UNIVERSITÀ di **VERONA**

DOCTORAL THESIS

Measuring and understanding light in real life scenarios

Author:
Theodoros TSESMELIS

Coordinator:
Prof. Marco CRISTANI
Tutors:
Dr. Alessio DEL BUE
Dr. Fabio GALASSO

*A thesis submitted in fulfillment of the requirements
for the degree of Doctor of Philosophy*

May 8, 2019

UNIVERSITA' DEGLI STUDI DI VERONA

DEPARTMENT OF

Sciences of Mathematics and Informatics

GRADUATE SCHOOL OF

Natural Sciences and Engineering

DOCTORAL PROGRAM IN

Computer Science

WITH THE FINANCIAL CONTRIBUTION OF
NAME OF THE FUNDING INSTITUTION

Cycle / year (1° year of attendance):   XXXI / 2015

TITLE OF DOCTORAL THESIS

Measuring and understanding light in real life scenarios

S.S.D.: INF/01 - INFORMATICA
(Please complete this space with the S.S.D. of your thesis - mandatory information)*

Coordinator: Prof./ssa Marco Cristani

             Signature _____

Tutor: Dr./ssa Alessio Del Bue

             Signature _____

Tutor: Dr./ssa   Fabio Galasso

             Signature _____

Doctoral Student: Dott./ssa Theodoros Tsesmelis

             Signature _____

iv

Declaration of Authorship

I, Theodoros TSESMELIS, declare that this thesis titled, "Measuring and understanding light in real life scenarios" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.

- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.

- Where I have consulted the published work of others, this is always clearly attributed.

- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.

- I have acknowledged all main sources of help.

- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

_____

Date:

_____

UNIVERSITY OF VERONA

Abstract

Dept. of Computer Science
Doctoral School of Natural Sciences and Engineering

Doctor of Philosophy

Measuring and understanding light in real life scenarios

by Theodoros TSESMELIS

Lighting design and modelling (the efficient and aesthetic placement of luminaires in a virtual or real scene) or industrial applications like luminaire planning and commissioning (the luminaire's installation and evaluation process along to the scene's geometry and structure) rely heavily on high realism and physically correct simulations. The current typical approaches are based only on CAD modeling simulations and offline rendering, with long processing times and therefore inflexible workflows. In this thesis we examine whether different camera-aided light modeling and numerical optimization approaches could be used to accurately understand, model and measure the light distribution in real life scenarios within real world environments. We show that factorization techniques could play a semantic role for light decomposition and light source identification, while we contribute a novel benchmark dataset and metrics for it. Thereafter we adapt a well known global illumination model (*i.e.* radiosity) and we extend it so that to overcome some of its basic limitations related to the assumption of point based only light sources or the adaption of only isotropic light perception sensors. We show that this extended radiosity numerical model can challenge the state-of-the-art in obtaining accurate dense spatial light measurements over time and in different scenarios. Finally we combine the latter model with human-centric sensing information and present how this could be beneficial for smart lighting applications related to quality lighting and power efficiency. Thus, with this work we contribute by setting the baselines for using an RGBD camera input as the only requirement to light modeling methods for light estimation in real life scenarios, and open a new applicability where the illumination modeling can be turned into an interactive process, allowing for real-time modifications and immediate feedback on the spatial illumination of a scene over time towards quality lighting and energy efficient solutions.

Acknowledgements

I would like to thank and acknowledge my supervisors, Dr. Alessio Del Bue, Dr. Fabio Galasso and Prof. Marco Cristani, from the Departments of Pattern Analysis and Computer Vision (PAVIS), at Istituto Italiano di Tecnologia (IIT), the sensors lab and CV group at OS-RAM GmbH and the Department of Computer Science, at University of Verona (UNIVR) respectively, for the freedom and guidance they provided me during my PhD and the corrections they made to all the manuscripts. They gave me the opportunity to perform my thesis in a field specific to scene understanding and analysis, with still many challenges unsolved and empty space for novel ideas.

I would like to thank my colleague and fellow ESR in this project, Irtiza Hasan, for his support, encouragement and contribution in order to successfully accomplish the goals of the SceneUnderLight project within these 3 years.

I would like to acknowledge Herbert Kaestle for the support they gave me during my PhD. He helped me identify the challenges in terms of computer vision in relation to the physical components of lighting as well as for all the interesting discussions we had together that have clarified difficult points.

I am very grateful to Luc Masset, from the Department of Aerospace and Mechanical Engineering at University of Liège, for his insightful feedback and technical support, the sharing of some of his source code snippets, and the thorough explanations for the implementation of the radiosity model from scratch.

I would like to thank all the people I met during my stay at Istituto Italiano di Tecnologia, OSRAM and Verona, for their help and friendship. I would like to thank Vasiliki Katsageorgiou for her encouragement and helping me understanding some theoretical aspects as well for her help for creating some media material within the requirements of this work; Ling Wang for her explanations about factorization techniques; Michael Eschey for his help when I was looking for some source code explanations; Abdul Rafey Aftab and Ahmed Talbi for the discussions and their help on software development; To all working students at the Computer Vision lab in OSRAM for giving me permission to include them into my experiments and recording that I have applied during my PhD; And I am grateful to all the other people I met and provided me with support and I forgot to include in these acknowledgements.

I would like to express my gratitude to Bianca Miene for all her support and guidance in the bureaucratic issues that I have faced during my PhD.

Finally, I want to thank my family and friends for the support and encouragement they have given me throughout this work. Without their help during moments of discouragement, this work would not have succeed.

Contents

List of Figures

List of Tables

List of Abbreviations

2D	**2 D**imensional
3D	**3 D**imensional
aka.	**a**lso **k**nown **a**s
cd	**c**an**d**ela
c.f.	**c**onfer/**c**on**f**eratur, *(both meaning "compare")*
e.g.	**e**xempli **g**ratia or **e**xample **g**iven, *("for example")*
et al	**et al**ia, *(meaning "and others")*
etc.	**et c**etera
etc.	**et c**etera
ft	**f**o**o**t
k	**k**elvin
lm	**l**u**m**en(s)
lx	**l**u**x**
sr	**s**t**e**r**a**dian
w.r.t.	**w**ith **r**espect **t**o
AI	**A**rtificial **I**nteligence
AoI	**A**ctivation (of) **I**mage
AP	**A**verage **P**recision
Avg	**A**vera**g**e
BRDF	**B**idirectional **R**eflection **D**istribution **F**unction
BPP-NMF	**B**lock **P**rincipal **P**ivoting **N**on-negative **M**atrix **F**actorization
CAD	**C**omputer **A**ided **D**esign
CCD	**C**harge-**C**oupled **D**evice
CV	**C**omputer **V**ision
DNA-NMF	**D**iagonalized **N**ewton **A**lgorithm (for) **N**on-negative **M**atrix **F**actorization
FOV	**F**ield (of) **V**iew
GT	**G**round **T**ruth
IALD	**I**nternation **A**ssociation (of) **L**ighting **D**esigners
ICA	**I**ndependent **C**omponent **A**nalysis
IESNA	**I**lluminating **E**ngineering **S**ociety (of) **N**orth **A**merica
ILS	**I**nvisible **L**ight **S**witch
LDC	**L**ight **D**istribution **C**urve
LIT	**LI**gh**T**
LIT-EST	**LI**gh**T**-**EST**imation
LIT-ID	**LI**gh**T**-**ID**entification
LSC	**L**uxmeter **S**ensitivity **C**urve
NMF	**N**on-negative **M**atrix **F**actorization
PCA	**P**rincipal **C**omponent **A**nalysis
PhD	**Ph**ilosophiae **D**octor
PNMU	**P**riors **N**on-negative **M**atrix **U**nder-approximation
RE-LIT	**RE**-**LI**gh**T**ing

RGB	**R**ed **G**reen **B**lue
RGBD	**R**ed **G**reen **B**lue **D**epth
SI	**S**ystème **I**nternational (d'unités) or **I**nternational **S**ystem (of) **U**nits
SSIM	**S**tructural **SIM**ilarity Index
ToF	**T**ime (**of**) **F**light
TV	**T**ele**V**ision
VFOA	**V**iew **F**rustum (of) **A**ttention
VPL	**V**irtual **P**oint **L**ights

Physical Constants

pi	$\pi = 3.141\,592\,653\,589\,79$
albedo	$\rho = [0 - 1]$
conversion factor	$\kappa = \frac{\text{ß}}{\text{æ}}$

Nomenclature

$\| \ \|_F$	Frobenius norm	
$\mathbf{a}$ or $\mathbf{a}_{ID}$	Activation vector of $\tilde{S}_i$ for each $\mathtt{I}_{ID}$	
$\mathbf{a}_{RE}$	Activation vector of $\tilde{S}_i$ for each $\mathtt{I}_{ID}$ obtained from $\mathtt{P}$	
b	Radiosity matrix	
b_i or b_j	Radiosity value of specific patch or surface	
cd	candela	[W / sr]
$dA^{\perp}$	Perpedicular projected area to a direction	[m^2]
dw	Solid angle	sr
$\mathbf{e}$	Vector of the self-emmitance scalar values of all patches	
e_i	Scalar for the self-emmitance of patch	[W / m^2]/lx
$f_{\mathtt{I}}$	Function of the image formation process	
f_{ρ}	Bidirectional Reflection Distribution Function (BRDF)	
$f_{\rho,d}$	Difuse component of the Bidirectional Reflection Distribution Function	
$f(\theta, \phi)$	Function of the direction of a ray/beam	
$\mathbf{h}$	Individual non-negative coefficients associated to each basis	
$\hat{j}$	Index to maximum activation of image pixel map	
k	Kelvin	[1.380649×10^{-23} J/K]
l	Total amount of light sources	
lm	lumen	[cd sr]
lx	lux	[lm / m^2 sr]
m	Total amount of vectorized image pixels	
n	Total amount of vectorized images in matrix $\mathtt{V}$	
p	Distance	[m]
r	Decomposition scalar value	
$\mathbf{v}_i$	Individual vectorized training image	
$\mathbf{w}$	Individual extracted weights (bases)	
A	Area of a surface	[m^2]
AoI	Activation of image pixel map	
B	Radiosity	[W / m^2]
B_e	Self-emitted radiosity	[W / m^2]
E	Irradiance/Illuminance	[W / m^2]/lx
F	Form factors matrix	
F_{ij} or F_{ji}	Form factor value between two patches or surfaces	
G	Geometry term between two surfaces	
$\mathtt{H}$	All non-negative coefficients associated to the bases	
I	Radiant intensity	[W / sr]
$\mathtt{I}$	Image	
$\mathtt{I}_{EST}$	Image in training set	
$\mathtt{I}_{ID}$	Image in testing set	
$\mathtt{I}_{RE}$	Ground truth synthesized image	
$\mathtt{I}_{\tilde{RE}}$	Estimated synthesized image	

K	Combination of the geometry, G, and visibility, Y, terms	$[G \times Y]$
L	Radiance	[W / m^2 sr]
$\mathcal{L}$	Set of l images of the individual light sources	
$\tilde{\mathcal{L}}$	Set of the estimated l images of the individual light sources	
L_d	Diffuse component of radiance	[W / m^2 sr]
L_e	Emitted component of radiance	[W / m^2 sr]
L_ρ	Reflected component of radiance	[W / m^2 sr]
N	Normal of surface	
P	Target image providing the desired light patern	
S_i	Light sources	
$\tilde{\mathrm{S}}_i$	Estimated light sources	
$SSIM$	Structural Similarity Index	
$\mathcal{T}$	Training image sequence	
V	Matrix of vectorized training images I	
$\hat{\mathrm{V}}$	Estimated matrix of vectorized training images I_{EST}	
W	All extracted weights (bases)	
α	Aperture angle	[*degrees*]
δt	Time	[seconds]
$\varepsilon_{\mathrm{EST}}$	Error metric for evaluating LIT-EST	
$\varepsilon_{\mathrm{ID}}$	Error metric for evaluating LIT-ID	
$\varepsilon_{\mathrm{RE}}$	Error metric for evaluating RE-LIT	
μ_j	Rays arriving patch or surface j	
μ_i	Rays starting from patch or surface i	
ρ	Reflectance (albedo)	
ρ_d	Diffuse reflectance (albedo)	
ρ_i	Reflectance (albedo) of specific patch or surface	
Γ	Set of surfaces	
$\Delta\varepsilon$	Energy emmited from source	[W (joules/sec)]
Λ	Vectorized set of the estimated light source images $\tilde{\mathrm{S}}$	
Υ	Visibility predicate between two surfaces	[0 or 1]
Φ	Flux/Radiant power	[W]
Ω	Hemisphere	

Chapter 1

Scene Composition Analysis

Contents

1.1 Introduction

1.1.1 Light quality of the visual environment

Light affects both our perceptions of the world and our emotional and physiological responses [PL00; PRDR03; SCK15; TW13; REA00], and thus it is essential in gathering information about the physical world. Good-quality lighting can support visual performance and interpersonal communication and improve our feelings of well-being. Thus, while good lighting in the workplace with well-lit task areas is essential for optimizing visual performance, visual comfort and ambience, especially with an aging workforce [KPH16; BOYC04; GGO06], poor-quality lighting can be uncomfortable and confusing and can inhibit visual performance.

The overall purpose of lighting is to serve the needs of people. The role of the lighting designer is to match and rank the needs of the people using the space with the economic and environmental considerations and the architectural objectives, and then to translate the results into a workable design and functional installation. The human needs served by lighting are identified in Figure 1.1a (see upper circle, *i.e.* "Individual Well-Being").

Thus, lighting can be used for much more than to illuminate. It can enhance productivity, creating flexible spaces that adapt to the task at hand; energy-efficient lighting solutions for industry can reduce environmental impact and save on costs, while at the same as increase life quality and productivity.

The International Association of Lighting Designers [IALD18], see Figure 1.1a, and based on Veitch's *et al.* research [VEIT00] states that good lighting amounts at *achieving an optimal balance among human needs, architectural considerations, and energy efficiency*.

Figure 1.1a, shows the effort from the lighting researchers to understand the variables that influence the goals of a lighting installation. As it is shown one can divide the broad

(a) Light quality framework [VEIT00].

(b) Energy saving strategies [TAK09].
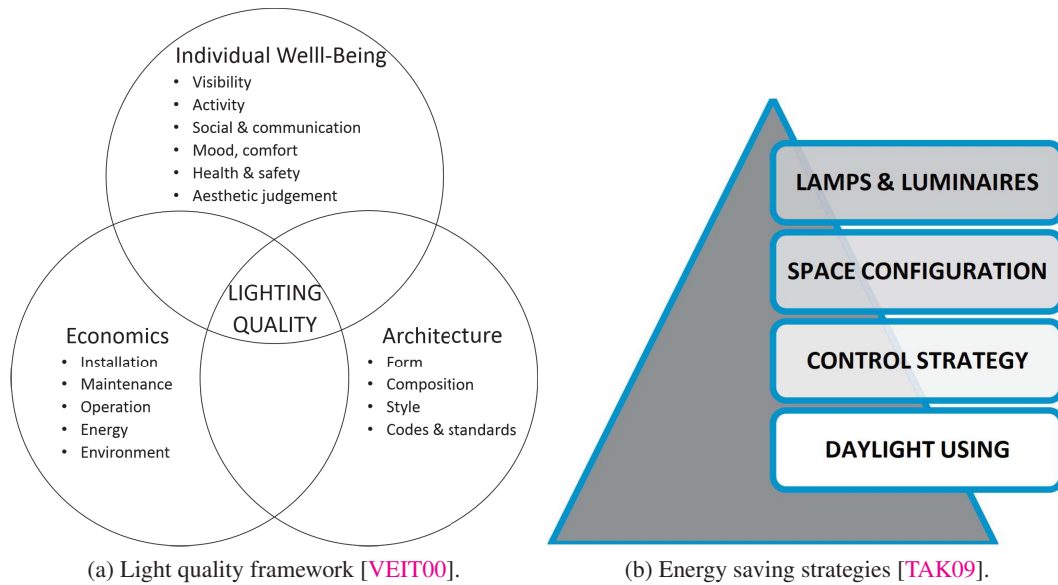
FIGURE 1.1: **(a)** The framework of lighting quality according to the International Association of Lighting Designers (IALD) [IALD18]. **(b)** The pyramidal scheme shows the energy saving strategies. From the bottom to the top, this implies the exploitation of natural light, the distribution and control of lighting locally, efficient space configuration with bright colors and open spaces for better light propagation and utilization of the light sources driven by energy saving customization. The order of the strategies in the pyramid shows the importance that each action should be applied in a green-oriented building maintenance.

domain of lighting research into three overlapping subfields, all of which have roles to play in the establishment of technologies, design processes, and knowledge pertaining how to improve lighting quality. These are: research aimed at developing lighting (and daylighting) technologies; architectural and design research; and, human factors research (encompassing biological and psychological processes and effects) [VEIT00].

