details: (a) dispositions of the lighting sources, the CCD camera and the object of interest, (b) the bar of diodes.

lower level (bottom annular) is direct, unlike the top annular illumination. The top annular lighting is obtained thanks to the reflection of the light from the reflective ceiling made with a special material. Therefore, the supplementary diffusion of light related to these reflections should be also included into image formation modeling process. The proposed photometric stereo approach contains a hypothesis of direct illumination. That is why in our applicative examples of the microbial colonies reconstruction we use only four images with the bottom annular illuminations.

Previously, to define lighting sources the term 'bar of diodes' was used. We will now detail the structure of these bars, presented in Fig. 5.4(b), taking into account that they are the same for all directions of the bottom and top annular illuminations. It

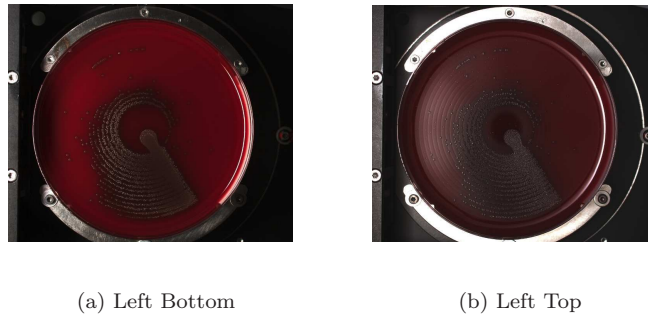


FIGURE 5.5: Examples of images with illumination of the two height levels from the left side.

is a volume bar which consists of back internal side with 10 RGB LEDs placed in form of zigzag. The front side of the bar is a light diffusing plastic slab which permits to blur directed diodes illumination and obtain soft side lighting. Each diode emits light with narrow bands of wavelengths corresponding to different lighting colors. The final RGB image, Fig. 5.5(a), is obtained as an aggregation of these three color channels captured by the CCD with three different wavelengths, see Fig. 5.6.

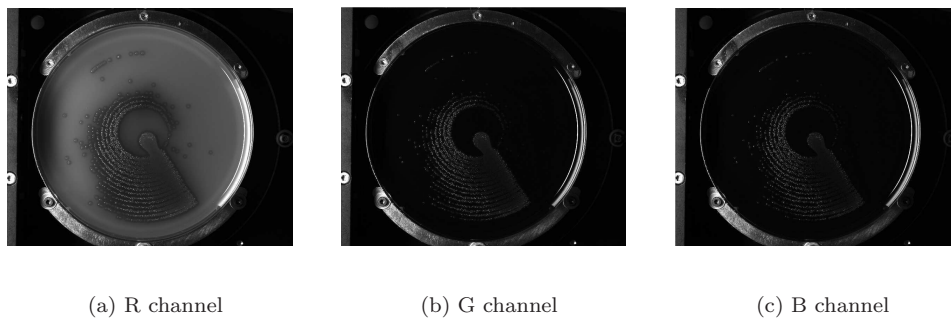


FIGURE 5.6: Channels of the RGB image taken with the left side bottom annular illumination.

Based on the description above, it can be concluded, that even if the Petri dish imaging device captures photographs which are suitable for a further photometric stereo procedure, the conditions of illumination used during acquisitions are very complex. Moreover lighting directions are not always well calibrated and the imaging system can evolve over time. Properties of the diffusing plastic slab, Fig. 5.4(b), are hard to model and can vary according to the mode of the device exploitation (e.g., these properties can vary depending on the temperature of the slab which is increased for a prolonged usage of the device). Thereby hypothesis of unknown illumination conditions for the photometric stereo problem used in our proposed algorithm, Chapter 4, is indispensable for this case. In addition, the objects inside

of the Petri dish are often specular and of unknown properties. Thus, the proposed method of reconstruction presented in Chapter 3 and Chapter 4 can be efficiently used for the presented kind of data.

5.3 Microbial colonies surface reconstruction

In this section we demonstrate several reconstruction examples using the proposed methods from Chapter 3 and Chapter 4. The variety of possible nutritious media and microorganisms in Petri dishes, is very large. In Fig. 5.7 we present several images of Petri dishes with different media and grown colonies, taken by the described device with left bottom illumination. The diameter of these dishes is 90 mm. The proposed method of uncalibrated non-Lambertian surface reconstruction should be able to cope with this diversity of the objects.

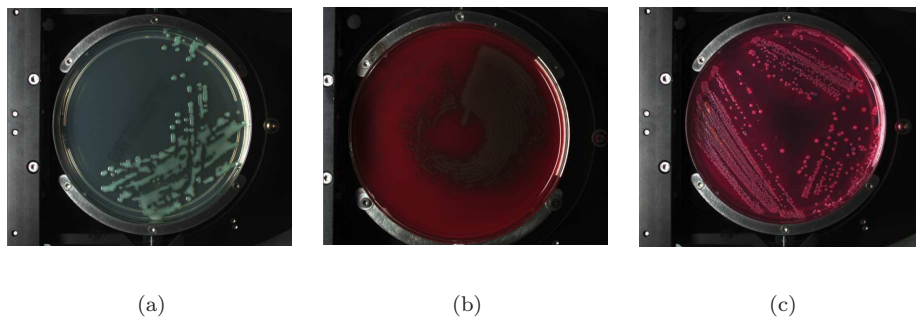


FIGURE 5.7: Diversity of media and microbial colonies.

5.3.1 Global Petri dish reconstruction

The surface reconstruction can be realized for the whole Petri dish or for a chosen region of the dish which is of interest for analysis. Let us present an example of the global reconstruction for the images of Fig. 5.7(c) by the proposed reconstruction chain. The presented reconstruction, Fig. 5.8(e), is made on the basis of four images with bottom illuminations from four different sides, Fig. 5.8(a) - 5.8(d).

Such a global reconstruction can be an interesting solution for the dish visualization in order to provide supplementary information on relative surface elevations, shape, etc. However, during the model component evaluation for the whole dish, we violate

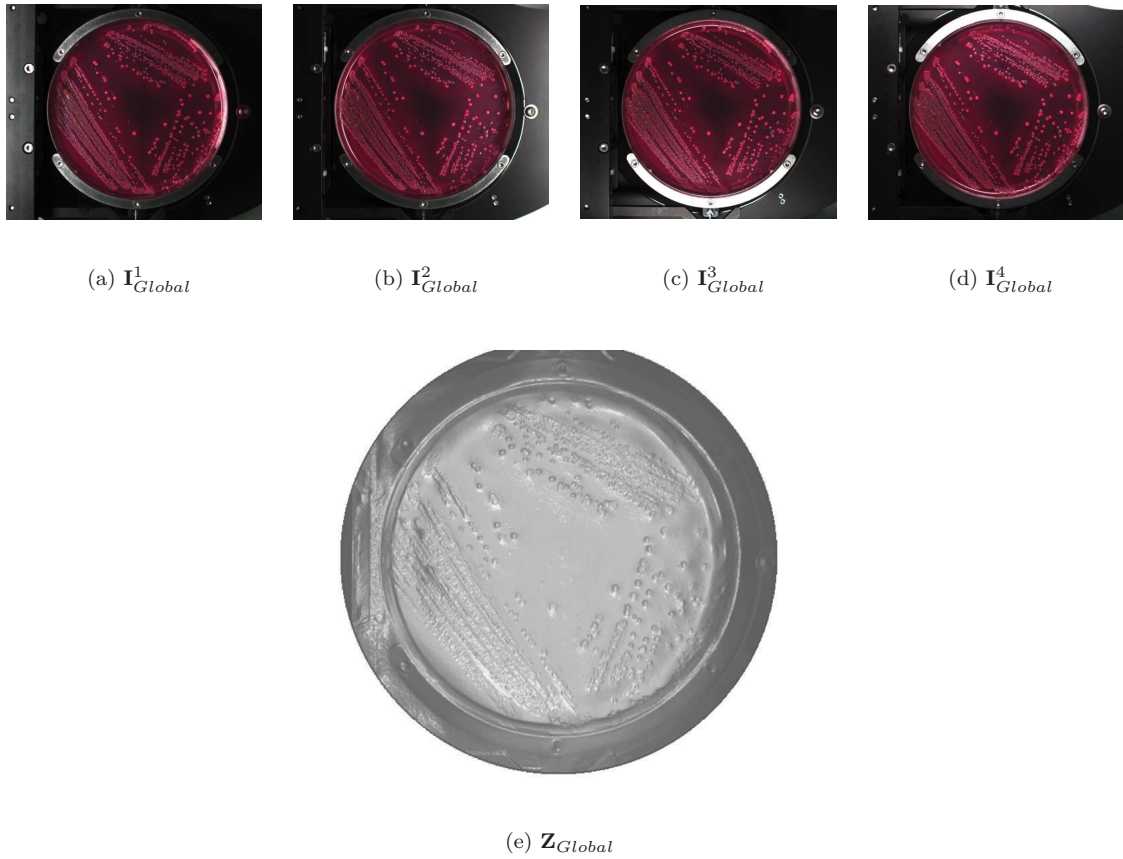


FIGURE 5.8: Global Petri dish reconstruction.

the condition of point light source. As a result, we obtain a distortion of the evaluated surface. When the object is small, the reconstruction procedure is not so sensible to this nonconforming, because the distance from the reconstructed object to the light source comparing to the size of the object itself is significant. The larger zone of the Petri dish image is used for the reconstruction, the more the hypothesis on the point light source is violated.

Moreover, images obtained in the utilized device is of a large size ($[2050 \times 2448 \times 3]$ each color image). The global dish reconstruction even with a re-scaling factor 0.25 takes 4 minutes. This computation time is not acceptable for the industrial system.

5.3.2 Microbial colonies surface reconstruction

Nevertheless, for automatic colonies analysis we are more interested in the reconstruction of only certain objects inside the dish. Frequently, microbiologists concentrate their attention on isolated microbial colonies characterization. The object of

interest can be chosen by a user on the basis of one image of the input sequence or automatically detected by a specific algorithm.

Let us present the processing chain which is used in all the experiments of this chapter, Fig 5.9. This chain consists of three principal phases: images acquisition, selection of the object of interest and evaluation of its mask and, finally, uncalibrated non-Lambertian photometric stereo reconstruction.

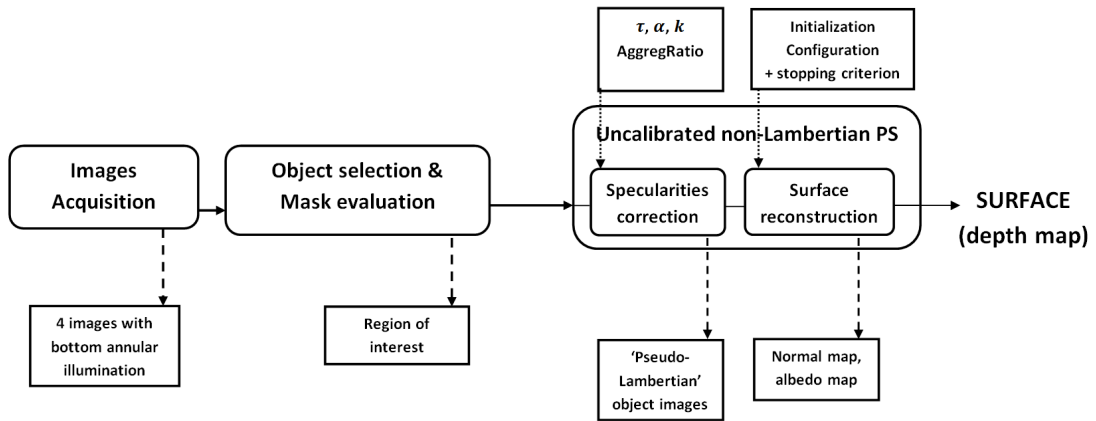


FIGURE 5.9: Processing chain for microorganisms surface reconstruction.

For example, in Fig. 5.10(a) there is one of the images of Fig. 5.3.1 with an additional rectangle inside. This rectangle is defined by the user as an object of interest which requires the reconstruction procedure by the proposed complete chain of processing. The zoomed image region to be processed for four images of the input sequence (bottom illumination from four different sides) is also demonstrated in Fig. 5.10.

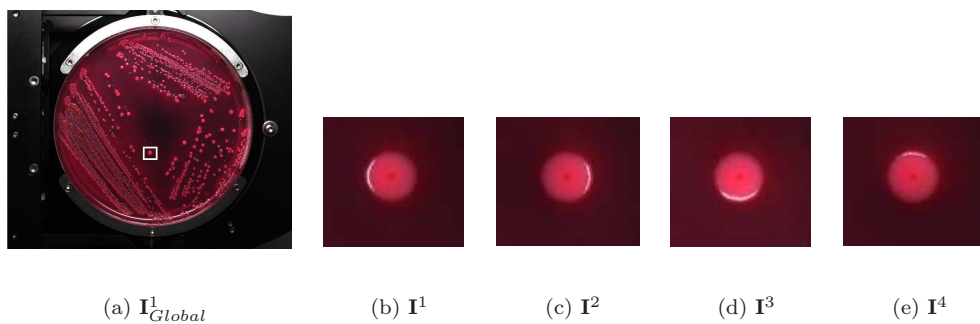


FIGURE 5.10: The chosen image region inside the Petri dish to be processed.

We do not detail here the algorithm of the object mask evaluation in order to distinguish the selected colony from the medium. This information is provided as a part of the input data for our reconstruction chain and can be done, for example, using

the algorithm of active contours, [117], or any other algorithm of objects detection. In this chapter we present examples with efficiently detected objects. However, the question of the object mask extraction can also be an important supplementary challenge, because the colonies in the medium can have a low contrast, small size or inhomogeneous coloration. For certain cases, colony can react with the media provoking changing of the color around itself. This creates supplementary difficulties during the automatic colonies detection and definition of its contours. We keep in perspectives of the work a universal algorithm of the object mask detection which can manage all these data particularities. In Fig. 5.11 there is an example of the chosen image region and the microbial colony object detected using the active contours approach.