For instance, the lighting in a surgical room will evaluated as good quality only if visibility of two- and three-dimensional objects is excellent in terms of both luminance and colour contrast. The cost of the lighting to achieve that will be relatively unimportant in regards to the importance of the task. On the other hand, the lighting in a living room will require less attention to the visibility of very fine details and more focus to the appropriate lighting for social interactions. In this case the cost of the lighting will be more important to some homeowners than to others.

1.1.2   Light energy management

Generally in order good-quality lighting to be ensured, we are pretty sensible on the illumination around us, especially during office and home activities, *e.g.* reading, writing, and studying or doing precision works in general. This brings to the tendency of keeping the environment around us all lit, increasing the energy spent. However, most of us do not consider the impact that indoor illumination has on the power consumption and/or the monthly costs of large environments like offices or warehouses in big buildings.

Moreover, if we look on the characteristics of most indoor office environments, these are usually large, with high spaces and plane offices, objects are usually with fixed positions (*e.g.* desks, couches, wardrobes, *etc.*), fixed light source installations with restricted relocation

opportunities. The above characteristics can lead clearly to significant energy consumption conditions for lighting [KAW15].

If we look in the literature, it is shown [KAW15; ZYHR15; REA00] that the lighting consumption of a building can take more than 15% of overall electricity consumption, while at peak periods; this can reach up to approximately a fourth or even more. It is obvious at energy audits that savings in the lighting are usually most evident and most easily feasible in environments where the human occupancy is limited. However, in dynamic environments where the human presence is more evident the power saving strategies are becoming more complex and harder to be addressed. The basic energy saving techniques and strategies, see Figure 1.1b, usually focus on the following principles: **a)** maximize the use of daylight, **b)** make lighting control as local as possible and get staff involved in energy saving planning, **c)** use bright coloured walls and ceilings **d)** utilize and adjust the light sources to the most energy efficient lamp/luminaire combinations.

1.2 Aim of study

In this work we try to encompass both fundamental research in computer vision and in smart lighting, in order to push forward for a new generation of smart lighting systems for a better quality living and energy saving solutions. SMART Lighting is a novel field of research, whereby we try and bring together the new capabilities of lighting and the novel achievements of computer vision.

This study aims to create a breakthrough in light management systems by enabling the understanding of the environment's lighting by using a single camera for smart illumination and energy saving applications. Thus, we applied research and developed novel autonomous tools using computer vision and machine learning approaches that seamlessly can integrate into a smart lighting system for indoor environments. This can be applied with the use of a sensing device and implemented as a part of an algorithmic analysis in a process unit (*i.e.* the processor). The former could be a simple RGB camera (or RGBD if including a depth sensor, *e.g.* ToF), co-located along with a luminaire installation to provide top-views of the scene. The latter involves the algorithms to understand the scene and to make decisions on lighting, which are then communicated to the lighting devices through a lighting control communication unit (*i.e.* the controller), see Fig. 1.2.

Towards the understanding of the scene, we distinguish the scene structure material properties and the human-centric scene. The first regards the scene composition: meaning its 3D structure, the objects materials, the light position and characterization (natural versus artificial) and their lighting patterns. The second regards the human activities and interactions, particularly the human-scene (walking, working at desk or reading, presenting at a board) and human-human interaction (where people meet, discuss, relax, *etc.*).

These two aspects are tightly intertwined, since the structure of the scene allows and constrains human activities, but at the same time the human activities influence the scene structure. Consider for example a warehouse as the static scene: its structure continuously changes due to the different arrangement of the goods, the latter being a direct consequence of the human activities carried out in the environment. In other words, the structure of the scene and the human have to be considered as parts of a whole, accounting in addition for their continued temporal evolution. For this reason, it appears convenient to deal with the two topics within the same research framework, for the first time in the literature.

Thereafter, the scene-centric composition analysis and the human-centric scene understanding are combined towards a smart lighting management system with the key idea of *"Invisible Light Switch"* (ILS), where the user have the sensation of *"all lit"* while actually
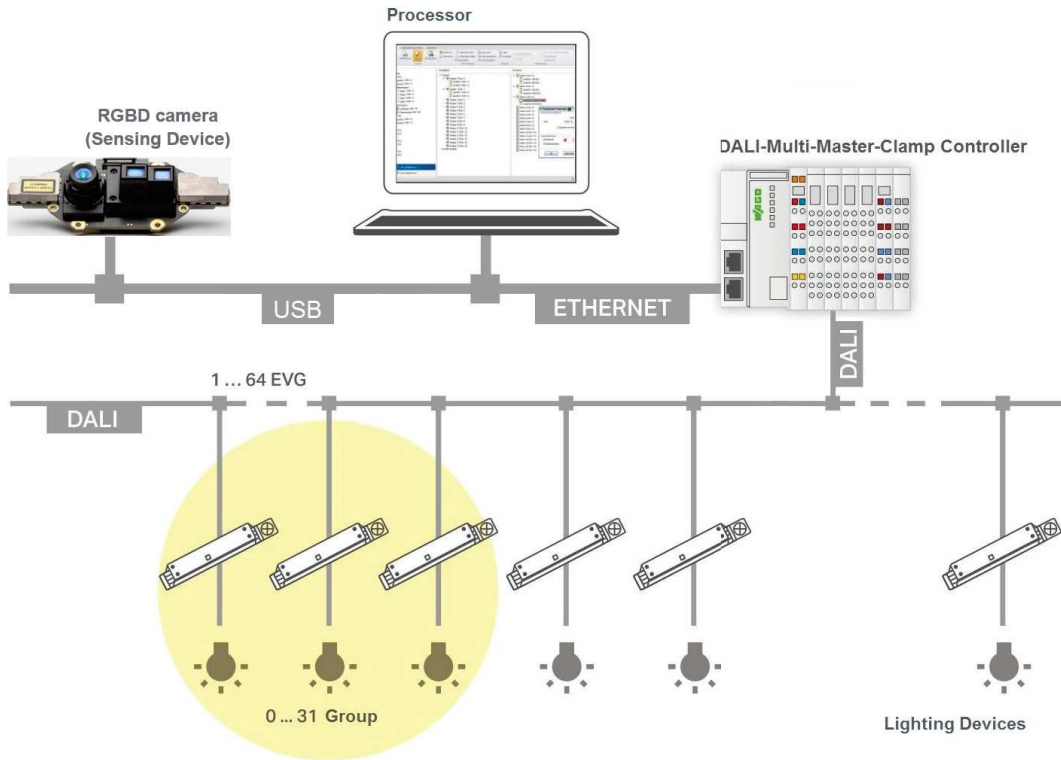
FIGURE 1.2: The end-to-end *ILS* architecture, starting from the sensing device (*top left*), the processor (*middle top*) where all the algorithmic analysis and computations are applied, to the lighting control communication unit (*top right*) and the actual lighting devices (*bottom*). Note the RGBD sensor connected to the processor via USB, we then use the WAGO DALI Configurator software to encode luminaire commands to a DALI (Digital Addressable Lighting Interface) Master Light Controller via TCP/IP. Thereafter the controller relays the commands to the luminaire and reads out their statuses via the DALI bus communication protocol.

the scene is optimally lit, therefore providing both notable energy saving and a productive environment.

The idea behind the *Invisible Light Switch* is straightforward: the proposed system controls and sets the illumination of the environment by taking into account the information regarding the part of the scene that the user can see or cannot see, by switching off or dimming down the lights outside the user's View Frustum of Attention (VFOA), and thus ensuring a consistent energy saving and productivity. The study of the scene as discussed above serves this goal: knowing the 3D geometry of the scene and the map of interreflectances will allow to understand how the different light sources impact each point of the space; knowing where a user is located and what is his posture serves to infer what he can see and what he cannot, individuating potential areas where the light can be switched off. Moreover, being able to forecast his future activities will help understand (in advance) which lights should be switched on, avoiding the user to continuously act on the illumination system, and showing the user the illumination scenario that he wants to have.

1.2.1 Scene-centric lighting composition analysis

The structure of a scene consists of a number of object material properties and their arrangement in the 3D space. This aspect is fundamental in order to understand the lighting propagation effects and the localization of the natural artificial sources. Within the 3D scene structure, light propagation can be defined as an inverse problem called inverse lighting [PP03;

CBNR11].

In this study, for the first time, inverse lighting is tackled in a real environment, typically indoor, that presents complex geometries and several types of lights (artificial and natural). Given the reconstructed 3D model of the scene from a top-view RGBD sensor, the problem will result in the estimation of the photometric properties of the scene and objects material together with a coarse localization of the lights and the spatial illumination of the scene over time. In particular, we rely on the fact that, given a large collection of images and based on the reconstructed geometry, inverse lighting becomes tractable and can be modeled through a visual computing pipeline (*e.g.* rendering).

By leveraging a larger number of images depicting the same scene over time (*i.e.* time-lapse) or considering the geometry of the scene together with estimated photometric properties as well as the light source characteristics, it is possible to reduce the ambiguity of the illumination modeling problem by studying the evolving lighting conditions.

To this end, the major goal is to provide a new theoretical framework for estimating the spatial illumination and scene properties over time from a single view.

In particular, the research was divided in three major stages:

1. Image spatio-temporal clustering for light source separation and identification with the aim to associate at each image pixel one or more light sources (artificial or natural), see Chapter 3

2. 3D light source localization, material photometric properties and inter-reflections / illumination map estimation given the RGBD images and the previous clustering results, see Chapter 4

3. Real-time inference of the current lighting conditions given a single image in order to make feasible ILS application from the human-centric light sensing perspective, see Chapter 5

1.3 Literature

In this section we provide an overview of the existing state-of-the-art literature related to the different aspects of light modeling and understanding that we are trying to address in this work, meaning light source calibration and light measurement methods based on the spatial and human-centric sensing.

1.3.1 Review on light sources calibration

Light source calibration in indoor environments is a key component of light modeling and it is a complicated procedure especially when it involves multiple light sources of different types (*e.g.* artificial, natural), the presence of objects of diverse sizes, shapes and spectral characteristics. Thus, most works assume known geometry or employ external sensors or tools, *e.g.* light probes, colorchecks. The existing methods can be roughly divided into three categories with the corresponding assumptions: 1. estimating light source directions for scenes with a distant point light source, 2. estimating illumination distributions in natural light source environments, and 3. estimating light source positions in scenes with a near point light source.

Reference spheres and a near light assumption are used in [TMNM09], achieving realistic results. However, without external sensor information, light source estimation is a complex and ill-posed problem. Lopez-Moreno *et al.* [LHRG10] and Lombardi *et al.* [LN16] make use of isolated reference objects (user input) to estimate multiple light sources from a single image. The approaches described in [PF92; WS03] estimate the illumination by taking

advantage of the cast shadows and specular reflections of known geometries in the scene. Extensions to the latter are introduced in [HNL05], which assume different reflectance models. Powell *et al.* [PSG01] extracts the positions of multiple light sources through a triangulation based on highlights of three specular spheres at known positions. On the other hand Ackermann *et al.* [AFG13] in contrast to a simple triangulation scheme models the light estimation as a nonlinear least squares minimization of the highlights reprojection error, yielding to an improved accuracy.

Karsch *et al.* [KHFH11] introduced the idea of light estimation and correction through a rendering-based optimization procedure. They optimize for the light position and intensity by a cost function, encoding the pixel difference between the real image and the simulated one. The rendering-based optimization procedure is also followed in [NMT12; BONM+15], coupled by depth information, with a very time-consuming solution and/or limited to a small number of light sources. Depth information and inverse rendering are also used in [OIDS15] for achieving illumination estimation and relighting. As for [LHRG10; KHFH11], these works also requires shape, position and color of light sources.

Regarding the estimation of light in outdoor environment, the work of Lalonde *et al.* [LEN12] estimates natural illumination from a single image using a sky model. This can allow to insert digital objects or to create animations with more realistic effects. Remarkably, the work of Tian *et al.* [TDRH+16] estimates the spectral power distribution of natural light in different weather conditions.

In recent works, Kasper *et al.* [KKSH17] tries to improve the computationally expensive technique of path-tracing for light estimation with the use of finite nonlinear parametrization while in [GSYS+17] they introduce a learning based approach. Santo *et al.* [SWSS+18] presents a practical method for geometric point light source calibration based on the simultaneous recovery of the light position and small shadow casters created from pins on a Lambertian plane at unknown positions in a structure from motion framework. Finally in [SSK17] similarly to our work (Chapter 3), light separation through a factorization technique is part of the computational imaging study on the electrical grid for bulb type and light phase estimation.

In the contrary of the above solutions in the proposed solution (Chapter 3) we do not require any scene nor light calibration. Additionally, to the best of our knowledge, we are the first to address non-synthetic scenes of entire offices.

1.3.2   Review on light measurement methods

Measuring light is a problem that can be addressed by different fields and it falls within the studies of light modelling and understanding. In the following we provide some notions of three major topics in the literature related to light measurements.

Image processing

According to Cuttle *et al.* [CUTT10] the lighting profession and lighting evaluation procedures are moving from the conventional illuminance-based towards the luminance-based. That means to move from assessing light incident on planes (*e.g.* illuminance) towards the assessing light arriving at the eye (*i.e.* luminance). Given this change, the works in [CAI16; CS15; HE14] take advantage of the emergence of camera-aided light measurement solutions. The white paper of Hiscocks *et al.* [HE14] provide an overall understanding of the luminance measuring procedure with a digital camera. Choo *et al.* [CS15] makes use of such a procedure for obtaining the luminance in a small simulated environment, structured from a carton box, a cheap web camera and a processing unit. On the other hand, Cai *et al.* [CAI16] instead focused on to a more advanced solution by taking advantage of the high dynamic range

(HDR) computational photography and the needed high quality equipment. However, all the previous mentioned solutions require a pre-calibration step of the camera sensor where pixel values are mapped according to a known luminance source. However, the light intensity illuminating a surface is not reliably recoverable from a pixel-like array of radiance values, because they are the product of the irradiance, the surface reflectance, and complex interreflections between all surfaces in the 3D scene. Untangling them is very challenging in the general case – as well explained in the pioneering works on *"inverse rendering"* by Ravi Ramamoorthi and Hanrahan [RH01], Steve Marschner [MG97].