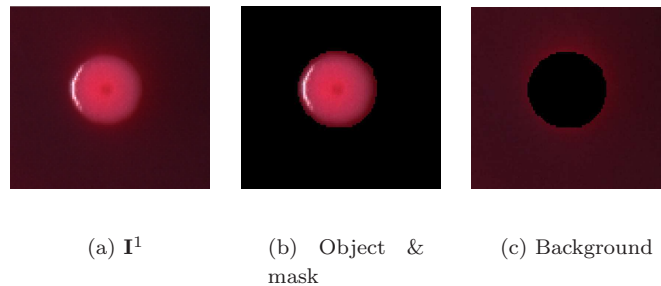


FIGURE 5.11: Image region with masked colony and masked background.

The first stage of the uncalibrated non-Lambertian photometric stereo, dedicated to the specular zones correction, is pixel-wise. Thus, the mask application decreases the computation time. For the second stage of the surface reconstruction, the calculations are matrix. So, the presence of the background pixels which are considered as the pixels of the object and taken into account for the model components evaluation, can affect the reconstruction results. In all the following examples we suppose to already have the object mask before the photometric stereo reconstruction.

On the basis of the obtained mask and four images with different illuminations we are able to apply the proposed uncalibrated photometric stereo reconstruction procedure. This procedure is done in two steps, and each of these steps needs its own parametrization. The specularities correction procedure needs parameters τ , α and k . We choose these parameters value exactly the same as for the global dish reconstruction ($\tau = 1.2$, $\alpha = 100$ and $k = 0.75$). Also the specularities correction algorithm need the type of the aggregation function used for the combination of ratio values; in our examples we used the median. The choice of these parameters

is the same for all the examples of this chapter. These values are not universal, but according to the usage of the same device with the same lighting directions and camera, and relative small size of objects, we can define the value of the parameter $\tau = 1.2$ and equal for all the reconstructions. Two other parameters are not related to the type of the reconstructed object, but the the desired quality of specularities correction.

As a result of the specularities correction algorithm, we obtain images of so-called 'pseudo-Lambertian' microbial colonies which are further reconstructed by the method of Chapter 4. The initialization, the configuration and the stopping criterion are chosen according to the prescriptions of Section 4.2, and are also the same as were used for the global dish reconstruction. On the basis of the calculated normals we reconstruction the depth map using Algorithm 6 of Basri, Jacobs and Kemelmacher.

As it could be seen from the Petri dish images, microbial colonies have their own particularities which can influence the quality of the surface reconstruction: size, inhomogeneity of the surface, large specular zones. First of all, we are going to show the importance of the first stage of specularities correction, presented in Chapter 3, for the complete reconstruction process.

5.3.3 Importance of specularities correction

There is no prior information about the colony surface, thus it can be either close to Lambertian without specular zones, or it can be highly specular and specular zone can occupy up to 20% of the object. By the following experiments we demonstrate the importance of the first phase of correction for the surface reconstruction results.

Let us have four input images of a colony, Fig. 5.12(a) - 5.12(d), as well as their corrected versions obtained using soft correction, Fig. 5.12(f) - 5.12(i). We provide the reconstruction results for both original, Fig. 5.12(e), and corrected, Fig. 5.12(j), images. For the presented reconstruction we used the processing chain of Fig. 5.9. If the proposed algorithm of uncalibrated photometric stereo reconstruction is applied directly to the input images, specular zones, situated on four different sides of the object depending on the illuminating directions, affect the resulting surface, Fig. 5.12(e). Instead of obtaining the hemispherical object, the resulting surface is presented by the parallelogram with rounded edges. After the proposed specularities correction, we overcome these negative effect, Fig. 5.12(j). As a result, the surface

evaluated after specularities correction is close to the sought hemispherical shape. Such type of shape is validated by the visual evaluation of microbiologists.

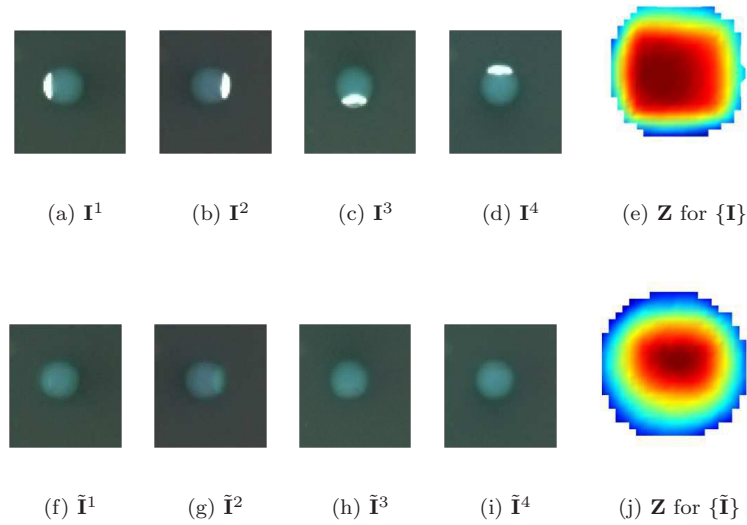


FIGURE 5.12: Partial colony distortion related to specularity.

The consequences of non-corrected specular zone may be even more severe. In Fig. 5.13 we demonstrate another example of reconstruction for the image without and with specularities correction. It can be seen that the optical effect such as specularities provokes important distortion of the object shape.

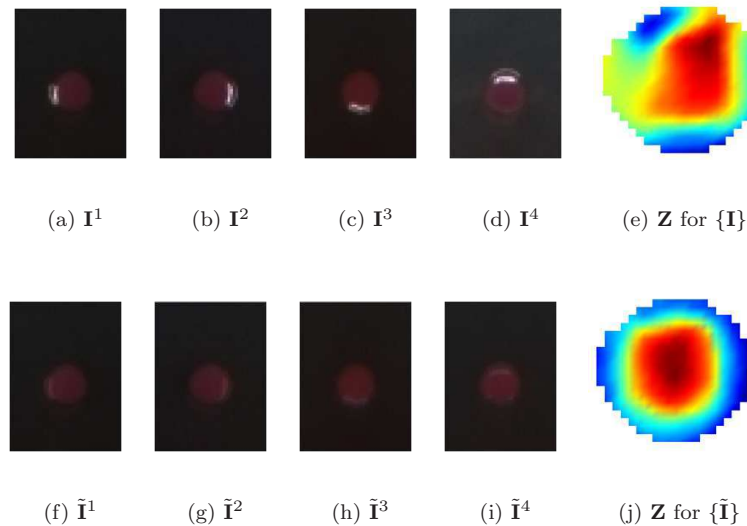


FIGURE 5.13: Important colony distortion related to specularity.

To conclude, specular zones correction is a very important phase of the reconstruction process for the microbial colonies. With the used imaging device and due to the

colonies surface particularities, specularities can occupy a rather large zone. This leads to the reconstructed surface total or partial distortion and makes impossible or false further process of the analysis.

5.3.4 Lambertian colonies surface reconstruction

Nevertheless, some colonies do not have specularities and may be mat and close to Lambertian objects. It is not known a priori if the object has Lambertian or non-Lambertian surface. So, we aim at using the proposed method without providing the algorithm information on the nature of the reconstructed surface. Let us verify if the proposed approach is capable to preserve a qualitative reconstruction even for Lambertian colonies.

In Fig. 5.14 using the same algorithms configuration as for the previous experiments, we present the results of the reconstruction approach application to microbial colony without specularities. In Fig. 5.14(a) - 5.14(d) there are four initial colony images obtained with the bottom illumination from four different sides. At the same time, in Fig. 5.14(f) - 5.14(i) we demonstrate these images after soft specular correction. The algorithm of the specularities correction provides no changes to the resulting corrected images. Consequently, the reconstructions of the surface obtained without, Fig. 5.14(e), and with, Fig. 5.14(j), specularities correction are equivalent.

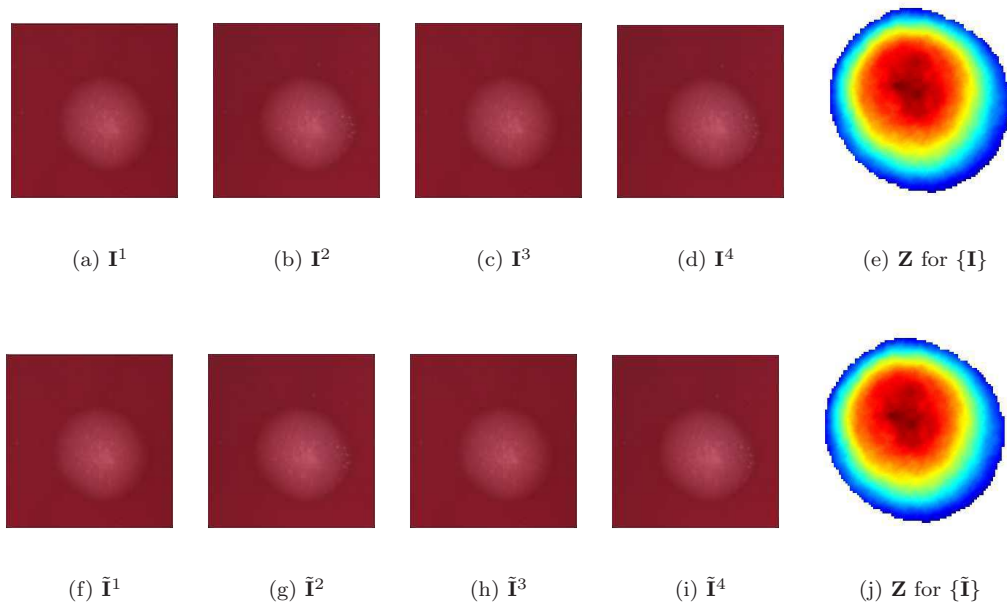


FIGURE 5.14: Lambertian surface reconstruction after specularities correction.

By the numerical experiments with the chosen algorithms parameters, we confirm the qualitative reconstruction of both types of surfaces, Lambertian and non-Lambertian. With the proposed method, it is not necessary to know a priori the type of surface. If it is Lambertian, the specular correction phase does not influence on the uncalibrated photometric stereo reconstruction. On the other hand, if the colony is non-Lambertian, the specular zones are corrected, and only after this correction the whole object is reconstructed.

5.3.5 Examples of different colonies reconstruction

The objective of this section is to demonstrate the reconstruction results for different types of microbial colonies. Due to the diversity of existing colonies elevation types, we demonstrate only a few possible groups of examples. These groups are the following: small objects, objects with concavities and sophisticated shapes. The first two groups are rather widespread in microbiological practice. The last group, dedicated to sophisticated shapes, includes colonies elevation types which are not frequent, but seem to be an interesting challenges for the reconstruction algorithm.

As for all the previous examples of this chapter, we use the proposed processing chain, Fig. 5.9. The results of reconstruction are visualized in a gray-scale mode with top and side views. This permits to better analyze the surface texture and elevation details.

Small objects

Microbial colony evolves in time. Thus, if the images of the Petri dish are taken for different incubation times, the size of the colony will be different. The shorter is the incubation time, the smaller will be the colony. On the other hand, there exist some colony species which rest small (0.5-2 mm) even with significant incubation time. The Petri dish imaging device has a high spatial resolution (1 pixel \approx 0.05 mm). However, even this resolution is not always sufficient to take into account the small colonies shading particularities. Here we demonstrate two examples of small colonies reconstruction. These examples do not generalize all possible cases of small colonies, but it already gives us ideas about the efficiency of the small colonies reconstruction by the proposed approach.

The first example of a small colony reconstruction, Fig. 5.15, is successful. This reconstruction is made for the colony of 22 pixels in diameter which approximately corresponds to 1.15 mm. In Fig. 5.15(a) there is one of four initial images (left bottom illumination). The resulting surfaces are shown in Fig. 5.15(b) and Fig. 5.15(c). The reconstructed colony is specular, and specularity correction procedure has also its influence on the reconstruction results. Even for such a small object, we managed to obtain the surface close to realistic one compared to microbiological experience.

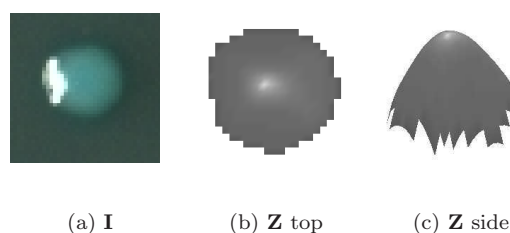


FIGURE 5.15: Example 1 of the reconstructed small colony (diameter= 22 pixels, 1.15 mm).

We also propose another example of the small colony reconstruction, Fig. 5.16(a). For this colony of 14 pixels in diameter (approximately 0.73 mm), we did not manage to obtain a qualitative reconstruction, Fig. 5.16(b) and Fig. 5.16(c). There are several possible reasons for such a failure. First of all, 14 pixels of diameter may be not sufficient to represent the object shading for the photometric stereo reconstruction. The second potential problem consists in a very small elevation of the object.

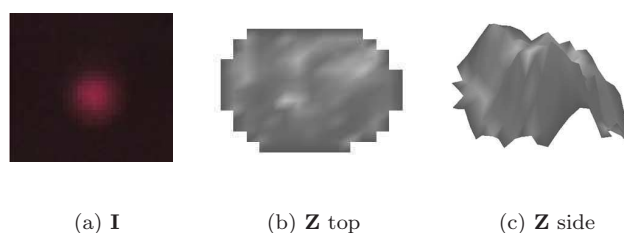


FIGURE 5.16: Example 2 of the reconstructed small colony (diameter= 14 pixels, 0.73 mm).

The size of the colony is important for the reconstruction. For too small colonies we just may have not enough information on the object shading due to different illuminations. This lack of information leads to inefficient reconstruction results.

Objects with concavities

Microbial colony surfaces can contain concave parts. In this section with the following examples we aim at investigating if the proposed approach is capable to reconstruct surfaces of such a type. In Fig. 5.17 there are an image of colony to be reconstructed, as well as the reconstruction results. The diameter of this colony is 2.25 mm. The obtained surface contains the sought concavity. It means that the reconstructed surface of this type can be used for an automated analysis.