Computer graphics

Light modeling and understanding have been also studied in computer graphics for the creation of photo-realistic renderings [HF14; MS15]. The forward and inverse light transport theory physically simulates the path of transmitted light in a 3D environment and models the image as an integration process. This formalization requires the 3D mesh, the material surfaces and many other physical properties [CBNR11; KKGK+14; VG95]. To this end, many light models have been proposed for retrieving and rendering scenes with as much as possible lifelike illumination. The most well known model is radiosity [CG85; GTGB84], popular for its simplicity and efficiency. Other more advanced and recent approaches are the instant radiosity [KELL97] with its bouncing energy and the Virtual Point Lights (VPL), photon mapping [JENS96] and progressive photon-tracing [HOJ08] with the idea of tracing photons from a light source through the scene and store their hits on diffuse surfaces in a so-called photon map. Deeb *et al.* [DMHT18] finally in his survey explored the importance of the interreflections, the consideration of only two bounces of light between surface elements and how could this improved based on radiometric definitions.

Light design software for light modeling

Relux [RELU10], DIALux [DIAL94] and AGi32 [LIGH90] are commercial CAD-design modelling software products that are commonly used in the lighting design field, for measurement and evaluation of lighting solutions. These software require the information of the luminaire specifications by the manufacturers, the material and photometric properties of objects in the scene that normally are retrieved from online libraries and finally the CAD model of the indoor environment. The light simulation process in all these software is based on the radiosity method, and different variants of it [REND06]. Recently two software prototypes, HILITE and LiteMaker [HILI16; KLSW17], are currently being developed in an academic environment and they try to combine a physically based real-time rendering with an interactive lighting simulation for complex architectural environments. HILITE uses a many-light global-illumination solution provided as light maps, including glossy material [LTHS+13; LTMS+14] where virtual point lights are clustered into a set of virtual polygon lights, which represent a compact description of the illumination in the scene, while LiteMaker combines a multi-resolution image filtering technique with an interactive, progressive photon-tracing algorithm [HOJ08].

All the above presented solutions feature limitations, *e.g.* user input, in advance information of the scene structure, geometry and photometric properties, time consuming solutions due to long rendering and simulation sessions, *etc*. To overcome some of these bottlenecks in this work we tried to bring together the best of visual computing and lighting design software; by using only RGBD input and the initial light source properties without any other user inputs; for obtaining a dense light intensity estimation of an indoor environment in realistic scenarios (Chapter 4).

1.3.3 Review on human-centric light sensing

Relationship between human activities and lights is a widely studied topic in perceptual sciences [AZ91; FHSM79; GIFF88]. Recently, it was shown by [XL13] that light intensifies people's perception by triggering the emotional system leading to intensified effective reactions. Light also changes our perception of space [GV06] and we tend to associate different illumination patterns to different social gatherings (musical concert vs candle light dinner). People seem to share more details in bright light than darkness [CD74], we as humans also rely on facial expressions which are only visible in light. Moreover, light provides sense of security [GIFF88], people adopted specific roads and streets in night due to the illumination [TS74]. Recently, studies targeting the office environments revealed a strong connection between people's productivity and the lights [KAW15; KPH16; SKTK12]. Eyeing the importance of lighting on humans, communities such as Human Computer Interaction (HCI) [PMAJ13] provided a sense of "belongingness" to the residents, by deploying interactive lighting installation in a city square. Furthermore, ubiquitous computing [GCGL$^+$11; IN06] and architectural design [MHF13] have also investigated this topic to an extent. However, there are also studies that question the relationship between the light perception and the actual measured spatial illumination [BERN17; REA00].

Despite receiving a wide scale attention, the literature in computer vision seems to have ignored the modelling of light and behaviour. Only recently we [HTDG$^+$17] presented the idea of jointly modeling the relationship of light and human behavior via long term time-lapse observation of the scene by recognizing and forecasting activities using the head pose estimation as a proxy for the gaze. To this end, the major goal in this work was to provide a new computer vision system for estimating the illumination map along with the human occupancy and attention from a single view. We did this by adapting and extending a well known light illumination algorithm (*aka.* radiosity) and combining it with the knowledge of human occupancy and posture [HTGD$^+$17] into a unique pipeline as we show in Chapter 5.

1.4 Thesis contributions

This research could be considered as unique, since for the first time we evaluate and test in practice how sufficiently reliable the actual use of computer vision and 3D modeling techniques could be for real scene light modeling. Our contributions are multi-fold and are summarized as follows:

1. We show how factorization techniques could play a semantic role for light decomposition and light source identification, introducing also a novel benchmark dataset and metrics.

2. We present and extend a well known global illumination model (*i.e.* radiosity) with real light source properties and by eliminating well known limitations (*e.g.* timely operational and simulation sessions, point based light sources, isotropic light perception, known photometric properties and geometry, *etc.*).

3. We show how the obtained spatial light estimation output could be combined with the aspect of the human occupancy and head pose estimation in order to retrieve information regarding the gaze-gathering light perception and how this could be used for ego-centric lighting and power saving solutions.

4. Finally, we combine and present all these together into a complete end-to-end system and set the baselines for adapting computer vision solutions for real life smart lighting applications.

The above contributions were further supported from the associated publications and patent disclosures listed below:

Publications:

[1] I. Hasan, **T. Tsesmelis**, A. Del Bue, F. Galasso and M. Cristani. "Don't Turn Off the Lights": Modelling of Human Light Interaction in Indoor Environments, In *Proc. IEEE International Conference on Image Analysis and Processing (ICIAP)*. Catania, Italy, September 2017.

[2] I. Hasan, **T. Tsesmelis**, A. Del Bue, F. Galasso and M. Cristani. Tiny head pose classification by bodily cues, In *Proc. IEEE International Conference on Image Processing (ICIP)*. Beijing, China, September 2017.

[3] red**T. Tsesmelis**, I. Hasan, M. Cristani, A. Del Bue and F. Galasso. LIT: a system and benchmark for light understanding, In *Proc. IEEE International Conference on Computer Vision, Workshop on Color and Photometry in Computer Vision (ICCVW)*. Venice, Italy, October 2017.

[4] I. Hasan, F. Setti, **T. Tsesmelis**, A. Del Bue, F. Galasso and M. Cristani. "Seeing is Believing": Pedestrian Trajectory Forecasting Using Visual Frustum of Attention, In *Proc. IEEE Winter Conference on Applications of Computer Vision (WACV)*. Lake Tahoe, USA, March 2018.

[5] I. Hasan, F. Setti, **T. Tsesmelis**, A. Del Bue, F. Galasso and M. Cristani. MX-LSTM: mixing tracklets and vislets to jointly forecast trajectories and head poses. In *Proc. IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*. Salt Lake City, USA, June 2018.

[6] **T. Tsesmelis**, I. Hasan, M. Cristani, A. Del Bue and F. Galasso. RGBD2Lux: Dense light intensity estimation with an RGBD sensor. In *Proc. IEEE Winter Conference on Applications of Computer Vision (WACV)*. Hawaii, USA, January 2019.

[7] **T. Tsesmelis**, I. Hasan, M. Cristani, A. Del Bue and F. Galasso. Human-centric light sensing and estimation from RGBD images: The Invisible Light Switch. In *Proc. IEEE Winter Conference on Applications of Computer Vision (WACV)*. Hawaii, USA, January 2019.

[8] I. Hasan, F. Setti, *T. Tsesmelis*, V. Belagiannis, S. Amin, A. Del Bue, M. Cristani, F. Galasso. Forecasting People Trajectories and Head Poses by Jointly Reasoning on Tracklets and Vislets. [**Under review**]

[9] **T. Tsesmelis**, I. Hasan, M. Cristani, A. Del Bue and F. Galasso. Machine vision for light measurement and human-centric light management systems. [**Under review**]

Patents:

[1] I. Hasan, F. Setti, **T. Tsesmelis**, H. Kaestle, M. Eschey, M. Cristani, A. Del Bue and F. Galasso. Patent disclosure with the title "Framework and Method for Tracking people with Frustum Detection". Filed to the Italian and International Patent Offices (IPO/PCT), Aug. 2016.

[2] **T. Tsesmelis**, I. Hasan, H. Kaestle, M. Eschey, M. Cristani, A. Del Bue and F. Galasso. Patent disclosure with the title "Framework for Light Source Identification for Indoor Application". Filed to the Italian and International Patent Offices (IPO/PCT), Aug. 2016.

[3] I. Hasan, F. Setti, **T. Tsesmelis**, H. Kaestle, M. Cristani, A. Del Bue and F. Galasso. Patent disclosure with the title "Forecasting trajectories and view frustum by LSTM". Filed to the Italian Patent Office (IPO), Feb. 2018.

[4] **T. Tsesmelis**, I. Hasan, H. Kaestle, M. Cristani, A. Del Bue and F. Galasso. Patent disclosure with the title "A camera-aided radiosity global illumination model". Filed to the Italian Patent Office (IPO), April 2018.

1.5   Structure of the thesis

The thesis comprises six chapters. The second chapter provides the necessary background related to the topic addressed in the thesis so that the reader can pursue the reading with all basic notions in mind.

Chapter 3 describes the **_LIT_** system. A computational system and benchmark for light understanding (light decomposition and identification). Using computational imaging and factorization techniques we model and benchmark the light variations of indoor scenes with different illuminations, including natural light and multiple luminaire setups.

Chapter 4 is dedicated to light modeling developments used to estimate the spatial illumination over time for indoor environments, as presented by the **_RGBD2Lux_** pipeline.

Chapter 5 presents the deployment of the estimation of the spatial illumination over time in accordance with gaze-gathered light estimation and a human-centric perspective.

Finally, Chapter 6 presents the conclusions and perspectives, as well as recommendations for further research.

Chapter 2

Basics on Light Transport

Contents

2.1 Light

"The quest to understand the nature of light has led curious human beings down into the innermost secrets of the atom and out to the farthest reaches of the starry universe."
–Ben Bova

According to the Illuminating Engineering Society of North America (IESNA) handbook [REA00] light is defined as the radiant energy that exits the human eye and creates a visual sensation. On the other hand, considering it as a physical quantity, light is defined in terms of its relative efficiency throughout the electromagnetic spectrum or radiant energy which lays between approximately 380nm and 780nm, see fig. 2.1.
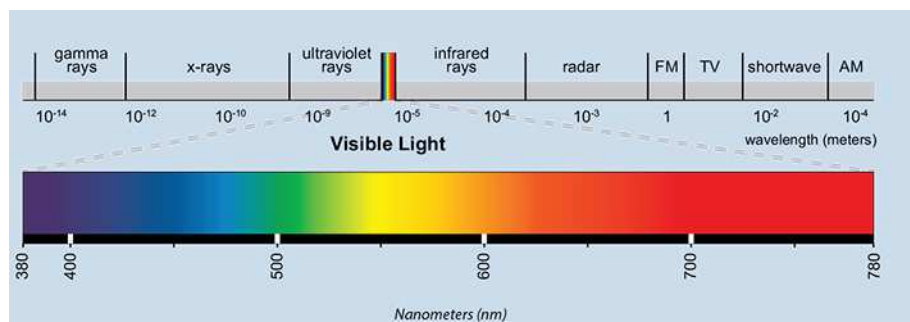
FIGURE 2.1: Visible light is the part of the spectrum from 380nm to 780nm.

One of the first theories regarding the entity of light was involving the notion that light was emitting from the human eyes, and that the environment was rendered visible when it

was struck by the emission. However, Aristotle rejected this theory when questioning the lack of vision in the dark. Since then, through the years, many alternative theories have been proposed and discussed. From a physical perspective, most of these theories generally regarded light as an energy transfer from one surface to another. We present some of these theories briefly below.

Quantum optics

Quantum optics is the fundamental model of light which considers and tries to explain its dual wave-particle nature. However, this model is generally considered too detailed for the purposes of image generation for typical computer graphics scenes and not only. Therefore it is not commonly used.

Wave model

The wave model is a simplification of the quantum model. Generally it is described by Maxwell's equations. With this model it is possible to capture effects, such as diffraction, interference, and polarization, that arise when light interacts with objects of size comparable to the wavelength of light. However, again for the purposes of image generation in computer graphics, the wave nature of light is also a model that in most cases is typically ignored.

Geometric optics

The geometric optics model is considered the simplest and most often used to model the light characteristics in imaging and visual computing. In the geometric optics model the light assumed to be emitted, reflected, and transmitted which leads to the following several assumptions regarding its behaviour:

- Light is emitted in straight lines, therefore effects such as diffraction where light "bends around" objects are out of the scope of this model.

- Light is emitted instantaneously through a medium; this is a practical assumption since it requires the global illumination algorithms to compute the steady-state distribution of light energy in scenes.

- Light is not affected by external physical quantities, such as gravity or magnetic fields.

2.2   Measurement of light

"When you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot express it in numbers, your knowledge is of a meager and unsatisfactory kind; it may be the beginning of knowledge, but you have scarcely, in your thoughts, advanced to the stage of science, whatever the matter may be."
–Lord Kelvin

Based on the literature [MAIT15; CWH93; DBBS06], the light energy distribution of a scene is measured from physical quantities that can represent light energy. Radiometry and photometry are the two related fields which have as a common subject to study the energy effects of the light radiation. Therefore, they describe two complementary aspects:

 - radiometry is the science that studies measurable and physical aspects of radiation;

- photometry studies and analyzes the subjective, perceptual aspect of the perceived light radiation, *e.g.* the light perceived from a reference observerThus, it deals with the quantification of the perception of light energy.

All forms of energy in principle can be inter-converted, a radiometric measurement is expressed in the SI units for energy or power as *joules* and *watts*, respectively. On the other hand the photometric unit for luminous power is the *lumen*, and the photometric unit for luminous energy is the *talbot*. The photometric quantities can be derived from the corresponding radiometric terms, thus most of the imaging and visual computing algorithms apply their computations on radiometric terms which we present below.

2.2.1 Radiometric quantities

Radiant Power/Energy or Flux

The main radiometric quantity is the radiant power, also called flux. The radiant power, is denoted as $\Phi = \frac{\Delta \varepsilon}{\delta t}$, where $\Delta \varepsilon$ consists the emitted energy by the source and expressed in watts, $W(joules/sec)$. In practise, this quantity expresses how much total energy flows from/to/through a surface per unit time. Note however that the flux does not consider any information regarding the sizes of the light source or the receiver (object surface), nor it includes any specification of their distance.

Radiant intensity

As radiant intensity of the light source is defined the flux, $d\Phi$, per solid angle towards a given direction of a small beam, dw:

$$I = \frac{d\Phi}{d\omega}. \tag{2.1}$$

Equally in photometry the corresponding quantity is the luminous intensity. The radiant intensity in a given direction is equal to the irradiance at a point on the unit sphere centered at the source. It is really common in the literature though the intensity to be defined as the power per unit area, rather than per unit solid angle.

Radiance

Radiance is one of the most important quantities regarding the visual computing algorithms. It is defined as the flux per unit projected area per unit solid angle $d\omega$ ($watts/(steradian \times m^2)$). Intuitively, radiance expresses how much power arrives at (or leaves from) a certain point on a surface, per unit solid angle, and per unit projected area and supposedly captures the "appearance" (*i.e.* the emitted or reflected energy) of the objects within the scene. Radiance is a quantity that varies with position x and direction vector Θ, and can be expressed as $L(x, \Theta)$:

$$L(x, \Theta) = \frac{d^2\Phi(x, \Theta)}{d\omega dA^{\perp}} = \frac{d^2\Phi(x, \Theta)}{d\omega dA \, cos\theta}. \tag{2.2}$$

The projected area $dA^{\perp}$ is the area of the surface projected perpendicular to the direction we are interested in. This derives from the fact that the power arriving at a glancing angle is scattered over a larger surface. Therefore, since we explicitly want to express the power per (unit) projected area and per (unit) direction, we need to consider the larger area, and thus this explains where the cosine term comes from withing the formulation.

In image generation and rendering, radiance is considered as a fundamental radiometric quantity based on the following property:

- Radiance is considered invariant along straight paths. Meaning that that the radiance leaving point x directed towards point y is equal to the radiance arriving at point y from the point x, this mathematically is expressed as $L(x \rightarrow y) = L(y \leftarrow x)$ with the assumption that there is not any medium in between (traveling through a vacuum).

Irradiance

The irradiance, E is defined as the incident radiant power on a surface, per unit surface area and it is measured in $watts/m^2$ and formulated as:

$$E = \frac{d\Phi}{dA}, \tag{2.3}$$

or

$$E = \int_\Omega L_i \cos\theta \, d\omega, \tag{2.4}$$

where L_i, is the integrated incident or incoming radiance over a hemisphere, Ω. Its corresponding photometric counterpart quantity is the illuminance.

Radiosity

Radiosity, denoted as B, is a very similar radiometric quantity to irradiance. While irradiance is the energy per unit area incident (arriving) onto a surface, radiosity is the energy per unit area that leaves a surface and it is given from the following formulation:

$$B = \int_\Omega L_o \cos\theta \, d\omega \tag{2.5}$$

where L_o is the outgoing radiance. Radiosity is also known as radiant exitance and its photometric equivalent is the luminosity.

To summarize, as we have seen there are mainly six radiometric (and their photometric counterparts) quantities that can be used in order to characterize the light distribution within an environment. These are the radiant energy (luminous energy), radiant power (luminous power), radiance (luminance), irradiance (illuminance), radiosity (luminosity), and radiant intensity (luminous intensity), see Table 2.1.

Radiometry	Radiometric Units	Photometry	Photometric Units
Radiant energy	joules [J = kg m^2 / s^2]	Luminous energy	talbot
Radiant power	watts [W = joules / s]	Luminous power	lumens [talbots / second]
Radiance	[W / m^2 sr]	Luminance	Nit [lumens / m^2 sr]
Irradiance	[W / m^2]	Illuminance	Lux [lumens / m^2 sr]
Radiosity	[W / m^2]	Luminosity	Lux [lumens / m^2 sr]
Radiant intensity	[W / sr]	Luminous intensity	Candela (cd) [lumens / sr]

TABLE 2.1: Radiometric and photometric light quantities [DBBS06; CWH93; MAIT15].

2.2.2   Light measuring equipment

Referring on the Illuminating Engineering Society of North America (IESNA) handbook [REA00] we see that the radiometric measurement instrumentation consists of several parts which are: a detector, a mean of conditioning or amplifying the output of the detector, a

FIGURE 2.2: Various commercial illuminance meters (luxmeters), that we also used in this thesis in order to obtain ground truth light measurements.

method of displaying or storing the measurement, and possibly an optical element or system of elements to collect the radiant quantity to be measured. Moreover, depending on the geometric relationship between the source and detector, the quantity measured is radiance, irradiance, or radiant intensity (or the corresponding photometric quantities: luminance, illuminance, and luminous intensity).
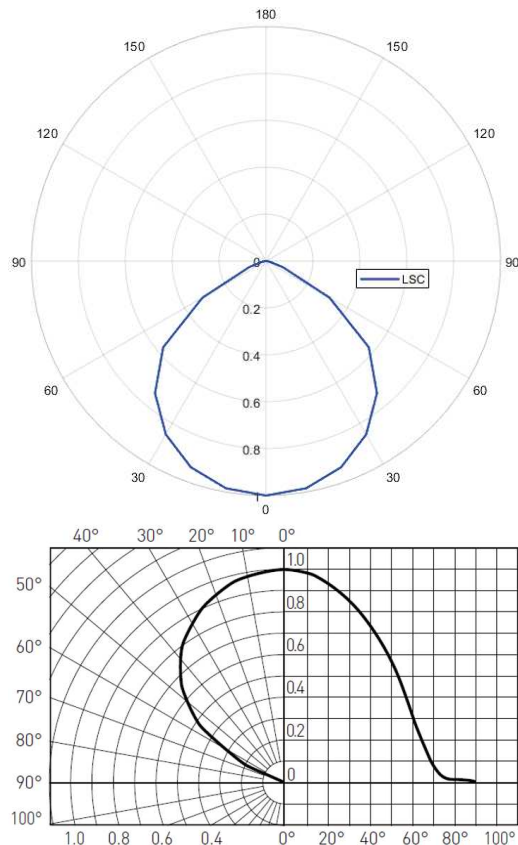
FIGURE 2.3: Characteristic curve (LSC) showing the sensor's sensitivity in function of the angle of the incident of the light. Both figure show the same LSC in different angular w.r.t. weight of light intensity perception forms.

There is a broad range of detector options available for use, choosing the correct detector it is a matter of the application that is going to be used for in terms of spectral response, environment, quality, accuracy and cost. Characteristics such as the signal-to-noise ratio, amplitude, time response, frequency bandwidth, the detector system's linearity range, field of view, noise equivalent power, window transmission, as well as other factors, are only a few that directly affect the choice of detector and the desired measurement for the needs of a corresponding application. Thus, a photometric instrument can be considered either as a stand-alone system such as an illuminance or luminance meter, or combined with additional equipment, *e.g.* an integrating sphere, in order to structure the light distribution properties and measurement system of a light source or the light propagation within an environment.

Illuminance meters (Luxmeters)

The illuminance meters or luxmeters as they are most commonly known, are structured from a photodiode with a photopic correction filter (the photopic response filters try simulate the spectral sensitivity curve of the human visual perception). Thereafter, the photodiode is connected to an operational amplifier together with a display where the final measurement is displayed. The luxmeters are frequently used to measure the illuminance incident on a surface such as a desk, a wall, *etc*. Luxmeters, are normally hand-held and

provide point-to-point measurements. Meaning that the operator in order to obtain a dense sampling of a scene needs to repeat the measurement several times, swaying for angular and spatial completeness and as a consequence turning out in a very time and cost expensive procedure (imagine applying this procedure in a 10-story building).

Luxmeters measure the incident light (illuminance) based on what is called the Luxmeter Sensitivity Curve (LSC) which most of the times is related with the curve of the photopic correction filter mentioned earlier. This distribution describes the luxmeter sensor's sensitivity to lighting, depending on the lighting angle and distance as well as to the manufacturing characteristics, see Fig. 2.3. The lines of the curve show how and where the light is measured from the detector. We can deduce the light intensity in the various angular orientations to which the light is perceived. This is always measured starting from the centre point of the light source. The weighted measuring distribution of the light intensity is indicated on the circles in the graph. The larger the circle, the higher the weighting value.

Luminance meters (Telephotometers / Lightmeters)

The luminance meters are essentially luxmeters with the addition of suitable optics in order to picture the targeting surface onto the detector. The possibility of viewing the surface is usually provided so that the user can see and target the area that he wants to measure. Due to the fact of the similarity of this sensor to a telescope, these instruments are also called telephotometers, see Fig. 2.4.

These sensors can change the focal length of the optic lens which by its turn changes the field of view and thus the size of the measurement area. Moreover, some systems have apertures of various sizes to further define the measuring area. The readings are usually in cd/m^2 or cd/ft^2 while the angular measurement varies from seconds of arc to several degrees, *e.g.* $0.25°$ to $10°$.

This sensor is known for its remarkable measurement accuracy, however it features some noteworthy disadvantages. Among these are the narrow angular field of view (FOV); and

FIGURE 2.4: An example of a luminance meter (lightmeter).

the difficulty of changing apertures and graticules for different measurement areas. The sensors of this class usually employ high-quality detectors as well as other features and capabilities which means that are coupled with a higher purchase cost.

Imaging photometers

On the other hand recent developments in imaging devices, *i.e.* cameras, have provided a powerful tool for dense luminance measurements of complete scenes. Cameras equipped with a charge-coupled device (CCD) array are able to capture and digitize (render) electronic images of visual scenes. Considering that the proper controls are applied (linearity of transfer function), the digital image can be used in order to determine the luminance at every point in the scene, corresponding to the pixels of the image captured from the camera's CCD array.

These sensors are operating in the photometric field and a complete photometric capture can be carried out and saved in seconds. This exploits the advantage that the information is provided in digital form, and thus complicated functions of luminance images can be analyzed and reported quickly for uniformity, contrast, spatial characteristics, and other photometric values. This form of photometry requires that many factors need to be handled and controlled within the sensor and its corresponding software (*e.g.* wavelength, aperture,

exposure-time, pixel-location within the image, camera sensitivity settings, and/or individual camera-to-camera differences), so that accurate and high quality measurements to be obtained.

2.3 Light emission devices

Following on the definition by [DBBS06], practically light is an electromagnetic radiation produced by accelerating a charge (*e.g.* electric discharge). Light can be produced in different ways; however the light sources are discriminated in two basic types: the incandescence and luminescence. The former describes the phenomenon when light is produced from a heated solid, *e.g.* by thermal sources such as the sun. The latter describes the light that is produced when an electron emits some of its energy as electromagnetic radiation, such as fluorescence light, where materials absorb energy at some wavelength and emit the energy at some other wavelength. In most rendering algorithms, light is assumed to be emitted from light sources at a particular wavelength and with a particular intensity described by the point based light sources (isotropic) assumption.

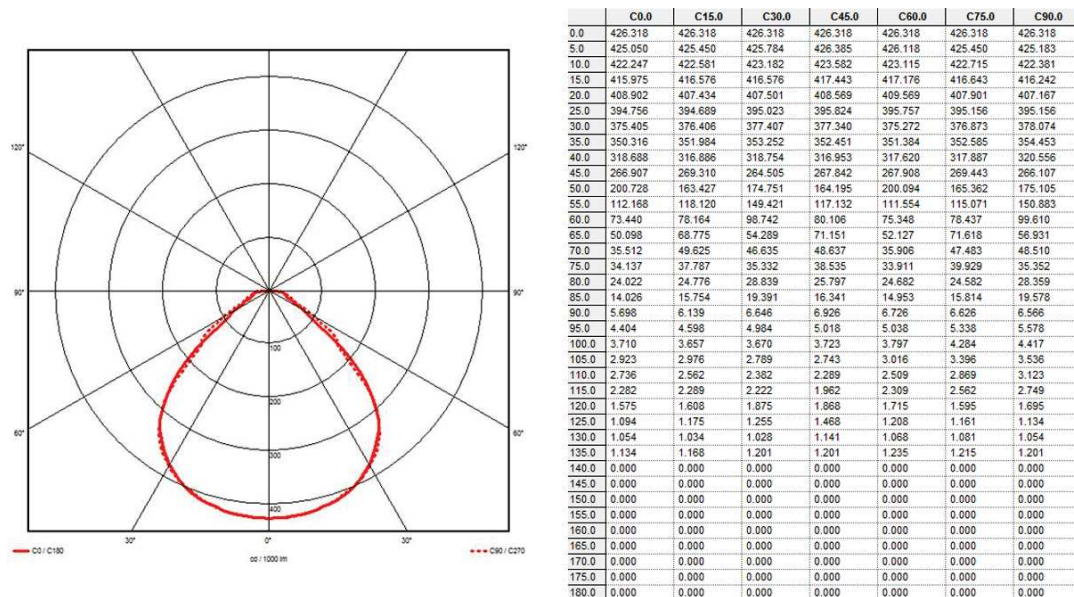|	C0.0	C15.0	C30.0	C45.0	C60.0	C75.0	C90.0
0.0	426.318	426.318	426.318	426.318	426.318	426.318	426.318
5.0	425.050	425.450	425.784	426.385	426.118	425.450	425.183
10.0	422.247	422.581	423.182	423.582	423.115	422.715	422.381
15.0	415.975	416.576	416.576	417.443	417.176	416.643	416.242
20.0	408.902	407.434	407.501	408.569	409.569	407.901	407.167
25.0	394.756	394.689	395.023	395.824	395.757	395.156	395.156
30.0	375.405	376.406	377.407	377.340	375.272	376.873	378.074
35.0	350.316	351.984	353.252	352.451	351.384	352.585	354.453
40.0	318.688	316.886	318.754	316.953	317.620	317.887	320.556
45.0	266.907	269.310	264.505	267.842	267.908	269.443	266.107
50.0	200.728	163.427	174.751	164.195	200.094	165.362	175.105
55.0	112.168	118.120	149.421	117.132	111.554	115.071	150.883
60.0	73.440	78.164	98.742	80.106	75.348	78.437	99.610
65.0	50.098	68.775	54.289	71.151	52.127	71.618	56.931
70.0	35.512	49.625	46.635	48.637	35.906	47.483	48.510
75.0	34.137	37.787	35.332	38.535	33.911	39.929	35.352
80.0	24.022	24.776	28.839	25.797	24.682	24.582	28.359
85.0	14.026	15.754	19.391	16.341	14.953	15.814	19.578
90.0	5.698	6.139	6.646	6.926	6.726	6.626	6.566
95.0	4.404	4.598	4.984	5.018	5.038	5.338	5.578
100.0	3.710	3.657	3.670	3.723	3.797	4.284	4.417
105.0	2.923	2.976	2.789	2.743	3.016	3.396	3.536
110.0	2.736	2.562	2.382	2.289	2.509	2.869	3.123
115.0	2.282	2.289	2.222	1.962	2.309	2.562	2.749
120.0	1.575	1.608	1.875	1.868	1.715	1.595	1.695
125.0	1.094	1.175	1.255	1.468	1.208	1.161	1.134
130.0	1.054	1.034	1.028	1.141	1.068	1.081	1.054
135.0	1.134	1.168	1.201	1.201	1.235	1.215	1.201
140.0	0.000	0.000	0.000	0.000	0.000	0.000	0.000
145.0	0.000	0.000	0.000	0.000	0.000	0.000	0.000
150.0	0.000	0.000	0.000	0.000	0.000	0.000	0.000
155.0	0.000	0.000	0.000	0.000	0.000	0.000	0.000
160.0	0.000	0.000	0.000	0.000	0.000	0.000	0.000
165.0	0.000	0.000	0.000	0.000	0.000	0.000	0.000
170.0	0.000	0.000	0.000	0.000	0.000	0.000	0.000
175.0	0.000	0.000	0.000	0.000	0.000	0.000	0.000
180.0	0.000	0.000	0.000	0.000	0.000	0.000	0.000

FIGURE 2.5: Characteristic curve (LDC) showing the spatial and directional distribution curve of the light source and the corresponding table with the candela values for each specific angular direction.

However users, such as the lighting design engineers, require accurate descriptions of light source distributions that match the physical light sources available in the real world. The directional distributions of typical luminaires is determined by the shape of their associated light fixtures usually provided from the manufacturers. All these distributions could be specified either as functions/distributions and/or as tables, and known as Light Distribution Curve (LDC) see Fig. 2.5. The LDC has a similar behaviour to the LSC described earlier and it shows how and where the light is distributed starting from the origin which is the center of the luminaire. In particular the light intensity can be deduced in the various angular orientations to which the light radiates and propagated into the space. The light intensity is expressed in Candela (cd) and is indicated on the circles in the graph. The larger the circle, the higher the candela value.

2.4　Lighting calculations

Different lighting calculations are applied during the design or light planning process in order to obtain information or to evaluate the performance of the installed lighting system. The lighting calculations are mathematical models of complex physical processes that occur within a lighted space and describe the light propagation. Due to the complexity that such models feature these calculations can never be accurate in every detail, therefore the computations are mainly approximations of the actual phenomena.

Estimating the light propagation from a source to a receiving surface is fundamental to all lighting calculations. This transport is assumed to be through air, which is considered to be non-absorbing and nonscattering. Furthermore, for calculations purposes, in most of the cases the luminaire (*i.e.* light source) is considered as a simple point. However, in reality the light sources have different properties and the photometric measurement of a luminaire (usually done from the manufacturers, using an integrated sphere) is a way to provide these information such as the luminous intensity distribution and the spatial distribution of flux of the luminaire.