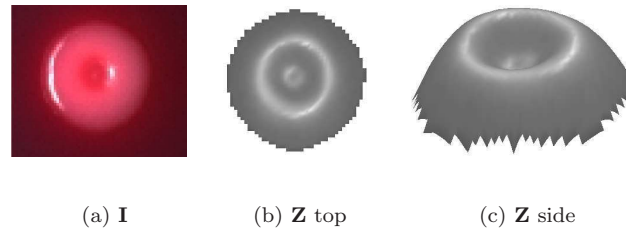


FIGURE 5.17: Example of the reconstructed concave colony (smaller diameter = 43 pixels, 2.25 mm).

Sophisticated shapes

In this section we present reconstructions of different more sophisticated colony surfaces. These surfaces are relatively large (up to several centimeters) and contain combinations of convex and concave regions.

The first example of reconstruction is given for the colony consisting of a convex hemispherical object without a section, Fig. 5.18. This colony is of 12 mm in diameter. In the place of the missing section a multi-layer inner structure of the colony is visible. Thanks to the obtained reconstruction, this multi-layer structure becomes more evident, especially on the side surface view, Fig 5.18(c). The 3D visualization of the surface may provide a useful support to microbiologist during such a complex object characterization.

In Fig. 5.19 we demonstrate the second example of the reconstructed sophisticated surface for the colony of 10.17 mm in diameter. This surface consists of a main convex part and a supplementary ring around the colony. Notable in this example is the colony rough surface. Using only the reconstructed surface we cannot say, if this particular texture is related to the reflections of the ground truth surface, or this roughness appears for smooth surface due to complex reflecting properties of

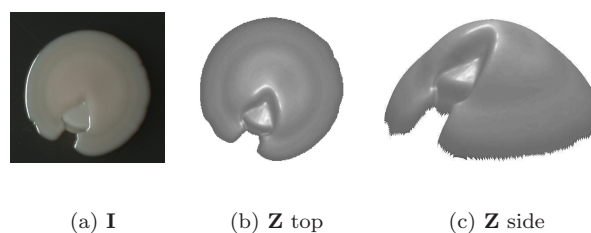


FIGURE 5.18: Example 1 of the reconstructed colony of sophisticated shape (diameter = 230 pixels, 12 mm).

the colony. In order to resolve this ambiguity, one needs supplementary images of the object, for example, taken with other illuminations.

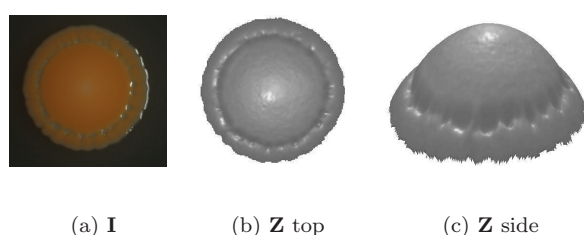


FIGURE 5.19: Example 2 of the reconstructed colony of sophisticated shape (diameter = 194 pixels, 10.17 mm).

The last example of this section, Fig. 5.20, corresponds to the colony with plural alternations of convex and concave regions. The smaller diameter of the colony is 17.45 mm. The proposed method of images processing cope with the task of this complex surface reconstruction, and its visualization can be very helpful for the object analysis and characterization.

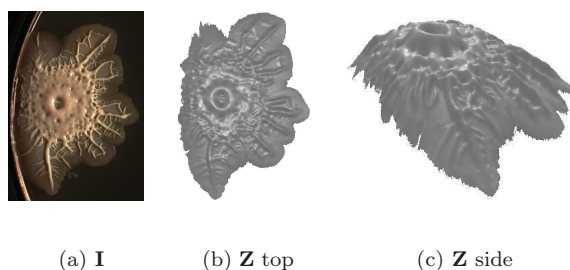


FIGURE 5.20: Example 3 of the reconstructed colony of sophisticated shape (smaller diameter = 333 pixels, 17.45 mm).

However, due to the bas-relief ambiguity we cannot estimate an absolute elevation or the deepness of the object concavities. Thus, the proposed reconstruction chain

may not be used in tasks, for which the absolute object size and the absolute height of different object parts are important. One of such a task, for example, consists in comparing the heights of the colonies obtained during two independent reconstruction procedures for two different Petri dishes.

In this section with several examples, we show how the proposed processing chain can cope with the task of microbial colonies surface reconstruction, which provides a useful tool for the colonies characterization. The more complex is the object, the more important becomes to complete the information present on images with the reconstructed surface.

5.4 Conclusions

This chapter is dedicated to the application of the proposed method of uncalibrated non-Lambertian photometric stereo to the images of Petri dish. 3D surface of the microbial colony inside the dish can be crucial for the analysis and diagnostics. The microbial colonies are specific objects to be reconstructed, they may be of small size and of complex non-homogeneous reflection properties. The microbial colonies can be either of Lambertian nature or not. If the colony is non-Lambertian, presence of specularities affects the resulting surface up to making impossible its further analysis. For such colonies, the proposed algorithm of specularities correction turns out to be mandatory. On the other hand, for the colony of properties close to Lambertian, the algorithm of correction have no action. Thus, the proposed method does not need information on the type of the colony surface a priori, and it is possible to obtain qualitative surfaces of both types.

We presented the reconstruction results only for a few types of microbial objects among all possible diversity. Nevertheless, these results confirm the viability of the proposed uncalibrated non-Lambertian photometric stereo reconstruction for real-world images obtained using industrial imaging device with complex illumination which is considered unknown.

Chapter 6

Conclusions and Perspectives

With the growing interest to the 3D technologies, surface from images reconstruction approaches have become more popular. The third dimension allows presenting the photographed objects more detailed, complementing the information on the object surface missing in a 2D image by the means of the existing 'shape-from-x' techniques.

This dissertation contributes to the field of the photometric stereo surface reconstruction and is dedicated to the task of uncalibrated non-Lambertian photometric stereo. We propose an original approach consisting in two stages: correction of specular zones as image intensity outliers and reconstruction of the object surface on the basis of these corrected images. The applicability of the developed approach for the real-world tasks is shown providing the reconstruction of microbial colonies on the basis of Petri dish images.

6.1 Conclusions

In course of the presented work we have developed an approach of non-Lambertian uncalibrated surface reconstruction. The processing is realized for the images taken from a fixed camera position and varying illuminations. The main advantage of the proposed chain consists in its independence from the prior knowledge on the light sources directions and intensities used during the input images acquisition as well as on the reconstructed surface properties. Such an independence enables the usage of the described method in industrial acquisition devices, when the characteristics of lighting and reconstructed objects are not known a priori or are difficult for modeling.

The first phase of the proposed approach permits to correct images of the input sequence from the negative optical effect such as specularity. This effect is crucial for the photometric stereo surface reconstruction. It is not only a property of the microbial colonies, reconstruction of which was in motivations of the present work, but also of many other real-world objects. The proposed algorithm of correction does not need prior information neither on the light sources intensities nor on the light source directions used during images acquisition. It permits to process only specular zones on the basis of all images of the sequence preserving shading particularities of the image which are important for the further reconstruction. Thanks to a sigmoid correcting function the algorithm of soft correction takes into account a continuous nature of specularity and avoids sharp transitions between corrected and non-corrected image regions. The correction of specularities transforms images of non-Lambertian objects to images of 'pseudo-Lambertian' ones, which could be further used as input data for Lambertian photometric stereo reconstruction.

The second phase of the proposed chain is dedicated to the uncalibrated Lambertian photometric stereo approach. We present an original alternating optimization algorithm which permits to realize matrix factorization and to separate the information contained in a set of the input images into albedo, normal map, light source intensities and light source directions matrices. The used multiplicative model allows simulating the images formation process for Lambertian object surfaces. This model is extended by the means of the value and quality constraints. The constraints allows not only to achieve the algorithm convergence, but also to obtain values of the model components which can be easily interpretable. The proposed reconstruction scheme is totally 'data-driven', and even the components initialization for the iterative algorithm is made on the basis of the input images. The quality of reconstruction by the algorithm depends on the uniformity of light sources spatial distribution around the object as well as the number of the input images.

In order to confirm the workability of the reconstruction, the proposed two-phase processing was applied to the images of Petri dish. Microbial colonies grown inside the dish can be of various properties, including various types of elevation and surface texture. The surface information is important for microorganisms analysis and can be hardly interpreted from an image. The Petri dish imaging device permits to obtain data with a fix camera position and different illuminations, which are suitable for the further photometric stereo. The illumination used in the imaging device is very complex for modeling and may be not stable. Thus, the proposed algorithm

of the uncalibrated non-Lambertian photometric stereo copes with this lack of information on lighting and provides an efficient reconstruction of microbial colonies with different properties including complex surfaces with alternation of convex and concave regions. It figures out to deal with small and highly specular objects. Moreover, it is not necessary to know a priori if the microbial colony has specularities or not, the method provides satisfactory results for both types of objects. The obtained results of reconstruction are promising and confirm the possibility to use the algorithm for the Petri images automated analysis.

6.2 Future Work

There are a lot of research directions which could be envisaged in perspectives of this thesis work. Some of these perspectives refer separately to different processing chain phases, others are dedicated to the complete chain of non-Lambertian uncalibrated photometric stereo.

The first of the perspective consists in the extension of specular removal algorithm for the case of shadows correction. We did not consider shadows in our reconstruction approach, but they can create the same inconvenient as specular zones for the photometric stereo. It is necessary to study in more details the shadows phenomenon, and then to adapt the proposed correction function and its parameters for shadows correction.

Another important question which should be studied in more details is related to the proposed uncalibrated Lambertian reconstruction algorithm. This is the question of bas-relief ambiguity. The surface obtained by this algorithm suffers from the three-parameter ambiguity, and its absolute characteristics of altitude can not be used for the surface analysis till this ambiguity is not resolved.

In real-world tasks, the reconstructed object is not the only object of the scene. There are also other foreign surfaces which creates supplementary reflections of light and violate the assumption of the directed point light source. The supplementary reflections can also be created by the object itself if it has a sophisticated shape with concavity and convexity. The problem of such a parasitic reflections (or interreflection) can be solved by implementation of the additional surface refinement procedure. As in [27] this procedure could be just a part of the iterative algorithm.

Another existing problem not yet solved of photometric stereo is transparent surfaces reconstruction. When the light interacts with a surface of this type, many supplementary events of scattering occur. Modeling of scattering is a difficult task, because this phenomenon is associated with the surface material and incident light physical properties in molecular level, and thus can be hardly independent from the prior knowledge on these two components. In order to include a possibility of transparent surfaces reconstruction, the proposed model and its components evaluation should be considerably transformed.

The proposed alternating optimization algorithm is iterative. With the practical examples we demonstrated that the error between the observed data and the multiplication of the found components decreases which each iteration. Mathematical analysis of the convergence (rate, order, etc.) is one of the possible directions of future work. This will permit to better characterize the efficiency of the presented algorithm.

The both stages of the proposed method (specularities correction and uncalibrated photometric stereo) require additional parameters for their efficient work. Some of these parameters can be easily chosen, and during the methods description we give our recommendations on their possible values. Other parameters need more efforts and their values are not evident. The automatic tuning of the parameters is in perspectives of our work.

There is another application of the proposed photometric stereo model. This model enables separation of the information of surface and illumination. In computer graphics and potentially in augmented reality task, the problem of re-lighting can also be demanded. The re-lighting consists in a replacement of the actual illumination sources used for the initial images acquisition by the sources with other directions and intensities. Thanks to the separation provided by the proposed algorithm, for the creation of new images it is sufficient to replace the appropriate matrices by the new ones and to re-utilize the proposed model for the images creation. Investigation of the re-lighting problem is also in perspectives of our work.

Appendix A

Depth map from normals reconstruction

In this appendix we present our implementation of the depth map from normals reconstruction on the basis of [95]. As described in Chapter 4, for each depth map pixel (x, y) it can be written the linear system of equations:

$$\begin{cases} n_z(x, y)\mathbf{z}(x + 1, y) - n_z(x, y)\mathbf{z}(x, y) = -n_x(x, y), \\ n_z(x, y)\mathbf{z}(x, y + 1) - n_z(x, y)\mathbf{z}(x, y) = -n_y(x, y), \\ n_z(x, y)\mathbf{z}(x - 1, y) - n_z(x, y)\mathbf{z}(x, y) = n_x(x, y), \\ n_z(x, y)\mathbf{z}(x, y - 1) - n_z(x, y)\mathbf{z}(x, y) = n_y(x, y), \end{cases} \quad (\text{A.1})$$

where $n_x(x, y)$, $n_y(x, y)$ and $n_z(x, y)$ correspond to the 3D normal and $\mathbf{z}(x, y)$ denotes the pixel depth.

In order to obtain the final surface, the global system of equations should be solved. In total, the number of pixels is P , then this global system consists of $4P$ equations. Let us see how this task can be algorithmically implemented for the most efficient solution.

After the photometric stereo reconstruction there are values of normals corresponding to each image pixel. For the depth map evaluation, let us firstly present all normals in the form of three vectors: \mathbf{n}_x , \mathbf{n}_y and \mathbf{n}_z . Each of this vector corresponds to unwrapped matrices of normals components \mathbf{N}_x , \mathbf{N}_y and \mathbf{N}_z respectively. In order to find the vectorized depth map $\mathbf{z}(x, y)$ we should resolve the linear system of equations.

As it can be seen in Eq. A.1, the depth map and normals components values are presented in 2D form. In order to transform them all, let us have the matrix of linear indices **IND**:

$$\mathbf{IND} = \begin{bmatrix} 1 & (X+1) & \dots & (Y-2)X+1 & (Y-1)X+1 \\ 2 & (X+2) & \dots & (Y-2)X+2 & (Y-1)X+2 \\ \cdot & \cdot & \dots & \cdot & \cdot \\ \cdot & \cdot & \dots & \cdot & \cdot \\ \cdot & \cdot & \dots & \cdot & \cdot \\ (X-1) & (X+X-1) & \dots & (Y-2)X+(X-1) & (Y-1)X+(X+1) \\ X & 2X & \dots & (Y-1)X & YX \end{bmatrix}, \quad (\text{A.2})$$

where X and Y correspond to the number of image lines and columns respectively. Into this matrix **IND** we can allocate 4 sub-matrices which correspond to the following intervals of indices:

$$\text{LeftTop matrix} \quad \mathbf{LT} = \mathbf{IND}(1 : X-1, 1 : Y-1),$$

$$\text{LeftBottom matrix} \quad \mathbf{LB} = \mathbf{IND}(2 : X, 1 : Y-1),$$

$$\text{RightTop matrix} \quad \mathbf{RT} = \mathbf{IND}(1 : X-1, 2 : Y),$$

$$\text{RightBottom matrix} \quad \mathbf{RB} = \mathbf{IND}(2 : X, 2 : Y).$$

These matrices serve for bias implementation: **LT** for $(x-1, y)$ pixel, **LB** for $(x+1, y)$ pixel, **RT** for $(x, y-1)$ pixel and **RB** for $(x, y+1)$ pixel. Unwrapped values of this matrices of index corresponds to vectors **linLT**, **linLB**, **linRT** and **linRB** respectively. Number of elements in each vector equals to ne .