Moreover, another important aspect is that often it is necessary to know how light is reflected from a point or an area of a surface. Exitance or radiosity is the simplest case that takes into account this aspect and is measured as flux density leaving the surface. A more complicated case is the luminance, which describes the flux leaving a surface in a particular direction. In both exitance and luminance calculations the reflecting properties of the surface must be known, thus methods for predicting illuminance are combined together with the reflectance information in order to predict exitances and luminances.

Interreflection is the procedure that describes the multiple reflection of the light among the different surfaces of the scene. It is an important aspect of most interior lighting systems, since it is by interreflection the way that light reaches to many interior architectural surfaces giving it its perceptual form, despite the main light source.
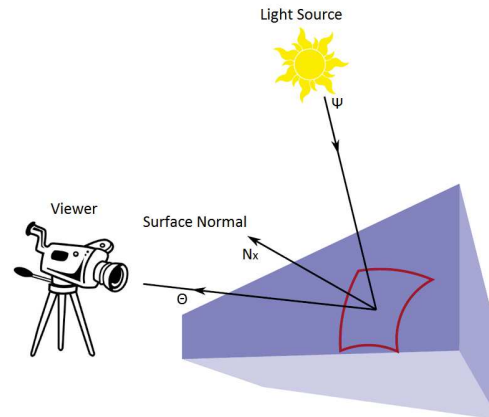
2.4.1　Materials photometric properties

Materials and objects in general interact with light in different ways, and their appearance usually differs given even the same lighting conditions [DBBS06]. Some objects tend to appear as mirrors; while others

FIGURE 2.6: Illustration of the bidirectional reflection distribution function principle.

as diffuse surfaces. The reflective properties of a surface affect the appearance of the object. In the most common scenario, light can reach a surface at a point x with incident direction Ψ and can leave the surface from the same point and exitant direction Θ. The function that defines this relation between the incident and reflected radiance is called the bidirectional reflectance distribution function (BRDF). In general, the BRDF at a point x is defined as the ratio of the differential radiance reflected in an exitant direction (Θ), and the differential irradiance incident through a differential solid angle ($d\omega_\Psi$). In this study we denote BRDF as $f_\rho(x, \Psi \to \Theta)$:

$$f_\rho(x, \Psi \to \Theta) = \frac{dL(x \to \Theta)}{dE(x \leftarrow \Psi)} = \frac{dL(x \to \Theta)}{L(x \leftarrow \Psi)\cos(N_x, \Psi)d\omega_\Psi}, \tag{2.6}$$

where $cos(N_x, \Psi)$ is the cosine of the angle formed by the normal vector at the point x, N_x, and the incident direction vector Ψ, see Fig. 2.6. Based on equation (2.6) we can then extract the relation between incident and reflected radiance as:

$$dL(x \to \Theta) = f_\rho(x, \Psi \to \Theta) \, L(x \leftarrow \Psi) cos(N_x, \Psi) d\omega_\Psi$$
$$L_\rho(x \to \Theta) = \int_{\Omega_x} f_\rho(x, \Psi \to \Theta) \, L(x \leftarrow \Psi) cos(N_x, \Psi) d\omega_\Psi. \tag{2.7}$$

The BRDF is a distribution function which is strictly positive and since it gives the concentration of flux per steradian, it may take any value between zero and infinity. However, often it is more intuitive to work with a quantity that is bounded between 0 and 1. This quantity is called simply *reflectance* (or also known as *albedo* for Lambertian surfaces).

Lambertian diffuse reflection

To define the relationship between the BRDF and reflectance, consider the case of the Lambertian diffuse reflectance [CWH93]. The diffuse reflectance is modeled by the assumption that light is scattered in an isotropic (equally) way in any direction, and regardless of the incident direction. In other words, the BRDF function is constant. Thus, considering the latter we have:

$$L_d(x \to \Theta) = \int_{\Omega_x} f_{\rho,d}(x, \Psi \to \Theta) \, L(x \leftarrow \Psi) cos(N_x, \Psi) d\omega_\Psi$$
$$= f_{\rho,d}(x, \Psi \to \Theta) \int_{\Omega_x} L(x \leftarrow \Psi) cos(N_x, \Psi) d\omega_\Psi \tag{2.8}$$
$$= f_{\rho,d}(x, \Psi \to \Theta) \, E(x \leftarrow \Psi),$$

from the above we can refer to two main outcomes a) that the value of the reflected radiance is proportional to the incident irradiance and b) that the reflected radiance is a constant and hence the same in all directions. This by its turn it can lead us to a relationship where the BRDF is parameterized in terms of the reflectance as follows:

$$f_{\rho,d}(x, \Psi \leftrightarrow \Theta) = \rho_d / \pi. \tag{2.9}$$

Moreover, since the outgoing radiance is constant, the radiosity $B = \pi L_d$ is related to the irradiance by the following formulation:

$$\rho_d = \frac{B}{E}, \tag{2.10}$$

which declares that for the case of the diffuse reflection, the reflectance is equal to the radiosity divided by the irradiance.

2.4.2 The rendering equation: direct reflection and inter-reflection

The reflectance formulation shows us that it is possible to compute the reflected light distribution by taking into account the incident light distribution and the BRDF of the material. Thus, what is left is the important task of computing, the incident light distribution. This is typically referred to as the illumination model [CWH93]. There are two main illumination models:

1. The case of no occlusion and only direct illumination from simple light sources (the simplest one), it is assumed that all the light arrives at the surface and therefore this model depends only on the individual properties of the light sources and the surfaces being lit. This model is often called the local or direct illumination model. Attributes

like shadowing are not initially considered but can be added later on by testing whether a point on a surface is visible to the light source or not.

2. The considerably more difficult case is what is called the indirect illumination. In this case light may derive from any surface in the scene, and it is very important to take into account the aspect of shadowing. Below we present the rendering equation as an extension of the relectance equation where the inter-reflections of the light between surfaces is taken into account.

Thus, the light transport in a three-dimensional environment can be described by the general rendering equation [DBBS06]:

$$L(x \to \Theta) = L_e(x \to \Theta) + L_\rho(x \to \Theta), \tag{2.11}$$

where the outgoing radiance $L(x \to \Theta)$ is expressed in terms of $L_e(x \to \Theta)$, *i.e.* the radiance emitted by the surface at the position x and in the outgoing direction Θ and $L_\rho(x \to \Theta)$, *i.e.* the radiance that is reflected by the surface at the position x in the same direction Θ. Including the reflected radiance definition, *i.e.* Eq. (2.7), the latter becomes:

$$L(x \to \Theta) = L_e(x \to \Theta) + \int_{\Omega_x} f_\rho(x, \Psi \leftrightarrow \Theta) L(x \leftarrow \Psi) cos(N_x, \Psi) dw_\Psi. \tag{2.12}$$

Considering now some simplifications, *e.g.* purely diffuse surfaces (*radiosity or Lambertian assumption*), results into the fact that the self-emitted radiance $L_e(x)$ and the BRDF $f_\rho(x)$ not to depend on directions Θ and Ψ and therefore the BRDF can be taken outside of the integral. The rendering equation then transforms as follows:

$$L(x \to \Theta) = L_e(x \to \Theta) + f_\rho(x) \int_{\Omega_x} L(x \leftarrow \Psi) cos(N_x, \Psi) dw_\Psi. \tag{2.13}$$

As we can see, the incident radiance $L(x \leftarrow \Psi)$ still depends on the incident direction. It actually corresponds to the exitant radiance $L(y)$ emitted towards x by the point y visible from x along the direction Ψ. However the integral above, over the hemisphere Ω_x, can be transformed into an integral over all surfaces S in the scene. The result is an integral equation in which no directions appear anymore:

$$L(x) = L_e(x) + \rho(x) \int_S K(x, y) L(y) dA_y, \tag{2.14}$$

where $K(x, y) = G(x, y) Y(x, y)$ with $G(x, y)$, being the geometry term, which describes the relative geometric dependency of the surfaces at points x and y:

$$G(x, y) = \frac{cos(N_x, \Psi) cos(N_y, -\Psi)}{\pi p_{xy}^2}, \tag{2.15}$$

Θ_{xy} is the direction originating from x to y, while p_{xy}^2 is the square distance between x and y. Finally, $Y(x, y)$ is the visibility predicate (1 if x and y are mutually visible, 0 otherwise).

The rendering equation is used to express the maintenance of light energy at all points of the surfaces in the space. The key feature of such a formulation is the quantity to be computed — *i.e.* in this case, the radiance — which shows up both in the left-hand side as well as under an integral on the right-hand side. Considering the latter, as well as the fact that integral equations are known to be difficult to be solved, since they rarely have a closed-form analytic solution, a numerical method must be used instead.

2.4.3 The radiosity equation

The radiosity equation, a simplified form of the rendering equation, is given by the assumption that all surfaces are Lambertian diffusive reflectors [CWH93].

In a diffusive environment, the radiometric quantities of radiosity and radiance are related as $B(x) = \pi L(x)$ and $B_e(x) = \pi L_e(x)$. If we now multiply both sides of the equation 2.14 by π we will result the following equation which is also know as the the radiosity integral equation [DBBS06]:

$$B(x) = B_e(x) + \rho(x) \int_\Gamma K(x,y)B(y)dA_y. \tag{2.16}$$

The radiosity $B(x)$ describes an random scalar function across the surfaces Γ (*i.e.*, the radiosity function defines a single value at each location on a surface). However, often integral equations like, Equation 2.16 are solved by reducing them to an approximate system of linear equations. This is the same principle as used in Finite Element Method (FEM) modeling [SF08].

If we assume that the radiosity $B(x)$ is constant on each surface i, $B(x) = b_i$, $B_e(x) = e_i$, where $x \in \Gamma_i$. The equation 2.16 can be converted into a linear system as follows:

$$
\begin{aligned}
B(x) &= Be(x) + \rho(x) \int_S K(x,y)B(y)dA_y \\
\Rightarrow \frac{1}{A_i}\int_{S_i} B(x)dA_x &= \frac{1}{A_i}\int_{S_i} Be(x)dA_x + \frac{1}{A_i}\int_{S_i}\int_S \rho(x)K(x,y)B(y)dA_y dA_x \\
\Leftrightarrow \frac{1}{A_i}\int_{S_i} B(x)dA_x &= \frac{1}{A_i}\int_{S_i} Be(x)dA_x + \sum_j \frac{1}{A_i}\int_{S_i}\int_{S_j} \rho(x)K(x,y)B(y)dA_y dA_x \\
\Leftrightarrow b_i &= e_i + \sum_j r_j \frac{1}{A_i}\int_{S_i}\int_{S_j} \rho(x)K(x,y)dA_y dA_x.
\end{aligned}
\tag{2.17}
$$

If we now also make the assumption that the reflectivity is constant over each patch, $\rho(x) = \rho_i$, $x \in S_i$, the following classical radiosity system of equations results in:

$$b_i = e_i + \rho_i \sum_j F_{ij}b_j. \tag{2.18}$$

Furthermore, it should also be considered that the resulting radiosities b_i (Eq. 2.18) are in principle only an approximation for the average radiosities (Eq. 2.16). The actual radiosity $B(y)$, which was replaced by b_j in the equations above, is actually only sometimes partly constant! In practice however, this difference between B_x and b_i rarely can be noticed [DBBS06].

The term F_{ij} is called *patch-to-patch* form factors. The meaning and properties of these form factors are discussed in Section 4.2.2 along with the adaptation of the linear radiosity formulation (Eq. 2.18) for estimating the spatial illumination in real scenes.

2.5 Conclusion

The purpose of this Chapter was to introduce and provide an overview and make aware to the reader all the different theoretical and practical definitions that we make use within this study. Therefore, in this Chapter, we started by presenting definitions to lighting standards and modeling and how are these related to the fields of computer vision and visual computing

based on the behavior of light related to the geometric optics. Thereafter, informational notions regarding the entity of light, the radiometric measuring quantities, equipment and properties of the light and its sources as well as basic lighting calculations have been detailed.

Most of the described definitions and especially the radiative quantities and the radiosity formulation are thoroughly used later as the key element in the developed algorithm in order to provide a camera-aided light estimation solution for real life case scenarios. Cases including real light sources, different scenes with different structure and material properties, have been evaluated as it will be presented in Chapter 4.

Of course there are more theoretical and practical definitions related to light modeling and that are not presented here, *e.g.* more complex BRDF models, local light propagation and illumination models, *etc*. However, we decided to keep this out of the scope of this study for reasons of simplicity.

Chapter 3

Light Source Calibration

Contents

3.1 Introduction

Light source calibration in indoor environments is a key component of light modeling and it is a complicated procedure especially when it involves multiple light sources of different types (*e.g.* artificial, natural), the presence of objects of diverse sizes, shapes and spectral characteristics. A modern lighting system should automatically calibrate itself (light commissioning), assess its own status (which lights are on/off and how dimmed), and allow for the creation or preservation of lighting patterns (adjustability), *e.g.* after the sunset. Such a system does not exist today, nor (real) data, labels, or metrics are available to compare with and foster progress.

Therefore, in this Chapter we introduce LIT, a novel computational framework and benchmark for complex indoor scenes to promote research on three necessary aspects of a smart lighting system. The first goal of the LIT system is to estimate the illumination layout, *i.e.* where the lights are positioned and what part of the scene they illuminate, *i.e.* automatic calibration, which we call *Light Source Estimation* (LIT-EST). Next, we name *Light Source Identification* (LIT-ID) the understanding of the current lighting pattern in the scene, status awareness, *i.e.* which allows the system to be aware about which lights are on and to what intensity level. Finally, we propose a third aspect which we name *ReLighting* (RE-LIT) as an application usage of the first two aspects, where the main target is to maintain or set an illumination pattern.

The proposed solution to these problems builds upon factorization techniques. In particular, given the additive and non-negative property of light, we consider Non-negative Matrix Factorization (NMF) [LS99] as a suitable decomposition technique for modelling the interaction of light sources in the scene. Concurrently we propose the LIT benchmark, the first applicable to a smart lighting system, comprising over ∼50K time-lapse images across the three defined tasks. We have acquired and manually labelled two time-lapse videos, in two scenes with natural light and different luminaire setups. Labels include the presence of a light source[1] and its dimming level.
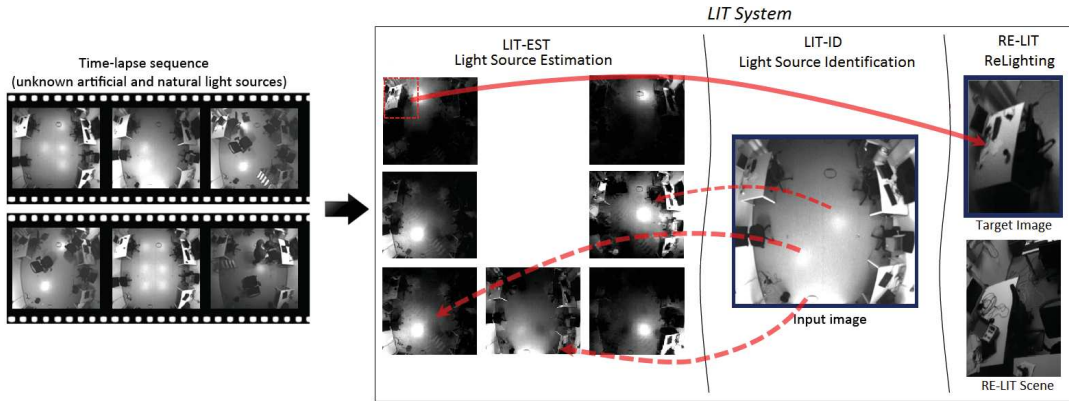

FIGURE 3.1: Overview of the LIT system. (*Left*) Time-lapse videos, depicting difference light sources active at different times. (*Mid-left*) LIT-EST estimates the single light sources of any type. (*Mid-right*) LIT-ID identifies which light sources are active and to what intensity. (*Right*) RE-LIT uses the sources to replicate the illumination pattern of a target image into a new scene.

3.2 Light sources estimation

LIT-EST models indoor lighting variations by estimating each light source independently from the other, in an unsupervised manner (without need for annotated data). More in details, we process time-lapse video sequences (**training set**, see Sec. A.1) and extract images of single lighting modules (luminaires or the natural light), although possibly no frame showed them in isolation.

More formally, we consider an image I to be the result of an image formation process, completely defined by the set $\mathcal{L}$ of l images of the individual light sources $\mathcal{L} = \{\mathtt{S}_i\}_{i=1}^{l}$ that are active for the image I. We assume that an unknown function exists, such that:

$$\mathtt{I} = f_I(\mathtt{S}_1, \ldots, \mathtt{S}_L), \qquad (3.1)$$

where f_I is the function of the image formation process given by the light sources $\mathtt{S}_i$.

Note that the above formalism and assumption is general and common to several lighting models as used in photometric stereo where each image $\mathtt{S}_i$ is a linear representation, for instance, of surface normals (surface normals and albedo [HAYA94], higher order spherical harmonics [BJK07], reflectance and illumination components [LN16]). Also note that a simple but effective model for f_I is a weighted linear summation *i.e.* $I = w_1\mathtt{S}_1 + \ldots + w_n\mathtt{S}_n$.

---

[1]We mean by *light sources* the natural light and the luminaires, providing light at a given intensity, but not a PC monitor for instance.

We name $\mathtt{I}_{EST}$ the distinct subset of the LIT dataset which is used for the LIT-EST light source estimation. Note that the LIT-EST process is unsupervised, but we collect the true light source appearances (ground truth) for benchmarking its quality.

Algorithm 1: LIT-EST algorithm overview.

1 Training image sequence $\mathcal{T} = \{\mathtt{I}_1, \ldots, \mathtt{I}_n\}$;

2 Create matrix $\mathtt{V} \in \mathbb{R}^{m \times n}$ such that: $\mathtt{V} = \begin{bmatrix} vec(\mathtt{I}_1) & \cdots & vec(\mathtt{I}_n) \end{bmatrix}$;

3 Estimate $[\mathtt{W}, \mathtt{H}]$ minimizing $\min\limits_{\mathtt{W},\mathtt{H} \geq 0} \|\mathtt{V} - \mathtt{W}\,\mathtt{H}\|^2$ where $\mathtt{W} \in \mathbb{R}^{m \times l}$, $\mathtt{H} \in \mathbb{R}^{l \times n}$ and
  $1 \leq l \leq \min(m, n)$;

4 **for** $i = 1 : l$ **do**

5   For each basis i, get the reconstructed images: $\hat{\mathtt{V}}_i = \mathbf{w}_i \mathbf{h}_i^\top$;

6   Compute the activation as $AoI_{(i)} = \|\hat{\mathtt{V}}_{(i)}\|$;

7   Find the maximum activation $idx_{max} = \arg\max(AoI_{(i)})$;

8   Assign $\tilde{\mathtt{S}}_i = \hat{\mathtt{V}}_{(i, idx_{max})}$;

9 $\tilde{\mathcal{L}} = \{\tilde{\mathtt{S}}_1, \ldots, \tilde{\mathtt{S}}_L\}$;

Thus, given a set of n $\mathtt{I}_{EST}$ images extracted from a time-lapse video

$$\mathcal{T} = \{\mathtt{I}_{EST1}, \ldots, \mathtt{I}_{ESTn}\}, \tag{3.2}$$

the goal of the LIT-EST application is to find, in an unsupervised way, the set $\mathcal{L}$ of l source images that explain the illumination of the scene (*c.f.* Eq. 3.1). This resembles the classical source separation problem, commonly approached with PCA, ICA and the like. For selecting the best technique of this family, we should note two main physical properties of light: **(a)** light from general illumination sources is incoherent and the light intensity from multiple sources combines linearly; **(b)** each light source contributes to the image formation with a non-negative weight term. This translates into an image formation function where non-negative light bases are combined with non-negative weights. Such conditions elects Non-negative Matrix Factorization (NMF) [LS99] as the natural choice for light decomposition.

3.2.1 Non-negative matrix factorization

The NMF problem statement is as follows: given an $m \times n$ non-negative matrix $\mathtt{V}$, NMF finds two compact non-negative matrices $\mathtt{W}_{m \times r}$ (containing r horizontally stacked m-dimensional bases), and $\mathtt{H}_{r \times n}$ (containing for each image the coefficients associated to the basis), such that:

$$\min\limits_{\mathtt{W},\mathtt{H}} \|\mathtt{V} - \mathtt{W}\,\mathtt{H}\|_F, \text{ subject to } \mathtt{W} \geq 0 \text{ and } \mathtt{H} \geq 0. \tag{3.3}$$

In our case the positive data matrix $\mathtt{V}$ is given by the images in the video sequence by vectorizing them (m being the number of pixels) and concatenating all n training dataset:

$$\mathtt{V} = \begin{bmatrix} vec(\mathtt{I}_1) & \cdots & vec(\mathtt{I}_n) \end{bmatrix} = \begin{bmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{bmatrix},$$

where $vec(\bullet)$ is the (column-wise) vectorizing operator. The positive matrix $\mathtt{W}$ contains the r light bases as a set of images containing each light sources (the value of r usually is predicted through a cross-validation procedure), while the matrix $\mathtt{H}$ holds the linear positive weights that sum each basis in order to obtain $\mathtt{V}$.

NMF has been successfully applied for a variety of problems including document clustering [XLG03], hyper-spectral image unmixing [CG15], sound separation [SMAR04]. The general NMF formulation has however drawbacks: it is a non-convex optimization problems,

thus subject to local solutions and its solution is not unique, unless the problem formulation
is complemented with additional constraints such as smoothness, sparsity, orthogonality, *etc*.
[VAVA07; SMAR04; HAMM13; GG10; CG15].

Motivated by their efficiency and robustness to noise, we evaluate three recent NMF
techniques:

- **Diagonalized newton algorithm for non-negative matrix factorization (DNA-NMF)
  [HAMM13].** The DNA-NMF shows a faster convergence compared to NMF thanks
  to the diagonalized Newton optimization. The estimation with DNA-NMF remains
  tractable even for high-rank problems.

- **Block principal pivoting non-negative matrix factorization (BPP-NMF) [KHP14].**
  BPP-NMF improves further the computational burden by allowing the exchanges of
  multiple variables between working sets. This method also allows a constrained solu-
  tion by inducing a sparsity prior with the use of l_1-norm.

- **Non-negative matrix under-approximation with priors (PNMU) [CG15].** PNMU
  extends standard NMF by solving it sequentially and incorporating spatial and sparsity
  priors. This biases the algorithm towards bases that have active pixels (those illumi-
  nated by the light source) spatially close, and in general fewer active pixels per basis.

Further to the NMF approaches, we experimented with PCA and ICA decompositions
as baseline approaches. The comparative evaluation should highlight the importance of the
non-negativity constraints in the LIT optimization.

3.2.2 Light source image reconstruction.

The light decomposition bases W computed from Eq.(3.3) via NMF techniques are important
for the light source estimation quality. The bases should in fact disentangle the light sources,
such that each basis would only contribute to a single light source.

Matrices W are however not yet light source images because their values need to be
weighted by the weight matrix H. For example, small bases values may still generate strong
light sources if their weight is large. This means that we should find a reference image for
each basis to drive the light source estimation, *i.e.* the weighted basis. Since the set of images
I_{EST} may contain light sources partially lit, it is important to identify the one image where
each weighted basis is strongest (across all pixels), therefore corresponding to one where the
light source is fully lit.

In more detail, let $\mathbf{w}_i$ be the basis vector W that contributes to the light source S_i and
let $\mathbf{h}_i$ be the row vector in H related to the corresponding coefficients. The reconstructed
images $\hat{V}_i = \mathbf{w}_i \mathbf{h}_i^\top$ is an $m \times n$ matrix for each of the i-th basis. By concatenating all $\hat{V}_i$,
we formulate the 3D matrix such that $\hat{V}_{k,j,i}$ correspond to the values of the k-th pixel in the
j-th image of the i-th basis. The Activation of Image (*AoI*) pixel map is given by evaluating,
for each image j in the training set, following $AoI_{j,i} = \|\hat{V}_{j,i}\|$ evaluated for all k. Hence, the
maximum activation would then be given by:

$$\bar{j} = \arg\max_j AoI_{(j,i)}, \tag{3.4}$$

for each i such that $\tilde{S}_i = \hat{V}_{\bar{j},i}$ is the desired light source image.

3.3 Light sources identification

LIT-ID identifies which light sources are active and to what intensity, *i.e.* the dimming level
of a luminaire. More formally, LIT-ID estimates a light activation vector $\mathbf{a}$ for each image I

of the **testing set**:

$$\mathbf{a} = [a_1, \ldots, a_l]^\top, \qquad a_i \in [0,1] \tag{3.5}$$

where a_i indicates whether the light source $\mathtt{S}_i$ is switched off ($a_i = 0$), lit to its maximum intensity ($a_i = 1$), or partially powered, $0 < a_i < 1$. Likewise to LIT-EST, the LIT-ID process is also unsupervised. We collect though a distinct set of images $\mathtt{I}_{ID}$, which we annotated with ground truth activation vectors for the sake of benchmarking.

The $\tilde{\mathcal{L}}$ light source images estimated in the previous phase are then used to identify which light sources are active for each test image (LIT-ID). This means estimating the activation vector $\mathbf{a}_{ID}$ indicating the intensity of the $\tilde{\mathcal{L}}$ light sources.

We formalized this step as a constrained Least Squares problem since $\mathbf{a}$ is bounded between 0 and 1. First, given all the light source images, we form the $n \times l$ matrix Λ as:

$$\Lambda = \left[\begin{array}{ccc} vec(\tilde{\mathtt{S}}_1) & \cdots & vec(\tilde{\mathtt{S}}_l) \end{array} \right].$$

Given a test image $\mathtt{I}_{ID}$, the estimated activations $\mathbf{a}_{ID}$ are computed as the solution of:

$$\min_{\mathbf{a}_{ID}} \frac{1}{2} \|\Lambda \cdot \mathbf{a}_{ID} - vec(\mathtt{I}_{ID})\|_2^2, \ \text{ subject to } \ 0 \le \mathbf{a}_{ID} \le 1.$$

The constrained optimization ensures meaningful non-negative activation of light sources.

3.4 Relighting

RE-LIT stands for the problem of setting the illumination pattern of a scene according to a target image or to a specific illumination pattern. For example this could be lighting a desk in the same way across different scenes or mimicking the natural light by only using the artificial lighting system. This procedure falls within the applications of a smart lighting system where there would be need for maintaining the spatial illumination in specific levels or to fit to the corresponding human activity within the environment.

We benchmark the latter quantitatively, by first estimating the light sources of the target scene with LIT-EST (from $\mathtt{I}_{EST}$, and then reconstructing the images of a distinct set $\mathtt{I}_{RE}$ that maintains the desired lighting pattern. In our experiments, $\mathtt{I}_{RE}$ is a subset of $\mathtt{I}_{ID}$ containing natural light or not according to the application. RE-LIT is benchmarked by directly comparing the target and re-lit scene images.

We assume here to have previously computed the $\tilde{\mathcal{L}}$ light source images with LIT-EST. However, differently from the LIT-ID, the target image (the one providing the lighting pattern to replicate) may in general have been acquired within a different reference system, *e.g.* different luminaire and furniture layout, or within the same scene under different conditions, *e.g.* the target image may contain natural light but require its replication after the sunset with luminaires only.

Let us denote Λ the $n \times l$ matrix, containing the vectorized $\tilde{\mathcal{L}}$ light sources. We further denote $\mathtt{P}$ the target image, providing the desired lighting pattern. The ReLighting activations $\mathbf{a}_{RE}$ are estimated by solving the constrained optimization problem:

$$\min_{\mathbf{a}_{RE}} \frac{1}{2} \|\Lambda \cdot \mathbf{a}_{RE} - vec(\mathtt{P})\|_2^2, \ \text{ subject to } \ 0 \le \mathbf{a}_{RE} \le 1.$$

The $\mathbf{a}_{RE}$ activations may be employed, as we do here, to synthesize the relit image: $\mathtt{I}_{RE} = \Lambda\, \mathbf{a}_{RE}$. The same $\mathbf{a}_{RE}$ vector may be input into the lighting system as luminaire dimming coefficients, for lighting the room as desired.

3.5   Metrics

In order to evaluate and test the performance, we defined metrics for each stage of the LIT system.

LIT-EST metric

This is given by the distance between the estimated light source images $\tilde{S}_i$ for $i = 1, \ldots, l$ and their ground truth images S_i (we extract the ground truth images either by a random selection of frames from the dataset corresponding to the individual light sources in the case of the static scene, or the average of a sequence of frames over time in the case of the dynamic scenes), which we define as the LIT-EST error measure ε_{EST}:

$$\varepsilon_{\text{EST}} = \frac{1}{l} \sum (1 - SSIM(S_i, \tilde{S}_i)), \qquad (3.6)$$

where SSIM stands for the Structural Similarity Index [WBSS04] defined as:

$$SSIM(S_i, \tilde{S}_i) = \frac{(2\mu_{S_i}\mu_{\tilde{S}_i} + c_1)(2\sigma_{S_i\tilde{S}_i} + c_2)}{(\mu_{S_i}^2 + \mu_{\tilde{S}_i}^2 + c_1)(\sigma_{S_i}^2 + \sigma_{\tilde{S}_i}^2 + c_2)} \qquad (3.7)$$

with μ, σ being the average and variance over the considered image, and $c_{1,2}$ a small constant for the division stability. The SSIM index specifically tests how similar two images are in terms of luminance, contrast and structure (*c.f.* details in [WBSS04]).

Prior to compare true and estimated light source images in Eq. (3.6), we first match them according to their best SSIM score. In case of competing matches (*e.g.* two sources matching the same true source), we leave the worse unassigned and penalize the term with a 0 similarity.

LIT-ID metric

We compare the estimated $\tilde{\mathbf{a}}$ and ground truth activation vector $\mathbf{a}$ in terms of normalized l_1-norm distance:

$$\varepsilon_{\text{ID}} = \frac{1}{n} \sum_{i=1}^{n} \frac{\|\mathbf{a}_i - \tilde{\mathbf{a}}_i\|_1}{l} \qquad (3.8)$$

where l is the number of light sources and n is the number of images in $\mathtt{I}_{ID}$.

RE-LIT metric

We measure the distance between a target image $\mathtt{I}_{RE}$ (*e.g.* an image containing the natural light) and the re-lit image estimate $\tilde{\mathtt{I}}_{RE}$ (*e.g.* generated with the sole artificial lights, estimated from $\mathtt{I}_{EST}$), in terms of SSIM such that:

$$\varepsilon_{\text{RE}} = \frac{1}{l} \sum (1 - SSIM(\tilde{\mathtt{I}}_{RE}, \mathtt{I}_{RE})) \qquad (3.9)$$

3.6   Dataset

The LIT system uses a set of time-lapse images showing varying light conditions and changes in the indoor area structure due to objects displacement and human activity. These images are used to first extract an indoor lighting model and then perform identification of the single sources for active relighting of the indoor scene. The system consists of three parts, as shown in Fig. 3.1.