Now using all the described notations and Algorithm 6, the depth map from normals can be efficiently found using a linear system $\mathbf{Pz} = \mathbf{b}$, the matrix **P** and the vector **b** should be properly filled to correspond to the global system of equations (Eq. A.1). The matrix **P** is sparse, so any existing sparse least mean square algorithm can be used for **z** finding. Moreover this algorithm can be presented in more compact matrix form suitable for the implementation with the Matlab system.

Algorithm 6 Depth map from normals reconstruction algorithm.

1: **In**:
 $\mathbf{n}_x, \mathbf{n}_y, \mathbf{n}_z$ - matrices of normals components;
 $\mathbf{linLT}, \mathbf{linLB}, \mathbf{linRT}, \mathbf{linRB}$ - linear indices of supplementary matrix **IND**;
 ne , - number of elements in \mathbf{linLT} (or \mathbf{linLB} , or \mathbf{linRT} , or \mathbf{linRB});
 X , - number of lines in the image;
 Y , - number of columns in the image.

2: **Out**:
 \mathbf{z} of size $XY \times 1$ - estimate of the unwrapped depth map.

3: $\mathbf{A} \leftarrow$ matrix of zeros of size $4ne \times XY$
4: $\mathbf{b} \leftarrow$ vector of zeros of size $4ne \times 1$
5: **for** $i = 1, \dots, ne$ **do**

6: $\mathbf{P}(i, \mathbf{linLT}(i)) \leftarrow -\mathbf{n}_z(\mathbf{linLT}(i))$
7: $\mathbf{P}(i, \mathbf{linLB}(i)) \leftarrow \mathbf{n}_z(\mathbf{linLT}(i))$
8: $\mathbf{b}(i) \leftarrow -\mathbf{n}_x(\mathbf{linLT}(i))$

9: $\mathbf{P}(ne + i, \mathbf{linLT}(i)) \leftarrow -\mathbf{n}_z(\mathbf{linLT}(i))$
10: $\mathbf{P}(ne + i, \mathbf{linRT}(i)) \leftarrow \mathbf{n}_z(\mathbf{linLT}(i))$
11: $\mathbf{b}(ne + i) \leftarrow -\mathbf{n}_y(\mathbf{linLT}(i))$

12: $\mathbf{P}(2ne + i, \mathbf{linRB}(i)) \leftarrow -\mathbf{n}_z(\mathbf{linRB}(i))$
13: $\mathbf{P}(2ne + i, \mathbf{linLB}(i)) \leftarrow \mathbf{n}_z(\mathbf{linRB}(i))$
14: $\mathbf{b}(2ne + i) \leftarrow \mathbf{n}_x(\mathbf{linRB}(i))$

15: $\mathbf{P}(3ne + i, \mathbf{linRB}(i)) \leftarrow -\mathbf{n}_z(\mathbf{linRB}(i))$
16: $\mathbf{P}(3ne + i, \mathbf{linRT}(i)) \leftarrow \mathbf{n}_z(\mathbf{linRB}(i))$
17: $\mathbf{b}(3ne + i) \leftarrow \mathbf{n}_y(\mathbf{linRB}(i))$

18: **end for**
19: $\mathbf{z} \leftarrow \text{SparseLeastMeanSquare}(\mathbf{P}, \mathbf{b})$

Appendix B

Photometric stereo model without constraints

The reflectance model is $\mathbf{M} = \mathbf{A} \odot \mathbf{L} \odot (\mathbf{V}^T \mathbf{N})$, where symbol ' \odot ' denotes the point-to-point or Hadamard product. If we include the structuring matrices \mathbf{W}_A , \mathbf{W}_L and \mathbf{W}_V into the model, it becomes $\mathbf{M} = (\mathbf{W}_A \tilde{\mathbf{A}}) \odot (\tilde{\mathbf{L}} \mathbf{W}_L) \odot ((\mathbf{W}_V \tilde{\mathbf{V}}^T) \mathbf{N})$.

B.1 Albedo values matrix evaluation

Let $\mathbf{F} = (\tilde{\mathbf{L}} \mathbf{W}_L) \odot ((\mathbf{W}_V \tilde{\mathbf{V}}^T) \mathbf{N})$ and $\mathbf{K} = \mathbf{W}_A \tilde{\mathbf{A}}$. We suppose that \mathbf{F} and \mathbf{W}_A are known. $\tilde{\mathbf{A}}$ is albedo values matrix which we want to estimate.

$$\begin{aligned} \mathbf{M} &= \mathbf{K} \odot \mathbf{F}, \\ \mathbf{K} &= \underset{\mathbf{K}}{\operatorname{argmin}} \|\mathbf{M} - \mathbf{K} \odot \mathbf{F}\|_2^2, \\ 2(\mathbf{M} - \mathbf{K} \odot \mathbf{F}) \odot \mathbf{F} &= 0, \\ \mathbf{M} \odot \mathbf{F} - \mathbf{K} \odot \mathbf{F} \odot \mathbf{F} &= 0, \\ \mathbf{K} &= \frac{\mathbf{M} \odot \mathbf{F}}{\mathbf{F} \odot \mathbf{F}} = \mathbf{M} \oslash \mathbf{F}, \end{aligned}$$

where symbol ' \oslash ' denotes the point-to-point division (analogue of the point-to-point or Hadamard product) of matrices. When \mathbf{K} is found we can return to the initial change of variables $\mathbf{K} = \mathbf{W}_A \tilde{\mathbf{A}}$ to find matrix $\tilde{\mathbf{A}}$ also using least square approximation:

$$\tilde{\mathbf{A}} = \underset{\tilde{\mathbf{A}}}{\operatorname{argmin}} \left\| \mathbf{K} - \mathbf{W}_A \tilde{\mathbf{A}} \right\|_2^2,$$

$$\begin{aligned}
2\mathbf{W}_A^T(\mathbf{K} - \mathbf{W}_A\tilde{\mathbf{A}}) &= 0, \\
\mathbf{W}_A^T\mathbf{K} - \mathbf{W}_A^T\mathbf{W}_A\tilde{\mathbf{A}} &= 0, \\
\mathbf{W}_A^T\mathbf{W}_A\tilde{\mathbf{A}} &= \mathbf{W}_A^T\mathbf{K}, \\
\tilde{\mathbf{A}} &= (\mathbf{W}_A^T\mathbf{W}_A)^{-1}\mathbf{W}_A^T\mathbf{K}.
\end{aligned}$$

Now we replace \mathbf{K} by the value already found on the previous step ($\mathbf{K} = \mathbf{M} \oslash \mathbf{F}$):

$$\tilde{\mathbf{A}} = (\mathbf{W}_A^T\mathbf{W}_A)^{-1}\mathbf{W}_A^T(\mathbf{M} \oslash \mathbf{F}).$$

It was defined at the beginning of the calculations that $\mathbf{F} = (\tilde{\mathbf{L}}\mathbf{W}_L) \odot ((\mathbf{W}_V\tilde{\mathbf{V}}^T)\mathbf{N})$. Thus matrix of albedo values can be found as:

$$\tilde{\mathbf{A}} = ((\mathbf{W}_A^T\mathbf{W}_A)^{-1}\mathbf{W}_A^T) \left(\mathbf{M} \oslash \left((\tilde{\mathbf{L}}\mathbf{W}_L) \odot ((\mathbf{W}_V\tilde{\mathbf{V}}^T)\mathbf{N}) \right) \right).$$

B.2 Light sources intensities vector evaluation

Let $\mathbf{F} = (\mathbf{W}_A\tilde{\mathbf{A}}) \odot ((\mathbf{W}_V\tilde{\mathbf{V}}^T)\mathbf{N})$ and $\mathbf{K} = \tilde{\mathbf{L}}\mathbf{W}_L$. We suppose that \mathbf{F} and \mathbf{W}_L are known. $\tilde{\mathbf{L}}$ is a light intensities vector which we want to estimate.

$$\begin{aligned}
\mathbf{M} &= \mathbf{K} \odot \mathbf{F}, \\
\mathbf{K} &= \underset{\mathbf{K}}{\operatorname{argmin}} \|\mathbf{M} - \mathbf{K} \odot \mathbf{F}\|_2^2, \\
2(\mathbf{M} - \mathbf{K} \odot \mathbf{F}) \odot \mathbf{F} &= 0, \\
\mathbf{M} \odot \mathbf{F} - \mathbf{K} \odot \mathbf{F} \odot \mathbf{F} &= 0, \\
\mathbf{K} &= \frac{\mathbf{M} \odot \mathbf{F}}{\mathbf{F} \odot \mathbf{F}} = \mathbf{M} \oslash \mathbf{F},
\end{aligned}$$

where symbol ' \oslash ' is the point-to-point division of matrices. When \mathbf{K} is found we can return to the initial change of variables $\mathbf{K} = \tilde{\mathbf{L}}\mathbf{W}_L$ to find matrix $\tilde{\mathbf{L}}$ also using least square approximation:

$$\begin{aligned}
\tilde{\mathbf{L}} &= \underset{\tilde{\mathbf{L}}}{\operatorname{argmin}} \left\| \mathbf{K} - \tilde{\mathbf{L}}\mathbf{W}_L \right\|_2^2, \\
2(\mathbf{K} - \tilde{\mathbf{L}}\mathbf{W}_L)\mathbf{W}_L^T &= 0, \\
\mathbf{K}\mathbf{W}_L^T - \tilde{\mathbf{L}}\mathbf{W}_L\mathbf{W}_L^T &= 0, \\
\tilde{\mathbf{L}}\mathbf{W}_L\mathbf{W}_L^T &= \mathbf{K}\mathbf{W}_L^T, \\
\tilde{\mathbf{L}} &= \mathbf{K}\mathbf{W}_L^T(\mathbf{W}_L\mathbf{W}_L^T)^{-1}.
\end{aligned}$$

Now we replace \mathbf{K} by the value already found on the previous step ($\mathbf{K} = \mathbf{M} \oslash \mathbf{F}$):

$$\tilde{\mathbf{L}} = (\mathbf{M} \oslash \mathbf{F}) \mathbf{W}_L^T (\mathbf{W}_L \mathbf{W}_L^T)^{-1}.$$

It was defined at the beginning of the calculations that $\mathbf{F} = (\mathbf{W}_A \tilde{\mathbf{A}}) \odot ((\mathbf{W}_V \tilde{\mathbf{V}}^T) \mathbf{N})$. Thus vector of light sources intensities can be found as:

$$\tilde{\mathbf{L}} = \left(\mathbf{M} \oslash \left((\mathbf{W}_A \tilde{\mathbf{A}}) \odot ((\mathbf{W}_V \tilde{\mathbf{V}}^T) \mathbf{N}) \right) \right) \mathbf{W}_L^T (\mathbf{W}_L \mathbf{W}_L^T)^{-1}.$$

B.3 Lighting sources directions matrix evaluation

Let $\mathbf{F} = (\mathbf{W}_A \tilde{\mathbf{A}}) \odot (\tilde{\mathbf{L}} \mathbf{W}_L)$ and $\mathbf{K} = \mathbf{W}_V \tilde{\mathbf{V}}^T \mathbf{N}$. We suppose that \mathbf{F} , \mathbf{W}_V and \mathbf{N} are known. $\tilde{\mathbf{V}}^T$ is lighting directions matrix which we want to estimate.

$$\begin{aligned} \mathbf{M} &= \mathbf{F} \odot \mathbf{K}, \\ \mathbf{K} &= \underset{\mathbf{K}}{\operatorname{argmin}} \|\mathbf{M} - \mathbf{F} \odot \mathbf{K}\|_2^2, \\ 2(\mathbf{M} - \mathbf{F} \odot \mathbf{K}) \odot \mathbf{F} &= 0, \\ \mathbf{M} \odot \mathbf{F} - \mathbf{F} \odot \mathbf{K} \odot \mathbf{F} &= 0, \\ \mathbf{K} &= \frac{\mathbf{M} \odot \mathbf{F}}{\mathbf{F} \odot \mathbf{F}} = \mathbf{M} \oslash \mathbf{F}, \end{aligned}$$

where symbol ' \odot ' denotes the point-to-point division of matrices. When \mathbf{K} is found we can return to the initial change of variables $\mathbf{K} = \mathbf{W}_V \tilde{\mathbf{V}}^T \mathbf{N}$ to find matrix $\tilde{\mathbf{V}}^T$ also using least square approximation:

$$\begin{aligned} \tilde{\mathbf{V}}^T &= \underset{\tilde{\mathbf{V}}^T}{\operatorname{argmin}} \left\| \mathbf{K} - \mathbf{W}_V \tilde{\mathbf{V}}^T \mathbf{N} \right\|_2^2, \\ 2\mathbf{W}_V^T (\mathbf{K} - \mathbf{W}_V \tilde{\mathbf{V}}^T \mathbf{N}) \mathbf{N}^T &= 0, \\ \mathbf{W}_V^T \mathbf{K} \mathbf{N}^T - \mathbf{W}_V^T \mathbf{W}_V \tilde{\mathbf{V}}^T \mathbf{N} \mathbf{N}^T &= 0, \\ \mathbf{W}_V^T \mathbf{W}_V \tilde{\mathbf{V}}^T \mathbf{N} \mathbf{N}^T &= \mathbf{W}_V^T \mathbf{K} \mathbf{N}^T, \\ \tilde{\mathbf{V}}^T &= (\mathbf{W}_V^T \mathbf{W}_V)^{-1} \mathbf{W}_V^T \mathbf{K} \mathbf{N}^T (\mathbf{N} \mathbf{N}^T)^{-1}. \end{aligned}$$