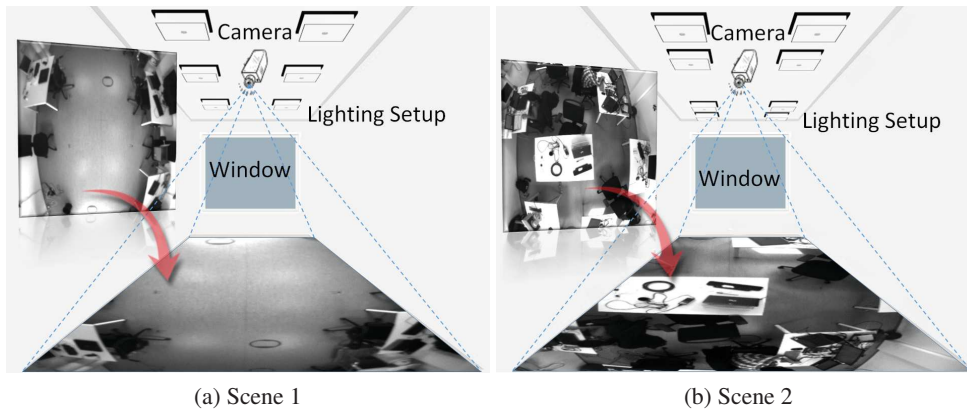(a) Scene 1   (b) Scene 2

FIGURE 3.2: 3D models of the two indoor scenes used for the data acquisition. Images are captured in raw format, presenting the raw luminance that reaches to the sensor CMOS.

We considered two scenes, as illustrated in Fig. 3.2, recorded time-lapse videos for an entire day using a fixed camera and labelled the active lights and their intensity for each frame. Both scene data is used for LIT-EST and LIT-ID. We only use Scene 2 data for RE-LIT because only luminaires in Scene 2 are dimmable, thus settable to floating point activations, necessary for re-lighting.

Scene 1 data consists of $33,280$ images ($15,310$ assigned to $\mathtt{I}_{EST}$ for light source estimation, $16,450$ in $\mathtt{I}_{ID}$ for identification). Scene 2 data contains $20,700$ images (9450 in $\mathtt{I}_{EST}$, 11250 in $\mathtt{I}_{ID}$). A subset of 2600 images from $\mathtt{I}_{ID}$ are used for $\mathtt{I}_{RE}$. The $\mathtt{I}_{EST}$ sets do not contain images of single sources (*i.e.* one luminaire active only), but we provide them separately for ablation studies (*c.f.* below).

Note that both $\mathtt{I}_{EST}$ and $\mathtt{I}_{ID}$ are ground-truth-annotated. Since we are targeting an unsupervised solution, we have used both entire sets (separately) for benchmarking the tasks. A train+test separation or each though is possible to allow learning-based techniques to be used.

3.6.1 Ablation data-subsets

We proposed and evaluated five subsets for ablation studies.

LIT-Artificial

This excludes all frames with natural light, which are supposedly more difficult for the processing. This contains $11,020$ $\mathtt{I}_{EST}$ images and $11,940$ $\mathtt{I}_{ID}$ frames for Scene 1. For Scene 2 the corresponding $\mathtt{I}_{EST}$ and $\mathtt{I}_{ID}$ frames are $16,500$ and $18,000$ respectively.

LIT-Static

This excludes all frames with motion of people, or objects. The Scene 1 *LIT-Static* comprises $7,240$ $\mathtt{I}_{EST}$ images and $7,490$ $\mathtt{I}_{ID}$ frames respectively. $\mathtt{I}_{EST}$ and $\mathtt{I}_{ID}$ sets for Scene 2 contain $8,900$ and $9,700$ respectively.

LIT-Artificial-Static

This is a subset of the *LIT-Artificial* where, having excluded frames featuring motion. This simplest scenario comprises $2,900$ I_{EST} and $3,400$ I_{ID} images for Scene 1, and $7,700$ I_{EST} and $8,500$ I_{ID} images for Scene 2.

LIT-Sources

This further adds to the full LIT the individual light sources (one light activated at a time), so as to test the decomposition techniques capability to disentangle lights. The *LIT-Sources* comprises $16,640$ I_{EST} and $16,450$ I_{ID} frames for Scene 1, $20,700$ I_{EST} images and $20,600$ I_{ID} images for Scene 2.

3.7 LIT system evaluation

Here we analyze the methods introduced in the previous sections for the tasks of LIT-EST, LIT-ID and RE-LIT.
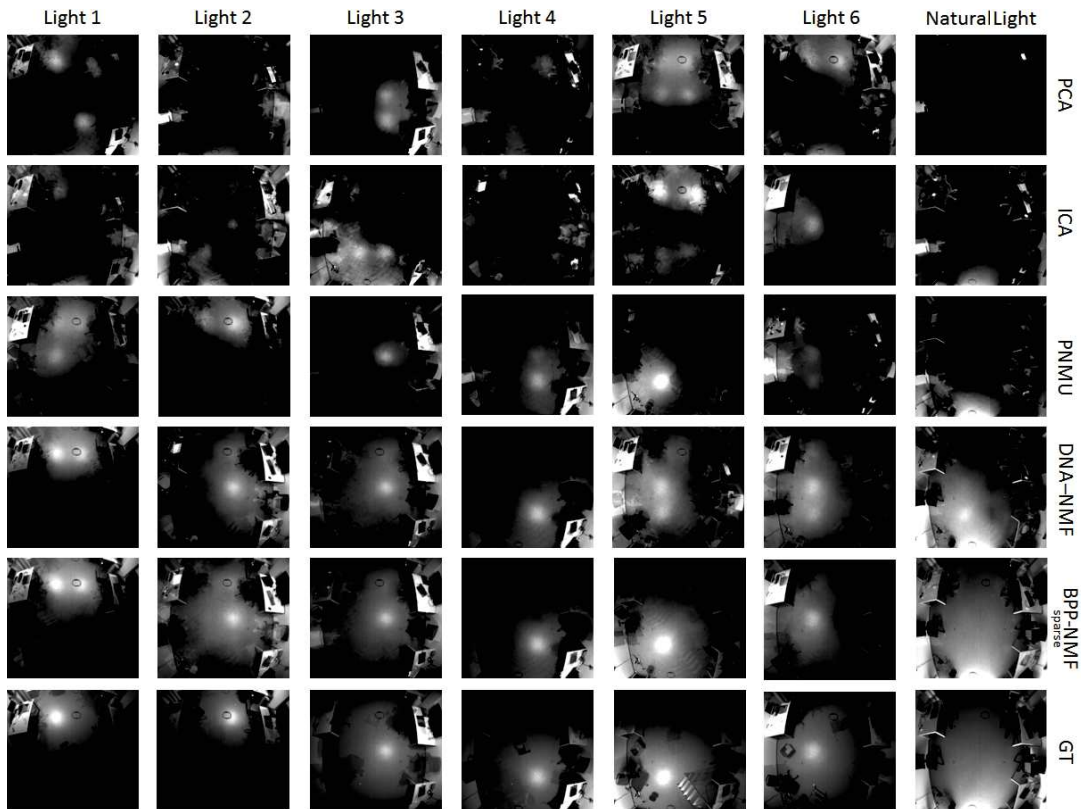

FIGURE 3.3: Light source estimates for each selected technique on the LIT-All set and subsets of Scene 1. PNMU and BPP-NMF techniques are shown only by the best performing alternatives, see Tables 3.1, 3.2. Column-wise, correct light source images should look alike to the ground truth, GT *(last row)*.

3.7.1 Light source estimation results (LIT-EST)

Table 3.1 reports the light source estimation errors ε_{EST}, measuring the quality of the computed light source set $\tilde{\mathcal{L}}$ with the source images of Scene 1 visualized in Figure 3.3.

In the first row of the table (Scene 1, LIT-All), one can immediately notice the inferiority of PCA and ICA, which confirms the importance of including non-negativity constraints in the modeling. Looking at the NMF techniques, BPP-NMF emerges as the best technique with an image reconstruction error of 31.2%. PNMU follows it tightly at the distance of $\sim$ 3%, while DNA-NMF under-performs by $\sim$ 13%. This is confirmed by the number of sources which the technique manages to reconstruct and match to the ground truth, 6 out of 7 for all but the DNA-NMF (5 out of 7).

Notably, sparsity may improve results, but only if carefully designed. As an example, the BPP-NMF-sparse improves by $\sim$ 3% on BPP-NMF, which results in the best reconstructions. By contrast, PNMU-sparse degrades PNMU by $\sim$ 4%. It is plausible that this is influenced by the type of sparsity, l_1-norm for the BPP-NMF-sparse (seemingly preferrable), l_2-norm for the PNMU-sparse.

Ablation studies analysis

First, we consider the influence of the natural light by removing it from the LIT-All benchmark (LIT-Artificial). The rank of the techniques does not change, confirming our earlier conclusions, but the decrease in performance of all techniques w.r.t. LIT-All (*e.g.* BPP-NMF from 28.9% to 31.4% error) is somewhat surprising. Intuitively, the natural light is more difficult to model, as coming from a larger source (the sun, through a large window) and changing in the course the day. We question the puzzling performance decrease by visual inspection of the data and we realize that most scene motion (people, chairs, objects) takes place after the sunset, apparently a more active working time in the room. We conclude from the LIT-Artificial that modelling the scene changes is a harder task for matrix factorization techniques, compared to the natural light variability.

This motivates the second ablation study (third row), whereby we remove any motion from the scene (LIT-Static). BPP-NMF-sparse remains the best with 24.1% error and improves w.r.t. LIT-All by $\sim$ 5%. Indeed motion is a difficult element to model and there is much motion in our LIT, which challenges the selected techniques. Somewhat surprisingly, the worse performers (PCA, ICA, PNMU-sparse) get worse, confirming the importance of non-negativity and of a correct sparsity prior.

Finally, in LIT-Artificial-Static we remove both motion and the natural light from LIT, which results in most techniques "solving" the decomposition correctly (*e.g.* a residual 3% error for BPP-NMF). This resonates with our earlier observations.

While in LIT-All the techniques should reconstruct the light sources without ever observing them in isolation (each frame features at least two lights switched on), LIT-Sources test

ε_{EST}	Scene 1							Scene 2
	PCA	**ICA**	**PNMU**	**PNMU sparse**	**DNA-NMF**	**BPP-NMF**	**BPP-NMF sparse**	**BPP-NMF sparse**
LIT-All	0.6526 (4/7)	0.6063 (5/7)	0.3457 (6/7)	0.3887 (6/7)	0.4468 (5/7)	0.3119 (6/7)	**0.2877** **(6/7)**	0.5872 (6/9)
LIT Artificial	0.5856 (4/6)	0.5912 (5/6)	0.4022 (5/6)	0.3981 (5/6)	0.5058 (4/6)	0.3298 (5/6)	**0.3147** **(5/6)**	0.5584 (5/8)
LIT Static	0.7382 (4/7)	0.6406 (5/7)	0.3315 (6/7)	0.5750 (4/7)	0.3121 (6/7)	0.3055 (6/7)	**0.2409** **(6/7)**	0.2928 (7/9)
LIT Artificial-Static	0.5867 (4/6)	0.7607 (3/6)	0.3188 (5/6)	0.3514 (5/6)	0.0578 (6/6)	**0.0267** **(6/6)**	0.0289 (6/6)	0.2775 (7/8)
LIT Sources	0.6018 (5/7)	0.5070 (6/7)	0.3500 (6/7)	0.3964 (6/7)	0.2912 (6/7)	0.2963 (6/7)	**0.2653** **(6/7)**	0.5333 (7/9)

TABLE 3.1: Comparative evaluation of light source reconstruction techniques over the LIT-All set and subsets. $\varepsilon_{EST} \in [0, 1]$ is the error measures *(lower is better)*. Number in parentheses indicate the successfully matched light sources, see Section 3.7.1.

ε_{ID}	Scene 1								Scene 2	
	PCA	**ICA**	**PNMU**	**PNMU sparse**	**DNA-NMF**	**BPP-NMF**	**BPP-NMF sparse**	**Oracle**	**BPP-NMF sparse**	**Oracle**
LIT-All	0.3647	0.3470	0.2727	0.3149	0.3153	0.2537	**0.2405**	0.0756	0.2346	0.1056
LIT Artificial	0.3754	0.3722	0.3523	0.3402	0.3621	0.2773	**0.2662**	0.0627	0.2527	0.0829
LIT Static	0.3969	0.3653	0.2807	0.3893	0.2526	0.2469	**0.2165**	0.0652	0.1503	0.0413
LIT Artificial-Static	0.3527	0.4110	0.2933	0.2531	0.1164	0.0920	**0.0375**	0.0271	0.1574	0.0316
LIT Sources	0.3811	0.32224	0.2921	0.3306	0.2721	0.2500	**0.2311**	0.0756	0.2318	0.1056

TABLE 3.2: Comparative evaluation of selected techniques in terms of light activation error ε_{ID} over the LIT-All set and subsets. $\varepsilon_{ID} \in [0,1]$ *(lower is better)*, see Section 3.7.2.

whether this would make a difference. The results match the intuition, though with marginal gain (*e.g.* BPP-NMF improving by 2%).

Scene 2

The same experiment, repeated on Scene 2, allows for the same conclusions as the errors of BPP-NMF-sparse, also here perform best (see Table 3.1 last column). Note the general higher error on LIT-All (58.7%), now larger than LIT-Artificial (by visual inspection, Scene 2 contained more motion in the day time). In fact, Scene 2 is more difficult "by design", with 2 additional luminaires, further away from the camera view. This is witnessed by the non-trivial error on LIT-Artificial-Static (27.7%). Complete results for Scene 2 are reported in Appendix A.

3.7.2   Light identification results (LIT-ID)

The light source activation errors ε_{ID} for the selected techniques, applied on Scene 1, are reported in Table 3.2. The scores question the capability to identify the lighting setups in all LIT-ID frames, given the reconstructed light sources.

Table 3.2 brings to similar conclusions as from Table 3.1, with BPP-NMF-sparse being the best performer (24% error, LIT-All row), both due to the superiority of the BPP-NMF technique (*c.f.* the higher errors of PNMU, DNA-NMF, PCA, ICA) and the type of adopted l_1-norm sparsity (*c.f.* the slightly worse performance of BPP-NMF). Again the LIT-Artificial appears to be more challenging for all techniques due to the after-sunset motion (*c.f.* second row); LIT-Static is in fact an easier task (BPP-NMF reducing its error to 21.6%); LIT-Static-Artificial just leaves most techniques with a residual error (3.7% for BPP-NMF).

Since the activation errors depend on the estimated light sources, we further compare the techniques to the Oracle (final column), which "knows" the true light sources. We notice though that the Oracle error on LIT-All is 7.6%, meaning that, even after a perfect light calibration, the light scene understanding still needs further research.

Second, the natural light emerges as the largest challenge (the Oracle error on LIT-Artificial drops to 6.2%). While the motion troubles the light source estimation the most (*c.f.* 3.7.1), the understanding of the scene illumination is most complex due to the natural light because it comes from a from the larger window area, it traverses the entire scene and it crosses all other artificial lights.