Now we replace \mathbf{K} by the value already found on the previous step ($\mathbf{K} = \mathbf{M} \oslash \mathbf{F}$):

$$\tilde{\mathbf{V}}^T = (\mathbf{W}_V^T \mathbf{W}_V)^{-1} \mathbf{W}_V^T (\mathbf{M} \oslash \mathbf{F}) \mathbf{N}^T (\mathbf{N} \mathbf{N}^T)^{-1}.$$

It was defined at the beginning of the calculations that $\mathbf{F} = (\mathbf{W}_A \tilde{\mathbf{A}}) \odot (\tilde{\mathbf{L}} \mathbf{W}_L)$. Thus, matrix of light sources directions can be found as:

$$\tilde{\mathbf{V}}^T = (\mathbf{W}_V^T \mathbf{W}_V)^{-1} \mathbf{W}_V^T \left(\mathbf{M} \oslash ((\mathbf{W}_A \tilde{\mathbf{A}}) \odot (\tilde{\mathbf{L}} \mathbf{W}_L)) \right) \mathbf{N}^T (\mathbf{N} \mathbf{N}^T)^{-1}.$$

B.4 Surface normals matrix evaluation

Let $\mathbf{F} = (\mathbf{W}_A \tilde{\mathbf{A}}) \odot (\tilde{\mathbf{L}} \mathbf{W}_L)$ and $\mathbf{K} = \mathbf{W}_V \tilde{\mathbf{V}}^T \mathbf{N}$. We suppose that \mathbf{F} , \mathbf{W}_V , and $\tilde{\mathbf{V}}^T$ are known. \mathbf{N} is surface normals matrix which we want to estimate.

$$\begin{aligned} \mathbf{M} &= \mathbf{F} \odot \mathbf{K}, \\ \mathbf{K} &= \underset{\mathbf{K}}{\operatorname{argmin}} \|\mathbf{M} - \mathbf{F} \odot \mathbf{K}\|_2^2, \\ 2(\mathbf{M} - \mathbf{F} \odot \mathbf{K}) \odot \mathbf{F} &= 0, \\ \mathbf{M} \odot \mathbf{F} - \mathbf{F} \odot \mathbf{K} \odot \mathbf{F} &= 0, \\ \mathbf{K} &= \frac{\mathbf{M} \odot \mathbf{F}}{\mathbf{F} \odot \mathbf{F}} = \mathbf{M} \oslash \mathbf{F}, \end{aligned}$$

where symbol ' \oslash ' is the point-to-point division of matrices. When \mathbf{K} is found we can return to the initial change of variables $\mathbf{K} = \mathbf{W}_V \tilde{\mathbf{V}}^T \mathbf{N}$ to find matrix \mathbf{N} also using least square approximation:

$$\begin{aligned} \mathbf{N} &= \underset{\mathbf{N}}{\operatorname{argmin}} \left\| \mathbf{K} - (\mathbf{W}_V \tilde{\mathbf{V}}^T) \mathbf{N} \right\|_2^2, \\ (\mathbf{W}_V \tilde{\mathbf{V}}^T)^T 2(\mathbf{K} - (\mathbf{W}_V \tilde{\mathbf{V}}^T) \mathbf{N}) &= 0, \\ (\mathbf{W}_V \tilde{\mathbf{V}}^T)^T \mathbf{K} - (\mathbf{W}_V \tilde{\mathbf{V}}^T)^T (\mathbf{W}_V \tilde{\mathbf{V}}^T) \mathbf{N} &= 0, \\ (\mathbf{W}_V \tilde{\mathbf{V}}^T)^T (\mathbf{W}_V \tilde{\mathbf{V}}^T) \mathbf{N} &= (\mathbf{W}_V \tilde{\mathbf{V}}^T)^T \mathbf{K}, \\ \mathbf{N} &= \left((\mathbf{W}_V \tilde{\mathbf{V}}^T)^T (\mathbf{W}_V \tilde{\mathbf{V}}^T) \right)^{-1} (\mathbf{W}_V \tilde{\mathbf{V}}^T)^T \mathbf{K}. \end{aligned}$$

Now we replace \mathbf{K} by the value already found on the previous step ($\mathbf{K} = \mathbf{M} \oslash \mathbf{F}$):

$$\mathbf{N} = \left((\mathbf{W}_V \tilde{\mathbf{V}}^T)^T (\mathbf{W}_V \tilde{\mathbf{V}}^T) \right)^{-1} (\mathbf{W}_V \tilde{\mathbf{V}}^T)^T (\mathbf{M} \oslash \mathbf{F}).$$

It was defined at the beginning of the calculations that $\mathbf{F} = (\mathbf{W}_A \tilde{\mathbf{A}}) \odot (\tilde{\mathbf{L}} \mathbf{W}_L)$. Thus matrix of surface normals can be found as:

$$\mathbf{N} = \left((\mathbf{W}_V \tilde{\mathbf{V}}^T)^T (\mathbf{W}_V \tilde{\mathbf{V}}^T) \right)^{-1} (\mathbf{W}_V \tilde{\mathbf{V}}^T)^T \left(\mathbf{M} \oslash \left((\mathbf{W}_A \tilde{\mathbf{A}}) \odot (\tilde{\mathbf{L}} \mathbf{W}_L) \right) \right).$$

Appendix C

Photometric stereo model with constraints

The reflectance model is $\mathbf{M} = \mathbf{A} \odot \mathbf{L} \odot (\mathbf{V}^T \mathbf{N})$, where symbol ' \odot ' denotes the point-to-point or Hadamard product. If we include the structuring matrices \mathbf{W}_A , \mathbf{W}_L , and \mathbf{W}_V into the model, it becomes $\mathbf{M} = (\mathbf{W}_A \tilde{\mathbf{A}}) \odot (\tilde{\mathbf{L}} \mathbf{W}_L) \odot ((\mathbf{W}_V \tilde{\mathbf{V}}^T) \mathbf{N})$.

C.1 Albedo values approximation with interval constraint

Albedo values matrix can be obtained as results of minimization of the following expression:

$$\tilde{\mathbf{A}} = \min_{\tilde{\mathbf{A}}} \left\| \mathbf{M} - f(\tilde{\mathbf{A}}, \tilde{\mathbf{L}}, \tilde{\mathbf{V}}, \mathbf{N}) \right\|_2^2$$

subject to $0 \leq a_{i,j} \leq 1$,

where $i = 1, \dots, 3$; $j = 1, \dots, P$; $f(\tilde{\mathbf{A}}, \tilde{\mathbf{L}}, \tilde{\mathbf{V}}, \mathbf{N})$ defines the model.

Let us have a change of variables $\mathbf{F} = (\tilde{\mathbf{L}} \mathbf{W}_L) \odot ((\mathbf{W}_V \tilde{\mathbf{V}}^T) \mathbf{N})$ and $\mathbf{B} = \mathbf{M} \oslash \mathbf{F}$, then the model can be rewritten as $\mathbf{B} = \mathbf{W}_A \tilde{\mathbf{A}}$. Schematically we obtain the following system of equations in matrix form:

$$\text{Case 1, } \lambda > 0, r = 0: \begin{cases} \lambda \geq 2 \sum_{i=1}^N b_{i,1}, \\ \lambda = 0, \text{ false.} \end{cases}$$

$$\text{Case 2, } \lambda > 0, r > 0: \begin{cases} \lambda \geq 2 \sum_{i=1}^N (b_{i,1} - r), \\ 1 - r = 0 \rightarrow r = 1, \text{ true.} \end{cases}$$

$$\text{Case 3, } \lambda = 0, r > 0: (-2 \sum_{i=1}^N (b_{i,1} - r))a = 0 \rightarrow r = \frac{\sum_{i=1}^N b_{i,1}}{N}, \text{ true.}$$

$$\text{Case 4, } \lambda = 0, r = 0: -2 \sum_{i=1}^N (b_{i,1} - r) + \lambda \geq 0, \rightarrow -2 \sum_{i=1}^N b_{i,1} \geq 0, \text{ true.}$$

Due to different cases we obtain three different solutions:

- 1) For lower bound Case 4 is used: when $\sum_{i=1}^N b_{i,1} < 0$ then $r = 0$.
- 2) For upper bound Case 2 is used: when $\sum_{i=1}^N b_{i,1} > 1$ then $r = 1$.
- 3) For all other situations Case 3 is used: $r = \frac{\sum_{i=1}^N b_{i,1}}{N}$. In other words, albedo equals to the mean value of the appropriate matrix \mathbf{B} elements.

C.2 Light sources intensities approximation with non-negativity constraint

$$\min_{\tilde{\mathbf{L}}} \left\| \mathbf{M} - f(\tilde{\mathbf{A}}, \tilde{\mathbf{L}}, \tilde{\mathbf{V}}, \mathbf{N}) \right\|_2^2$$

subject to $0 > l_i > H$,

where $i = 1, \dots, 3N$; H is the maximal possible value of image intensity;

$f(\tilde{\mathbf{A}}, \tilde{\mathbf{L}}, \tilde{\mathbf{V}}, \mathbf{N})$ defines the model.

Let us have a change of variables $\mathbf{F} = (\mathbf{W}_A \tilde{\mathbf{A}}) \odot ((\mathbf{W}_V \tilde{\mathbf{V}}^T) \mathbf{N})$ and $\mathbf{B} = \mathbf{M} \oslash \mathbf{F}$, then the model can be rewritten as $\mathbf{B} = \tilde{\mathbf{L}} \mathbf{W}_L$. Schematically we obtain the following system of equations in matrix form:

$$\begin{array}{c} \underbrace{\hspace{10em}}_P \\ \left. \begin{array}{c|c|c|c} \hline b_{1,1} & b_{1,2} & \dots & b_{1,P} \\ \hline b_{2,1} & b_{2,2} & \dots & b_{2,P} \\ \hline b_{3,1} & b_{3,2} & \dots & b_{3,P} \\ \hline \vdots & \vdots & \vdots & \vdots \\ \hline b_{(3N-2),1} & b_{(3N-2),2} & \dots & b_{(3N-2),P} \\ \hline b_{(3N-1),1} & b_{(3N-1),2} & \dots & b_{(3N-1),P} \\ \hline b_{3N,1} & b_{3N,2} & \dots & b_{3N,P} \\ \hline \end{array} \right\} \begin{array}{l} R_1 \\ G_1 \\ B_1 \\ R_2 \\ G_2 \\ B_2 \\ \vdots \\ R_N \\ G_N \\ B_N \end{array} \\ \mathbf{B} \end{array} = \begin{array}{c} \underbrace{\hspace{1em}}_1 \\ \left. \begin{array}{c|c} \hline l_{R,1} \\ \hline l_{G,1} \\ \hline l_{B,1} \\ \hline \vdots \\ \hline l_{R,N} \\ \hline l_{G,N} \\ \hline l_{B,N} \\ \hline \end{array} \right\} \begin{array}{l} R_1 \\ G_1 \\ B_1 \\ R_2 \\ G_2 \\ B_2 \\ \vdots \\ R_N \\ G_N \\ B_N \end{array} \\ \tilde{\mathbf{L}} \end{array} \quad 1 \quad \left\{ \begin{array}{c} \underbrace{\hspace{10em}}_P \\ \begin{array}{c|c|c|c} \hline 1 & 1 & \dots & 1 \\ \hline \end{array} \\ \mathbf{W}_L \end{array} \right.$$

For simplification we will write equations only for the 1st value of intensity for the 1st image channel (e.g., red), $l_{R,1} \mathbf{W}_L = b_{1,j}$, $j = 1, \dots, P$. The same calculations can

be made for all other images channels. Loss function with non-negativity constraint for one resulting light intensity value is the following:

$$\min_{l_{R,1}} f = \sum_{j=1}^P (b_{1,j} - l_{R,1} \mathbf{W}_L)^2, \quad \text{subject to } 0 > l_{R,1} > H.$$

The general representation of KKT conditions is the following:

$$\begin{cases} \frac{\partial f(x)}{\partial x} + \sum \lambda_i \frac{\partial g_i(x)}{\partial x} \geq 0, \\ \lambda_i (b_i - g_i(x)) = 0, \\ (\frac{\partial f(x)}{\partial x} + \sum \lambda_i \frac{\partial g_i(x)}{\partial x}) x = 0, \\ \lambda_i \geq 0. \end{cases}$$

For simplification of notation let us have the following change of variable $l_{R,1} = l$. The supplementary matrix elements \mathbf{W}_L are equal to ones. According to our loss function and constraints representation: $x = l$, $\frac{\partial f}{\partial l} = -2 \sum_{j=1}^P (b_{1,j} - l)$ and $\frac{\partial g}{\partial l} = H$. KKT conditions can be rewritten as:

$$\begin{cases} -2 \sum_{j=1}^P (b_{1,j} - l) + \lambda H \geq 0, \\ \lambda (H - l) = 0, \\ (-2 \sum_{j=1}^P (b_{1,j} - l) + \lambda H) l = 0, \\ \lambda \geq 0. \end{cases}$$

We should test different sign-value combinations of λ and l , to define the cases for which KKT conditions are hold and when they are not hold. The cases, when they are hold will give us solutions of our minimization task.

$$\begin{aligned} \text{Case 1, } \lambda > 0, l = 0: & \begin{cases} \lambda H \geq 2 \sum_{j=1}^P b_{1,j}, \\ \lambda = 0, \quad \text{false.} \end{cases} \\ \text{Case 2, } \lambda > 0, l > 0: & \begin{cases} \lambda H \geq 2 \sum_{j=1}^P (b_{1,j} - l), \\ H - l = 0 \rightarrow l = H, \quad \text{true.} \end{cases} \\ \text{Case 3, } \lambda = 0, l > 0: & \end{aligned}$$

$$\begin{aligned} (-2 \sum_{j=1}^P (b_{1,j} - l)) l &= 0, \\ -2 \sum_{j=1}^P (b_{1,j} - l) &= 0, \\ l &= \frac{\sum_{j=1}^P b_{1,j}}{P}, \quad \text{true.} \end{aligned}$$

Case 4, $\lambda = 0$, $l = 0$:

$$\begin{aligned} -2 \sum_{j=1}^P b_{1,j} &\geq 0, \\ \sum_{j=1}^P b_{1,j} &\leq 0, \quad \text{true.} \end{aligned}$$

Due to different cases we obtain three different solutions:

- 1) For lower bound Case 4 is used: when $\sum_{j=1}^P b_{1,j} \leq 0$ then $l = 0$.
- 2) For upper bound Case 2 is used: when $\sum_{i=1}^N b_{1,i} \leq H$ then $l = H$.
- 3) For all other situations Case 3 is used: $l = \frac{\sum_{j=1}^P b_{1,j}}{P}$.

C.3 Lighting sources directions approximation with norm constraint

$$\begin{aligned} \min_{\tilde{\mathbf{V}}} & \left\| \mathbf{M} - f(\tilde{\mathbf{A}}, \tilde{\mathbf{L}}, \tilde{\mathbf{V}}, \mathbf{N}) \right\|_2^2 \\ \text{subject to} & v_{i,x}^2 + v_{i,y}^2 + v_{i,z}^2 = 1, \\ & \text{where } i = 1, \dots, N; f(\tilde{\mathbf{A}}, \tilde{\mathbf{L}}, \tilde{\mathbf{V}}, \mathbf{N}) \text{ denotes the model.} \end{aligned}$$

Let us have a change of variables $\mathbf{F} = (\mathbf{W}_A \tilde{\mathbf{A}}) \odot (\tilde{\mathbf{L}} \mathbf{W}_L)$, and matrix

$\mathbf{B} = ((\mathbf{W}_V^T \mathbf{W}_V)^{-1} \mathbf{W}_V^T \mathbf{M}) \odot \mathbf{F}$. Schematically we obtain the following system of equations in matrix form:

$$\begin{array}{c} \left. \begin{array}{|c|} \hline \overbrace{\begin{array}{cccc} b_{1,1} & b_{1,2} & \dots & b_{1,P} \\ b_{2,1} & b_{2,2} & \dots & b_{2,P} \\ \vdots & \vdots & \vdots & \vdots \\ b_{N,1} & b_{N,2} & \dots & b_{N,P} \end{array}}^P \\ \hline \end{array} \right\} N \quad = \quad \left. \begin{array}{|c|} \hline \overbrace{\begin{array}{ccc} v_{1,x} & v_{1,y} & v_{1,z} \\ v_{2,x} & v_{2,y} & v_{2,z} \\ \vdots & \vdots & \vdots \\ v_{N,x} & v_{N,y} & v_{N,z} \end{array}}^3 \\ \hline \end{array} \right\} N \quad \approx \quad \left. \begin{array}{|c|} \hline \overbrace{\begin{array}{cccc} n_{x,1} & n_{x,2} & \dots & n_{x,P} \\ n_{y,1} & n_{y,2} & \dots & n_{y,P} \\ n_{z,1} & n_{z,2} & \dots & n_{z,P} \end{array}}^P \\ \hline \end{array} \right\} N \end{array}$$

For simplification we will write equations only for the lighting source direction of one image, $\mathbf{v}_1^T \mathbf{n}_j = \mathbf{b}_{1,j}$. The same calculations can be made for all other lighting source directions used during initial images acquisition. Loss function with Lagrange multipliers of the norm constraint for one resulting illumination direction is the following:

$$e = \sum_{j=1}^P (v_{1,x}n_{x,j} + v_{1,y}n_{y,j} + v_{1,z}n_{z,j} - b_{1,j})^2 + \lambda(v_{1,x}^2 + v_{1,y}^2 + v_{1,z}^2 - 1).$$

In order to find the final solution of our minimization problem we should calculate partial derivatives of the loss function from all possible variables $(v_{1,x}, v_{1,y}, v_{1,z})$, including Lagrange multiplier (λ) :

$$\begin{cases} 2 \sum_{j=1}^P (v_{1,x}n_{x,j} + v_{1,y}n_{y,j} + v_{1,z}n_{z,j} - b_{1,j})n_{x,j} + 2\lambda v_{1,x} = 0, \\ 2 \sum_{j=1}^P (v_{1,x}n_{x,j} + v_{1,y}n_{y,j} + v_{1,z}n_{z,j} - b_{1,j})n_{y,j} + 2\lambda v_{1,y} = 0, \\ 2 \sum_{j=1}^P (v_{1,x}n_{x,j} + v_{1,y}n_{y,j} + v_{1,z}n_{z,j} - b_{1,j})n_{z,j} + 2\lambda v_{1,z} = 0, \\ v_{1,x}^2 + v_{1,y}^2 + v_{1,z}^2 - 1 = 0. \end{cases}$$

$$\begin{cases} v_{1,x}(\sum_{j=1}^P n_{x,j}^2 + \lambda) + v_{1,y} \sum_{j=1}^P n_{y,j}n_{x,j} + v_{1,z} \sum_{j=1}^P n_{z,j}n_{x,j} = \sum_{j=1}^P b_{1,j}n_{x,j}, \\ v_{1,x} \sum_{j=1}^P n_{x,j}n_{y,j} + v_{1,y}(\sum_{j=1}^P n_{y,j}^2 + \lambda) + v_{1,z} \sum_{j=1}^P n_{z,j}n_{y,j} = \sum_{j=1}^P b_{1,j}n_{y,j}, \\ v_{1,x} \sum_{j=1}^P n_{x,j}n_{z,j} + v_{1,y} \sum_{j=1}^P n_{y,j}n_{z,j} + v_{1,z}(\sum_{j=1}^P n_{z,j}^2 + \lambda) = \sum_{j=1}^P b_{1,j}n_{z,j}, \\ v_{1,x}^2 + v_{1,y}^2 + v_{1,z}^2 - 1 = 0. \end{cases}$$

Let us have the following change of variables:

$$\mathbf{Y} = \begin{bmatrix} y_{1,1} & y_{1,2} & y_{1,3} \\ y_{2,1} & y_{2,2} & y_{2,3} \\ y_{3,1} & y_{3,2} & y_{3,3} \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^P n_{x,j}^2 & \sum_{j=1}^P n_{x,j}n_{y,j} & \sum_{j=1}^P n_{x,j}n_{z,j} \\ \sum_{j=1}^P n_{y,j}n_{x,j} & \sum_{j=1}^P n_{y,j}^2 & \sum_{j=1}^P n_{y,j}n_{z,j} \\ \sum_{j=1}^P n_{z,j}n_{x,j} & \sum_{j=1}^P n_{z,j}n_{y,j} & \sum_{j=1}^P n_{z,j}^2 \end{bmatrix},$$

$$\mathbf{d} = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^P b_{1,j}n_{x,j} \\ \sum_{j=1}^P b_{1,j}n_{y,j} \\ \sum_{j=1}^P b_{1,j}n_{z,j} \end{bmatrix}.$$

Then our system of equations can be rewritten as:

$$\begin{cases} v_{1,x}(y_{1,1} + \lambda) + v_{1,y}y_{1,2} + v_{1,z}y_{1,3} = d_1, \\ v_{1,x}y_{2,1} + v_{1,y}(y_{2,1} + \lambda) + v_{1,z}y_{2,3} = d_2, \\ v_{1,x}y_{3,1} + v_{1,y}y_{3,2} + v_{1,z}(y_{3,3} + \lambda) = d_3, \\ v_{1,x}^2 + v_{1,y}^2 + v_{1,z}^2 - 1 = 0. \end{cases}$$

Or in matrix form we can present this system as:

$$\begin{cases} (\mathbf{Y} + \lambda \mathbf{I}_3) \mathbf{v} = \mathbf{d}, \\ \mathbf{v}^T \mathbf{v} = 1. \end{cases}$$

where \mathbf{I}_3 signifies identity matrix of 3 by 3 size and $\mathbf{v} = [v_{1,x}, v_{1,y}, v_{1,z}]^T$.

This system of equations contains quadratic components and can be solved by the means of nonlinear least squares minimization using the Levenberg-Marquardt algorithm. The general case of this algorithm is presented in Appendix D.

C.4 Surface normals matrix approximation with norm, direction and smoothness constraints

$$\begin{aligned} & \min_{\mathbf{N}} \left\| \mathbf{M} - f(\tilde{\mathbf{A}}, \tilde{\mathbf{L}}, \tilde{\mathbf{V}}, \mathbf{N}) \right\|_2^2 \\ & \text{subject to } \begin{cases} n_{x,j}^2 + n_{y,j}^2 + n_{z,j}^2 = 1, \\ \sum_{k=-1}^1 \sum_{l=-1}^1 \mathbf{n}^T \mathbf{n}^{(k,l)} = 8, \end{cases} \end{aligned}$$

where $j = 1, \dots, P$; $\mathbf{n}^{(k,l)}$ is the j th pixel neighbors in 2D image space; $f(\tilde{\mathbf{A}}, \tilde{\mathbf{L}}, \tilde{\mathbf{V}}, \mathbf{N})$ defines the model.

Let us have a change of variables $\mathbf{F} = (\mathbf{W}_A \tilde{\mathbf{A}}) \odot (\tilde{\mathbf{L}} \mathbf{W}_L)$, and matrix $\mathbf{B} = ((\mathbf{W}_V^T \mathbf{W}_V)^{-1} \mathbf{W}_V^T \mathbf{M}) \odot \mathbf{F}$. Schematically we obtain the following system of equations in matrix form:

$$\begin{array}{c} \begin{array}{c} \underbrace{\hspace{10em}}_P \\ \begin{array}{|c|c|c|c|} \hline b_{1,1} & b_{1,2} & \dots & b_{1,P} \\ \hline b_{2,1} & b_{2,2} & \dots & b_{2,P} \\ \hline \vdots & \vdots & \vdots & \vdots \\ \hline b_{N,1} & b_{N,2} & \dots & b_{N,P} \\ \hline \end{array} \\ \underbrace{\hspace{10em}}_N \\ \mathbf{B} \end{array} \quad = \quad \begin{array}{c} \begin{array}{c} \underbrace{\hspace{3em}}_3 \\ \begin{array}{|c|c|c|} \hline v_{1,x} & v_{1,y} & v_{1,z} \\ \hline v_{2,x} & v_{2,y} & v_{2,z} \\ \hline \vdots & \vdots & \vdots \\ \hline v_{N,x} & v_{N,y} & v_{N,z} \\ \hline \end{array} \\ \underbrace{\hspace{3em}}_N \\ \tilde{\mathbf{V}}^T \end{array} \quad \approx \quad \begin{array}{c} \begin{array}{c} \underbrace{\hspace{10em}}_P \\ \begin{array}{|c|c|c|c|} \hline n_{x,1} & n_{x,2} & \dots & n_{x,P} \\ \hline n_{y,1} & n_{y,2} & \dots & n_{y,P} \\ \hline n_{z,1} & n_{z,2} & \dots & n_{z,P} \\ \hline \end{array} \\ \underbrace{\hspace{10em}}_N \\ \mathbf{N} \end{array} \end{array}$$

For simplification let us present a solution only for one pixel 3D normal. They same optimization procedure should be applied to each pixel of image matrix. Minimization function for finding $(n_{x,1}, n_{y,1}, n_{z,1})^T = (n_x, n_y, n_z)^T$ can be written as:

$$e = \sum_{i=1}^N (\mathbf{v}_i^T \mathbf{n} - b_{i,1})^2 + \lambda_1 (n_x^2 + n_y^2 + n_z^2 - 1) + \lambda_2 \left(\sum_{k=-1}^1 \sum_{l=-1}^1 (n_x n_x^{(k,l)} + n_y n_y^{(k,l)} + n_z n_z^{(k,l)}) - 8 \right),$$

where λ_1 and λ_2 denote Lagrange multipliers. In order to find the final solution of our minimization problem we should calculate partial derivatives of the loss function from all possible variables (n_x, n_y, n_z) , including Lagrange multipliers (λ_1, λ_2) :

$$\left\{ \begin{array}{l} \frac{\partial e}{\partial n_x} = 2 \sum_{i=1}^N (v_{i,x} n_x + v_{i,y} n_y + v_{i,z} n_z - b_{i,1}) v_{i,x} + 2\lambda_1 n_x + \lambda_2 \sum_{k=-1}^1 \sum_{l=-1}^1 n_x^{(k,l)}, \\ \frac{\partial e}{\partial n_y} = 2 \sum_{i=1}^N (v_{i,x} n_x + v_{i,y} n_y + v_{i,z} n_z - b_{i,1}) v_{i,y} + 2\lambda_1 n_y + \lambda_2 \sum_{k=-1}^1 \sum_{l=-1}^1 n_y^{(k,l)}, \\ \frac{\partial e}{\partial n_z} = 2 \sum_{i=1}^N (v_{i,x} n_x + v_{i,y} n_y + v_{i,z} n_z - b_{i,1}) v_{i,z} + 2\lambda_1 n_z + \lambda_2 \sum_{k=-1}^1 \sum_{l=-1}^1 n_z^{(k,l)}, \\ \frac{\partial e}{\partial \lambda_1} = n_x^2 + n_y^2 + n_z^2 - 1, \\ \frac{\partial e}{\partial \lambda_2} = \sum_{k=-1}^1 \sum_{l=-1}^1 (n_x n_x^{(k,l)} + n_y n_y^{(k,l)} + n_z n_z^{(k,l)}) - 8. \end{array} \right.$$

Let us have the following change of variables: $g_x = \frac{1}{2} \sum_{k=-1}^1 \sum_{l=-1}^1 n_x^{(k,l)}$,

$$g_y = \frac{1}{2} \sum_{k=-1}^1 \sum_{l=-1}^1 n_y^{(k,l)}, \text{ and } g_z = \frac{1}{2} \sum_{k=-1}^1 \sum_{l=-1}^1 n_z^{(k,l)}.$$

$$\left\{ \begin{array}{l} (n_x \sum_{i=1}^N v_{i,x} + n_y \sum_{i=1}^N v_{i,y} + n_z \sum_{i=1}^N v_{i,z} - \sum_{i=1}^N b_{i,1}) v_{i,x} + n_x \lambda_1 + \lambda_2 g_x = 0, \\ (n_x \sum_{i=1}^N v_{i,x} + n_y \sum_{i=1}^N v_{i,y} + n_z \sum_{i=1}^N v_{i,z} - \sum_{i=1}^N b_{i,1}) v_{i,y} + n_y \lambda_1 + \lambda_2 g_y = 0, \\ (n_x \sum_{i=1}^N v_{i,x} + n_y \sum_{i=1}^N v_{i,y} + n_z \sum_{i=1}^N v_{i,z} - \sum_{i=1}^N b_{i,1}) v_{i,z} + n_z \lambda_1 + \lambda_2 g_z = 0, \\ n_x^2 + n_y^2 + n_z^2 = 1, \\ n_x g_x + n_y g_y + n_z g_z = 4. \end{array} \right.$$

$$\left\{ \begin{array}{l} n_x \sum_{i=1}^N v_{i,x}^2 + n_y \sum_{i=1}^N v_{i,x} v_{i,y} + n_z \sum_{i=1}^N v_{i,x} v_{i,z} + n_x \lambda_1 + \lambda_2 g_x = \sum_{i=1}^{3N} b_{i,1} v_{i,x}, \\ n_x \sum_{i=1}^N v_{i,x} v_{i,y} + n_y \sum_{i=1}^N v_{i,y}^2 + n_z \sum_{i=1}^N v_{i,y} v_{i,z} + n_y \lambda_1 + \lambda_2 g_y = \sum_{i=1}^{3N} b_{i,1} v_{i,y}, \\ n_x \sum_{i=1}^N v_{i,x} v_{i,z} + n_y \sum_{i=1}^N v_{i,y} v_{i,z} + n_z \sum_{i=1}^N v_{i,z}^2 + n_z \lambda_1 + \lambda_2 g_z = \sum_{i=1}^{3N} b_{i,1} v_{i,z}, \\ n_x^2 + n_y^2 + n_z^2 = 1, \\ n_x g_x + n_y g_y + n_z g_z = 4. \end{array} \right.$$

$$\begin{cases} n_x(\sum_{i=1}^N v_{i,x}^2 + \lambda_1) + n_y \sum_{i=1}^N v_{i,x}v_{i,y} + n_z \sum_{i=1}^N v_{i,x}v_{i,z} + \lambda_2 g_x = \sum_{i=1}^{3N} b_{i,1}v_{i,x}, \\ n_x \sum_{i=1}^N v_{i,x}v_{i,y} + n_y(\sum_{i=1}^N v_{i,y}^2 + \lambda_1) + n_z \sum_{i=1}^N v_{i,y}v_{i,z} + \lambda_2 g_y = \sum_{i=1}^{3N} b_{i,1}v_{i,y}, \\ n_x \sum_{i=1}^N v_{i,x}v_{i,z} + n_y \sum_{i=1}^N v_{i,y}v_{i,z} + n_z(\sum_{i=1}^N v_{i,z}^2 + \lambda_1) + \lambda_2 g_z = \sum_{i=1}^{3N} b_{i,1}v_{i,z}, \\ n_x^2 + n_y^2 + n_z^2 = 1, \\ n_x g_x + n_y g_y + n_z g_z = 4. \end{cases}$$

In order to solve this system of equation let us pass to a matrix form representation with the following change of variables:

$$\mathbf{Y} = \begin{bmatrix} y_{1,1} & y_{1,2} & y_{1,3} \\ y_{2,1} & y_{2,2} & y_{2,3} \\ y_{3,1} & y_{3,2} & y_{3,3} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^N v_{i,x}^2 & \sum_{i=1}^N v_{i,x}v_{i,y} & \sum_{i=1}^N v_{i,x}v_{i,z} \\ \sum_{i=1}^N v_{i,y}v_{i,z} & \sum_{i=1}^N v_{i,y}^2 & \sum_{i=1}^N v_{i,y}v_{i,z} \\ \sum_{i=1}^N v_{i,z}v_{i,x} & \sum_{i=1}^N v_{i,z}v_{i,y} & \sum_{i=1}^N v_{i,z}^2 \end{bmatrix},$$

$$\mathbf{d} = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^N b_{i,1}v_{i,x} \\ \sum_{i=1}^N b_{i,1}v_{i,y} \\ \sum_{i=1}^N b_{i,1}v_{i,z} \end{bmatrix}.$$

$$\mathbf{g} = \begin{bmatrix} g_x \\ g_y \\ g_z \end{bmatrix}.$$

If \mathbf{I}_3 is an identity matrix of $[3 \times 3]$ size, then the system of equations has the following form:

$$\begin{cases} \mathbf{Y}\mathbf{n} + \lambda_1 \mathbf{I}_3 \mathbf{n} = \mathbf{d} - \lambda_2 \mathbf{g}, \\ \mathbf{n}^T \mathbf{I}_3 \mathbf{n} = 1, \\ \mathbf{g}^T \mathbf{n} = 4. \end{cases}$$

$$\begin{bmatrix} \mathbf{Y}\mathbf{n} + \lambda_1 \mathbf{I}_3 \mathbf{n} \\ \mathbf{n}^T \mathbf{I}_3 \\ \mathbf{g}^T \end{bmatrix} \mathbf{n} = \begin{bmatrix} \mathbf{d} - \lambda_2 \mathbf{g} \\ 1 \\ 4 \end{bmatrix}.$$

This system of equations contains quadratic components and can be solved by the means of nonlinear least squares minimization using the Levenberg-Marquardt algorithm. The general case of this algorithm is presented in Appendix D.

Appendix D

The Levenberg-Marquardt Algorithm

The Levenberg-Marquardt algorithm (LVA) was firstly proposed by K. Levenberg [115] and then improved by D. Marquardt [118]. It is one of the mostly used iterative optimization algorithms for the non-linear problems solution. It can be seen as a mixture between simplicity of the basic gradient descent method and convergence of the Gauss-Newton iterative method, [119].

Let us have the function to be minimized $f(\mathbf{x}) = \sum_{j=1}^m y_j^2(\mathbf{x})$, where $\mathbf{x} = [x_1, x_2, \dots, x_n]$ and $m \geq n$. So, $y(\mathbf{x}) = [y_1(\mathbf{x}), y_2(\mathbf{x}), \dots, y_m(\mathbf{x})]$, and the initial problem becomes $f(\mathbf{x}) = \|y(\mathbf{x})\|_2^2$.

The set of partial derivatives of $f(\mathbf{x})$ can be written as the Jacobian matrix \mathbf{J} :

$$\mathbf{J} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \dots & \frac{\partial y_1}{\partial x_n} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \dots & \frac{\partial y_2}{\partial x_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial y_m}{\partial x_1} & \frac{\partial y_m}{\partial x_2} & \dots & \frac{\partial y_m}{\partial x_n} \end{bmatrix}.$$

In order to solve the quadratic problem let us find the first and the second derivatives:

$$\begin{aligned} \nabla f(\mathbf{x}) &= \sum_{j=1}^m y_j(\mathbf{x}) \nabla y_j(\mathbf{x}) = \mathbf{J}(\mathbf{x})^T y(\mathbf{x}), \\ \nabla^2 f(\mathbf{x}) &= \mathbf{J}(\mathbf{x})^T \mathbf{J}(\mathbf{x}) + \sum_{j=1}^m y_j(\mathbf{x}) \nabla^2 y_j(\mathbf{x}). \end{aligned}$$

If $\nabla^2 y_j(\mathbf{x})$ or $y_j(\mathbf{x})$ are small, the the Hessian can be simply evaluated as:

$$\mathbf{H}(\mathbf{x}) = \nabla^2 f(\mathbf{x}) = \mathbf{J}(\mathbf{x})^T \mathbf{J}(\mathbf{x}). \quad (\text{D.1})$$

Due to the LMA the resulting vector \mathbf{x} on the $(i + 1)$ iteration can be found as:

$$\mathbf{x}_{i+1} = \mathbf{x}_i - (\mathbf{H}(\mathbf{x}_i) + \lambda \text{diag}[\mathbf{H}(\mathbf{x}_i)])^{-1} \nabla f(\mathbf{x}_i), \quad (\text{D.2})$$

where \mathbf{H} is evaluated as in Eq.D.1 and λ denotes the step of descent as in simple gradient descent method.

Coefficient λ can be found using the gradient or steepest descent algorithm, [119].

Bibliography

- [1] bioMérieux. bioMérieux France, 2014. URL <http://www.biomerieux.fr/>.
- [2] L. Allano, L. Jacolin, D. Decaux, F. Pinston, D. Desseree, G. Boissier, C. Tachier, C. Fulchiron, L. Lapierre, and G. Sutra. Bio-imaging method and system, May 2011.
- [3] T. Moons, L.J. Van Gool, and M. Vergauwen. 3D reconstruction from multiple images: part 1 - principles. *Foundations and Trends in Computer Graphics and Vision*, 4(4):287–404, 2009.
- [4] M.Z. Brown, D. Burschka, and G.D. Hager. Advances in computational stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(8):993–1008, 2003.
- [5] R. Rosenholtz and J. Malik. Surface orientation from texture: Isotropy or homogeneity (or both)? *Vision Research*, 37:2283–2293, 1997.
- [6] A.M. Loh and R. Hartley. Shape from non-homogeneous, non-stationary, anisotropic, perspective texture. In *British Machine Vision Conference*, 2005.
- [7] A.P. Witkin. *Shape from contour*. AI-TR. Artificial Intelligence Laboratory, Massachusetts Institute of Technology, 1980.
- [8] M. Brady and A.L. Yuille. An extremum principle for shape from contour. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(3):288–301, 1984.
- [9] F. Ulupinar and R. Nevatia. Shape from contour: Straight homogeneous generalized cylinders and constant cross section generalized cylinders. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(2):120–135, 1995.
- [10] S.K. Nayar and Y. Nakagawa. Shape from focus. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(8):824–831, 1994.

- [11] S. Pertuz, D. Puig, and M.A. García. Analysis of focus measure operators for shape-from-focus. *Pattern Recognition*, 46(5):1415–1432, 2013.
- [12] I. Shlyakhter, M. Rozenoer, J. Dorsey, and S. Teller. Reconstructing 3D tree models from instrumented photographs. *IEEE Computer Graphics and Applications*, 21(3):53–61, 2001.
- [13] Y. Furukawa and J. Ponce. Carved visual hulls for image-based modeling. In *European Conference on Computer Vision*, pages 564–577, 2006.
- [14] Z. Song and Y. Shing-Tung. High-speed three-dimensional shape measurement system using a modified two-plus-one phase-shifting algorithm. *Optical Engineering*, 46(11), 2007.
- [15] Kinect for Windows. Microsoft corporation, 2013. URL <http://www.microsoft.com/en-us/kinectforwindows/>.
- [16] O. Aubreton, A. Bajard, B. Verney, and F. Truchetet. Infrared system for 3D scanning of metallic surfaces. *Machine Vision and Applications*, 24(7):1513–1524, 2013.
- [17] M. Gupta, A. Agrawal, A. Veeraraghavan, and S.G. Narasimhan. A practical approach to 3D scanning in the presence of interreflections, subsurface scattering and defocus. *International Journal of Computer Vision*, 102(1-3):33–55, 2013.
- [18] B.K.P. Horn. Obtaining shape from shading information. *The Psychology of Computer Vision*, pages 115–155, 1975.
- [19] B.K.P. Horn. Understanding image intensities. *Artificial Intelligence*, 8(2):201–231, 1977.
- [20] R. Zhang, P.-S. Tsai, J.E. Cryer, and M. Shah. Shape from shading: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(8):690–706, 1999.
- [21] J.T. Barron and J. Malik. Shape, albedo, and illumination from a single image of an unknown object. In *Computer Vision and Pattern Recognition*, 2012.
- [22] J.T. Barron and J. Malik. Color constancy, intrinsic images, and shape estimation. In *European Conference on Computer Vision*, 2012.

- [23] J.T. Barron and J. Malik. Shape, illumination, and reflectance from shading. Technical Report UCB/EECS-2013-117, EECS, UC Berkeley, May 2013.
- [24] R.J. Woodham. Photometric method for determining surface orientation from multiple images. *Optical Engineering*, 19(1):139–144, 1980.
- [25] C.L. Lawson and R.J. Hanson. *Solving least squares problems*. Englewood Cliffs: Prentice-Hall, 3^d edition, 1995.
- [26] P.N. Belhumeur, D.J. Kriegman, and A.L. Yuille. The bas-relief ambiguity. *International Journal of Computer Vision*, 35(1):33–44, 1999.
- [27] S.K. Nayar, K. Ikeuchi, and T. Kanade. Shape from interreflections. *International Journal of Computer Vision*, 6(3):173–195, 1991.
- [28] T. Papadhimetri and P. Favaro. A new perspective on uncalibrated photometric stereo. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1474–1481, 2013.
- [29] H.W. Jensen, S.R. Marschner, M. Levoy, and P. Hanrahan. A practical model for subsurface light transport. In *Special Interest Group in GRAPHics*, pages 511–518, 2001.
- [30] J.L. Shumaker. A survey of photon mapping for realistic image synthesis. Technical Report ARL-TR-3608, U.S. Army Research Laboratory, September 2005.
- [31] N.G. Aldrin, T. Zickler, and D.J. Kriegman. Photometric stereo with non-parametric and spatially-varying reflectance. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2008.
- [32] H. Murase. Surface shape reconstruction of a nonrigid transport object using refraction and motion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(10):1045–1052, 1992.
- [33] W. Matusik, H. Pfister, R. Ziegler, A. Ngan, and L. McMillan. Acquisition and rendering of transparent and refractive objects. In *Eurographics Workshop on Rendering*, pages 267–278, 2002.
- [34] D. Miyazaki and K. Ikeuchi. Inverse polarization raytracing: Estimating surface shapes of transparent objects. In *IEEE Computer Vision and Pattern Recognition*, pages 910–917, 2005.

- [35] M. Saito, Y. Sato, K. Ikeuchi, and H. Kashiwagi. Measurement of surface orientations of transparent objects using polarization in highlight. *Systems and Computers in Japan*, 32(5):64–71, 2001.
- [36] G.A. Atkinson and E.R. Hancock. Shape estimation using polarization and shading from two views. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(11):2001–2017, 2007.
- [37] T. Chen, H.P.A. Lensch, C. Fuchs, and H.-P. Seidel. Polarization and phase-shifting for 3D scanning of translucent objects. In *IEEE Computer Vision and Pattern Recognition*, 2007.
- [38] B. Trifonov, D. Bradley, and W. Heidrich. Tomographic reconstruction of transparent objects. In *Rendering Techniques*, pages 51–60. Eurographics Association, 2006.
- [39] N.J.W. Morris and K.N. Kutulakos. Reconstructing the surface of inhomogeneous transparent scenes by scatter-trace photography. In *IEEE International Conference on Computer Vision*, pages 1–8, 2007.
- [40] R. Rantson, C. Stolz, D. Fofi, and F. Meriaudeau. Non contact 3D measurement scheme for transparent objects using UV structured light. In *International Conference on Pattern Recognition*, pages 1646–1649, 2010.
- [41] S.A. Shafer. Using color to separate reflection components. Technical Report TR136, Computer Science Department, University of Rochester, April 1984.
- [42] J.F. Blinn. Models of light reflection for computer synthesized pictures. In *Conference on Computer Graphics and Interactive Techniques*, pages 192–198, 1977.
- [43] G.J. Ward. Measuring and modeling anisotropic reflection. In *Conference on Computer Graphics and Interactive Techniques*, pages 265–272, 1992.
- [44] P. Beckmann and A. Spizzichino. *The Scattering of Electromagnetic Waves from Rough Surfaces*. Artech Print on Demand, 1987.
- [45] R.L. Cook and K.E. Torrance. A reflectance model for computer graphics. *Special Interest Group in GRAPHics*, 15(3):307–316, 1982.
- [46] C. Wynn. An introduction to BRDF-based lighting. Technical report, NVIDIA Corporation, 2000.

- [47] T. Zickler, P.N. Belhumeur, and D.J. Kriegman. Helmholtz stereopsis: Exploiting reciprocity for surface reconstruction. *International Journal of Computer Vision*, 49(2-3):215–227, 2002.
- [48] S.C. Foo. A gonioreflectometer for measuring the bidirectional reflectance of materials for use in illumination computations. Master’s thesis, Cornell University, 1997.
- [49] B.T. Phong. Illumination for computer generated pictures. *Communications of the ACM*, 18(6):311–317, 1975.
- [50] M. Oren and S.K. Nayar. Generalization of Lambert’s reflectance model. In *Special interest group in GRAPHics*, pages 239–246, 1994.
- [51] I. Ihrke, K.N. Kutulakos, H.P.A. Lensch, M.A. Magnor, and W. Heidrich. Transparent and specular object reconstruction. *Computer Graphics Forum*, 29(8):2400–2426, 2010.
- [52] A. Zisserman, P. Giblin, and A. Blake. The information available to a moving observer from specularities. *Image and Vision Computing*, 7(1):38–42, 1989.
- [53] J.Y. Zheng and A. Murata. Acquiring a complete 3D model from specular motion under the illumination of circular-shaped light sources. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8):913–920, 2000.
- [54] S. Magda and D.J. Kriegman. Beyond Lambert: Reconstructing surfaces with arbitrary BRDFs. In *IEEE International Conference on Computer Vision*, pages 391–398, 2001.
- [55] L. Wang, R. Yang, and J.E. Davis. BRDF invariant stereo using light transport constancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(9):1616–1626, 2007.
- [56] S. Savarese and P. Perona. Local analysis for 3D reconstruction of specular surfaces. In *IEEE Computer Vision and Pattern Recognition*, volume 2, pages 738–745, 2001.
- [57] S. Savarese, M. Chen, and P. Perona. Local shape from mirror reflections. *International Journal of Computer Vision*, 64(1):31–67, 2005.
- [58] M. Tarini, H.P.A. Lensch, M. Goesele, and H.-P. Seidel. 3D acquisition of mirroring objects using striped patterns. *Graphical Models*, 67(4):233–259, 2005.

- [59] D.B. Goldman, B. Curless, A. Hertzmann, and S.M. Seitz. Shape and spatially-varying BRDFs from photometric stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(6):1060–1071, 2010.
- [60] A. Hertzmann and S.M. Seitz. Example-based photometric stereo: shape reconstruction with general, varying BRDFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(8):1254–1264, 2005.
- [61] J. Wang and K.J. Dana. Relief texture from specularities. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(3):446–457, 2006.
- [62] K.J. Dana. BRDF/BTF measurement device. In *IEEE International Conference on Computer Vision*, pages 460–466, 2001.
- [63] S. Roth and M.J. Black. Specular flow and the recovery of surface structure. In *IEEE Computer Vision and Pattern Recognition*, pages 1869–1876, 2006.
- [64] Y. Adato, Y. Vasilyev, O. Ben-Shahar, and T. Zickler. Toward a theory of shape from specular flow. In *IEEE International Conference on Computer Vision*, pages 1–8, 2007.
- [65] T. Higo, Y. Matsushita, and K. Ikeuchi. Consensus photometric stereo. In *IEEE Computer Vision and Pattern Recognition*, pages 1157–1164, 2010.
- [66] G. Healey and T.O. Binford. Local shape from specularity. *Computer Vision, Graphics, and Image Processing*, 42(1):62–86, 1988.
- [67] E.N. Coleman and R. Jain. Obtaining 3-dimensional shape of textured and specular surfaces using four-source photometry. *Computer Graphics and Image Processing*, 18:309–328, 1982.
- [68] S. Barsky and M. Petrou. The 4-source photometric stereo technique for three-dimensional surfaces in the presence of highlights and shadows. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(10):1239–1252, 2003.
- [69] V. Argyriou, M. Petrou, and S. Barsky. Photometric stereo with an arbitrary number of illuminants. *Computer Vision and Image Understanding*, 114(8):887–900, 2010.
- [70] H.E. Rushmeier, G. Taubin, and A. Guézic. Applying shape from lighting variation to bump map capture. In *Eurographics Workshop*, pages 35–44, 1997.

- [71] J. Sun, M.L. Smith, L.N. Smith, S. Midha, and J. Bamber. Object surface recovery using a multi-light photometric stereo technique for non-Lambertian surfaces subject to shadows and specularities. *Image and Vision Computing*, 25(7):1050–1057, 2007.
- [72] S.P. Mallick, T. Zickler, D.J. Kriegman, and P.N. Belhumeur. Beyond Lambert: reconstructing specular surfaces using color. In *IEEE Computer Vision and Pattern Recognition*, pages 619–626, 2005.
- [73] A. Blake and H. Bülthoff. Shape from specularities: computation and psychophysics. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 331:237–252, 1991.
- [74] J.E. Solem, H. Aanæs, and A. Heyden. PDE based shape from specularities. In *International Conference on Scale Space Methods in Computer Vision*, pages 401–415, 2003.
- [75] M. Osadchy, D.W. Jacobs, and R. Ramamoorthi. Using specularities for recognition. In *IEEE International Conference on Computer Vision*, pages 1512–1519, 2003.
- [76] A. Artusi, F. Banterle, and D. Chetverikov. A survey of specular removal methods. *Computer Graphics Forum*, 30(8):2208–2230, 2011.
- [77] S.K. Nayar, X.-S. Fang, and T.E. Boult. Separation of reflection components using color and polarization. *International Journal of Computer Vision*, 21(3):163–186, 1997.
- [78] R. Feris, R. Raskar, K. Tan, and M. Turk. Specular reflection reduction with multi-flash imaging. In *IEEE Brazilian Symposium on Computer Graphics and Image Processing*, pages 316–321, 2004.
- [79] Y. Xu, F. Wang, and Y. Zhao. Matching based highlight removal. In *International Conference on Multimedia Technology*, pages 1–4, 2010.
- [80] P. Tan, S. Lin, L. Quan, and H.-Y. Shum. Highlight removal by illumination-constrained inpainting. In *IEEE International Conference on Computer Vision*, pages 164–169, 2003.
- [81] L.I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Journal of Physics D*, 60(1-4):259–268, 1992.

- [82] Q. Yang, S. Wang, and N. Ahuja. Real-time specular highlight removal using bilateral filtering. In *European Conference on Computer Vision*, pages 87–100, 2010.
- [83] R.T. Tan and K. Ikeuchi. Separating reflection components of textured surfaces using a single image. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 870–877, 2003.
- [84] H.-L. Shen and Q.-Y. Cai. Simple and efficient method for specular removal in an image. *Applied Optics*, 48:2711–2719, 2009.
- [85] C. Capanna, G. Gesquière, L. Jorda, P. Lamy, and D. Vibert. Three-dimensional reconstruction using multiresolution photoclinometry by deformation. *The Visual Computer*, 29(6-8):825–835, 2013.
- [86] I. Avcibas, B. Sankur, and K. Sayood. Statistical evaluation of image quality measures. *Journal of Electronic Imaging*, 11(2):206–223, 2002.
- [87] M.F. Hansen, G.A. Atkinson, L.N. Smith, and M.L. Smith. 3D face reconstructions from photometric stereo using near infrared and visible light. *Computer Vision and Image Understanding*, 114(8):942–951, 2010.
- [88] M.K. Johnson and E.H. Adelson. Retrographic sensing for the measurement of surface texture and shape. In *IEEE Computer Vision and Pattern Recognition*, pages 1070–1077, 2009.
- [89] W.M. Silver. Determining shape and reflectance using multiple images. Master’s thesis, Massachusetts Institute of Technology. Dept. of Electrical Engineering and Computer Science, 1980.
- [90] T.-P. Wu and C.-K. Tang. Photometric stereo via expectation maximization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(3):546–560, 2010.
- [91] S. Biswas, G. Aggarwal, and R. Chellappa. Robust estimation of albedo for illumination-invariant matching and shape recovery. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(5):884–899, 2009.
- [92] R.J. Woodham, Y. Iwahori, and R.A. Barman. Photometric stereo: Lambertian reflectance and light sources with unknown direction and strength. Technical report, Vancouver, Canada, 1991.

- [93] H. Hayakawa. Photometric stereo under a light source with arbitrary motion. *Journal of the Optical Society of America*, 11(11):3079–3089, 1994.
- [94] D. Miyazaki and K. Ikeuchi. Photometric stereo under unknown light sources using robust SVD with missing data. In *IEEE International Conference on Image Processing*, pages 4057–4060, 2010.
- [95] R. Basri, D.W. Jacobs, and I. Kemelmacher. Photometric stereo with general, unknown lighting. *International Journal of Computer Vision*, 72(3):239–257, 2007.
- [96] K. Kanatani. *Geometric computation for machine vision*. Oxford University Press, Inc., 1993.
- [97] J.C. Bezdek and R.J. Hathaway. Some notes on alternating optimization. In *The 2002 AFSS International Conference on Fuzzy Systems*, pages 288–300, 2002.
- [98] I. Csiszár and G. Tusnády. Information geometry and alternating minimization procedures. *Statistics & Decisions, Supplement Issue*, 1:205–237, 1984.
- [99] R.J. Hathaway, Y. Hu, and J.C. Bezdek. Local convergence of tri-level alternating optimization. *Neural Parallel & Scientific Comp.*, 9:19–28, 2001.
- [100] A. Cichocki, R. Zdunek, A.H. Phan, and S. Amari. *Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-way Data Analysis and Blind Source Separation*. Wiley, John & Sons, Incorporated, 1st edition, November 2009.
- [101] H. Lee, A. Battle, R. Raina, and A. Ng. Efficient sparse coding algorithms. In *Neural Information Processing Systems*, pages 801–808, 2007.
- [102] H.-S. Chung and J. Jia. Efficient photometric stereo on glossy surfaces with wide specular lobes. In *IEEE Computer Vision and Pattern Recognition*, 2008.
- [103] M.J. Brooks and B.K.P. Horn. Shape and source from shading. In *International Joint Conference on Artificial Intelligence*, pages 932–936, 1985.
- [104] Z. Zhou and P. Tan. Ring-light photometric stereo. In *European Conference on Computer Vision*, pages 265–279, 2010.

- [105] R.T. Frankot and R. Chellappa. A method for enforcing integrability in shape from shading algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(4):439–451, 1988.
- [106] P. Kovesi. Shapelets correlated with surface normals produce surfaces. In *IEEE International Conference on Computer Vision*, pages 994–1001, 2005.
- [107] A.K. Agrawal, R. Chellappa, and R. Raskar. An algebraic approach to surface reconstruction from gradient fields. In *IEEE International Conference on Computer Vision*, pages 174–181, 2005.
- [108] J. Draréni. Reconstruction par stéréo photométrie. Technical report, Université de Montréal, 2007.
- [109] D. Reddy, A.K. Agrawal, and R. Chellappa. Enforcing integrability by error correction using l1-minimization. In *IEEE Computer Vision and Pattern Recognition*, pages 2350–2357, 2009.
- [110] A. Cichocki and R. Zdunek. Regularized Alternating Least Squares algorithms for non-negative matrix/tensor Factorization. In *International Symposium on Neural Networks*, pages 793–802, 2007.
- [111] R. Grosse, M.K. Johnson, E.H. Adelson, and W.T. Freeman. Ground-truth dataset and baseline evaluations for intrinsic image algorithms. In *IEEE International Conference on Computer Vision*, pages 2335–2342, 2009.
- [112] N.G. Alldrin and D.J. Kriegman. Toward reconstructing surfaces with arbitrary isotropic reflectance : A stratified photometric stereo approach. In *IEEE International Conference on Computer Vision*, pages 1–8, 2007.
- [113] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, New York, NY, USA, 2004.
- [114] R.T. Rockafellar. Lagrange multipliers and optimality. *SIAM Review*, 35(2): 183–238, 1993.
- [115] K. Levenberg. A method for the solution of certain non-linear problems in least-squares. *Quarterly of Applied Mathematics*, 2(2):164–168, 1944.
- [116] J.M. Willey, L.M. Sherwood, and C.J. Woolverton. *Prescott's Microbiology*. McGraw-Hill, 8th edition, 2011.

-
- [117] T. Chan and L. Vese. An active contour model without edges. In *International Conference on Scale-Space Theories in Computer Vision*, pages 141–151, 1999.
- [118] D.W. Marquardt. An algorithm for the least-squares estimation of nonlinear parameters. *SIAM Journal of Applied Mathematics*, 11(2):431–441, 1963.
- [119] J.A. Snyman. *Practical mathematical optimization: an introduction to basic optimization theory and classical and new gradient-based algorithms*, volume 97 of *Applied Optimization*. Springer-Verlag New York, Inc., 2nd edition, 2005.