Scene 2

Testing on Scene 2 (Table 3.2, last two columns) brings to similar conclusions. As discussed in Sec. 3.7.1 and we show in Appendix A, Scene 2 is more difficult, which is witnessed by the larger error in the LIT-Artifical-Static (15.7%).
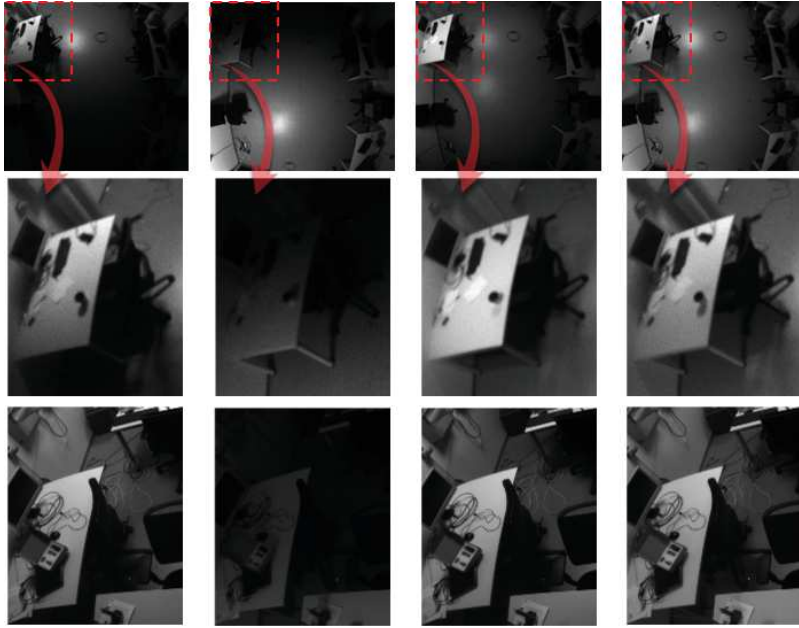
FIGURE 3.4: Relighting sample. Scene 1 lighting scenario *(top row, zoom in middle row)* is re-created in Scene 2 *(bottom row)*, with different luminaires.

3.7.3 Relighting results (RE-LIT)

We proposed two sample applications for the introduced system: a) maintaining a lighting pattern when the illumination setup changes and b) compensating the lack of natural light, by synthesizing its effect with the installed luminaires.

Towards the first application, in Figure 3.4 we show qualitatively how the illumination level on a desk in Scene 1 (first two rows) can be replicated onto a corresponding desk in Scene 2 (third row). The gradient and overall illumination level should be as similar as possible along the columns (for each light scenario).

(a) Target natural light   (b) Synthesis with artificial lights
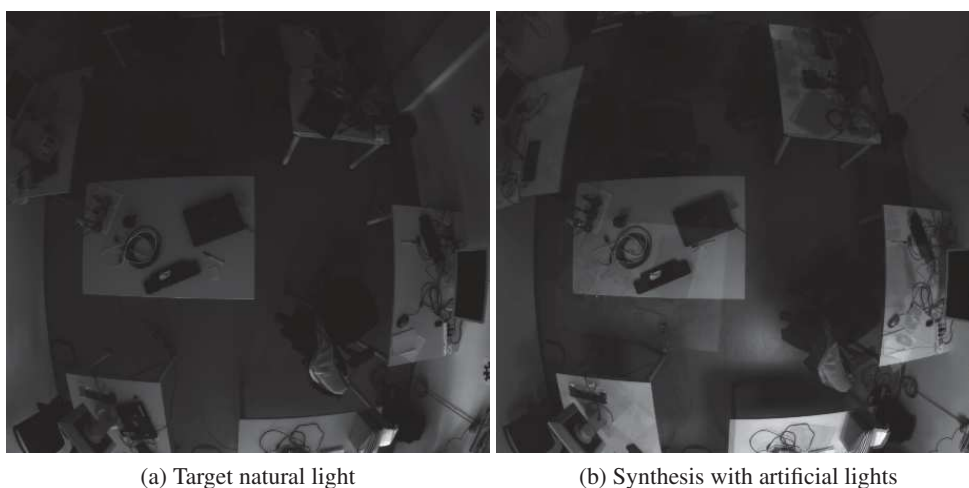
FIGURE 3.5: Reconstruction of natural light sample by the artificial sources, estimated with BPP-NMF for Scene 2.

For the second application task, the natural light synthesis, we consider Scene 2, because

it offers dimmable luminaires (where their light intensity can be adjusted). As described in Section 3.4, the artificial light sources (estimated from the LIT-Artificial set) are combined to compensate for the missing natural light. Figure 3.5 illustrates one such example, whereby the sole natural source is synthesized.

We measured the average relighting error ε_{RE} over all $\mathtt{I}_{ID}$ images (those containing natural lights). BPP-NMF-sparse achieves a **14%** error, while the oracle (*i.e.* the true artificial light sources) gives an error of **13%**. The sole 1% gain of the oracle indicates that the light source estimation of BPP-NMF-sparse is not the performance bottleneck. Rather most error is due to the complex reflection patterns which the natural light casts in the scene.

3.8   Conclusion

In this Chapter, we presented a system for light decomposition and identification based on factorization techniques. We have shown how the natural features of light, *i.e.* of being non-negative and additive, could be used by considering the NMF algorithm as a suitable technique for modeling the interaction of light sources in the scene in the 2D image space.

Thereafter, we have introduced the a benchmarking pipeline: data, labels and metrics, in order to evaluate and foster the performance and the applicability of the proposed system in this new field of research. The proposed solution introduces a new way for light calibration, meaning the localization of the light sources in the space as well as the estimation of the luminous intensity. As we will see later on, in Chapter 4, these are two key components of the model that we used for light estimation. Therefore, this study could help towards a fully automatic end-to-end solution for light commissioning, monitoring and light management. Furthermore, with the use of the light identification output we attain to select the image-set of single-light-source-lit images in order to extract the albedo map of the visible scene surfaces.

Chapter 4

Camera-aided Light Modeling and Estimation

Contents

4.1 Introduction

Most lighting systems around us are the result of a careful design by lighting professionals and architects. This is the case of offices, whereby the level of light on each desk is regulated by ISO standards and of industrial sites, whereby lighting means safety and well-being for the human operators. In recent years, most sectors have also changed to more agile strategies, *e.g.* reconfiguring the position of office desks in due course according to the need, which is not reflected any more in the previously designed lighting system.

The industrial-driven process of designing and creating lighting systems relies heavily on computer graphic based simulations. These approaches have been given an increased scientific attention [LHHC+13; SW14; SOLS+16] with modern systems where the light design process can be carried out in a completely interactive way using novel methods from the field of visual computing [HILI16; KLSW17]. Besides such improvements, this process requires a strong manual intervention, using tools that are essentially custom-build computer graphics and computer-aided design (CAD) software [DIAL94; LIGH90; RELU10]. All methods provide dense light intensity measurements through simulation, but they are based on simplified CAD models of a scene and they require to manually assign the reflectance properties of each structure and object present in the environment.

On the other hand, the only alternative for measuring light is hardware-based and it relies on the use of luxmeters, a device that normally provide a point-to-point measurements (*i.e.* sparse) of light intensity. This means that the operator needs to repeat the measurements in

several positions of the environment, swaying for angular and spatial completeness and as a consequence being a very time and cost-expensive procedure.

Thus in this chapter, we propose an efficient solution to light measurement which uses commodity hardware (an RGBD sensor) and it provides dense pixel-level measurements of the light intensity in real environments. We leverage best features from CAD-based simulation approaches, *i.e.* to achieve a dense lighting measurement, but we bypass the need of user input of the CAD and reflectance models by means of a RGBD camera observing the scene. This results in an automatic procedure, which reaches the accuracy required by lighting designers.
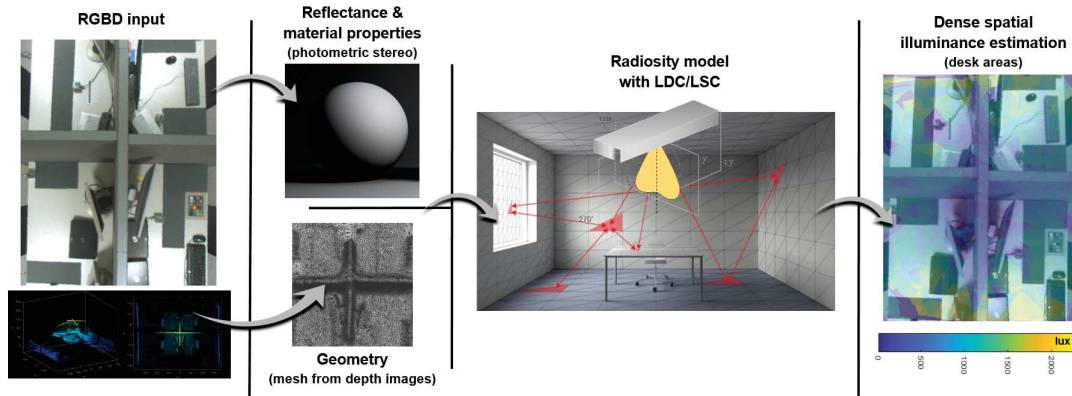
FIGURE 4.1: Pipeline of the RGBD2Lux system. We first acquire the RGBD input from the camera system (*left*), we then use the RGB images for extracting the albedo values of the surfaces based on a photometric stereo solution and the depth images for extracting the actual surfaces and the partial geometry of the scene (*center*). Both are then fed into our radiosity model using known light sources distribution curves. Lastly we measure the estimated illumination over the visible scene. See the estimated illumination over the desk areas (*right*).

Figure 4.1 shows a graphical description of our system (RGBD2lux). In general, the RGBD camera is installed so that to provide a top-view of the indoor environment[1]. Given the depth information, we extract a partial mesh in the field of view of the camera. At the same time, for each patch of the mesh we estimate the reflectance as a scalar value (albedo) from multiple images with different light sources activated. Finally, the 3D mesh and reflectance are then used in a radiosity model that provides us with the overall estimation of the scene illuminance. The proposed system is then evaluated against ground truth readings from a set of luxmeters showing its potential as a new system for dense light measurement.

4.2   Light modeling

4.2.1   Camera-aided 3D and reflectance modelling

Most simulated environments used for light intensity estimation need the *a priori* information of the CAD model and the material properties of the objects in the room. This is often hard to obtain in every lighting setup scenario and, when the CAD model is available, it is common to contain structural inaccuracies of the environment. To this end, we use a depth sensor for retrieving the coherent information of the indoor environment 3D structure without using a CAD model. The main goal is to reconstruct the surface from the point cloud, represent it as

---

[1]This is a practical solution since it provides the least occluded view of an indoor environment, however our method by design could work from any viewpoint.

a pair of vertices and faces (also known as patches) and associate at each face a scalar albedo value.

The main issues to deal with is that depth sensors provide a sparse and nosier point cloud from which it is difficult to obtain a complete CAD like representation. Thus, given an exemplar point cloud from the RGBD sensor, we first apply a denoising procedure to remove any outlier points, we then apply a mesh reconstruction solution based on Poisson surface reconstruction [KH13] approach and finally we post process the mesh with a Laplacian smoothing filter.

Using instead the images from the RGBD sensor, we record a time lapse video of the scene undergoing light variations given by the activation of the different light sources (*i.e.* luminaries) in the room. Then we select a subset of images with the LIT method as we described in the previous Chapter for which a single light source only is active. We use such images to estimate the pixel-wise albedo ρ using a first-order spherical harmonics model [BJK07] with known surface normals obtained from the previous surface reconstruction step. Then these values are mapped to the surface faces as the mean value of the of the albedo related to each pixel which are falling within the area of each triangle.

Given the partial 3D surface of the indoor environment and the scalar reflectance we can now apply the radiosity model in order to estimate the illuminance at each surface patch.

4.2.2 Light modeling - radiosity

The radiosity model [CWH93; MAIT15] is the current state-of-art model in all commercial lighting simulation software in the lighting field (Relux [RELU10], DIALux [DIAL94], AGi32 [LIGH90]).

As we've seen in Chapter 2, radiosity represents the light arriving at each point within a 3D scene by considering both direct light sources and inter-reflections. In this case given our 3D mesh discretized into a set of n triangular faces (or patches), we aim to compute the radiosity b_i at each patch $i = 1 \ldots n$. The scalar b_i (measured in $[W/m^2]$) is given by the direct emission of the room active sources (*e.g.* the room luminaires) plus the inter-reflection given by the other patches. We consider the information about the light source position in 3D and their luminous intensity as known since luminaries may hardly be moved after their installation. Thus, the patches corresponding to the light sources are assigned with an emitting intensity value equal to the luminous intensity of the luminaires.

Given the available information, and considering the radiosity formulation, *i.e.* Eq. 2.18, we can re-write the expression for radiosity with linear equations for the scene patches as:

$$b_i = e_i + \rho_i \sum_{j=1}^{n} F_{ij} b_j, \tag{4.1}$$

where i and j index the n patches, e_i is a scalar for the self-emittance of patch i ($e_i = 1$ for light source patches in the unit measure, 0 otherwise), ρ_i is the isotropic reflectivity of patch i and F_{ij} is the form factor between patch i and j.

Form factors encode two main aspects:

- **Visibility**, if two patches are visible one from each other, *i.e.* this value is equal to zero if there is no line of sight between them;

- **Distance and orientation**, encoding how well two patches "see each other", *i.e.* low values correspond to very far patches with an oblique line of sight, high values instead refers to close fronto-parallel patches.

They are further constrained to be strictly non-negative and satisfy the reciprocity relation $A_i F_{ij} = A_j F_{ji}$, where A is the area of each patch (note that the reciprocity relation is not symmetric unless the patches have the same size).

Eq. (4.1) can be expressed as a global model by stacking all the equations into a matrix (radiosity matrix) and thus obtaining a linear system $\mathsf{F}\,\mathbf{b} = \mathbf{e}$ such as:

$$
\underbrace{\begin{bmatrix}
1 - \rho_1 F_{11} & -\rho_1 F_{12} & \cdots & -\rho_1 F_{1n} \\
-\rho_2 F_{21} & 1 - \rho_1 F_{22} & \cdots & -\rho_2 F_{2n} \\
\vdots & \vdots & \ddots & \\
-\rho_n F_{n1} & -\rho_n F_{n2} & \cdots & 1 - \rho_1 F_{nn}
\end{bmatrix}}_{\mathsf{F}}
\underbrace{\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}}_{\mathbf{b}}
=
\underbrace{\begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix}}_{\mathbf{e}}
\tag{4.2}
$$

where F is a $n \times n$ square matrix describing the geometry of the whole room, the n-vector $\mathbf{b}$ contains the associated radiosities at each patch and the self-emission $\mathbf{e}$ contains non-zero values at patches corresponding to active light sources. Now, solving for $\mathbf{b}$ requires the knowledge of geometry F, object reflectance ρ_i and the luminous intensity of the light sources $\mathbf{e}$.

The calculation of the form factors in its simplest analytical form is formulated by integrating the differential areas of the two in relation surfaces i,j as shown in the following equation:

$$
F_{ij} = \frac{1}{A_i} \int_{Ai} \int_{Aj} \frac{cos\theta_i cos\theta_j}{\pi p_{ij}^2} Y_{ij} dA_j dA_i
\tag{4.3}
$$

This approximation is easy to be implemented and works well for simple scenes and as long as the elements i and j are well-separated from each other (the distance p between them is large relative to their sizes), however in more complex scenes this analytical solution tends to be time consuming and with less accurate output.

Therefore, in our case in order to compute the form factors we adapted a ray casting approach, where we uniformly sample rays within the unit disc (*i.e.* the orthogonal projection of the unit sphere), whereby each point on the unit disc defines the direction of a ray in the space. Thus, following [MALL88], we compute F_{ij} as the ratio:

$$
F_{ij} = \frac{\mu_j}{\mu_i},
\tag{4.4}
$$

where μ_j stands for the number of rays emitted by patch i that reaches patch j, and μ_i is the total number of rays emitted by facet i.

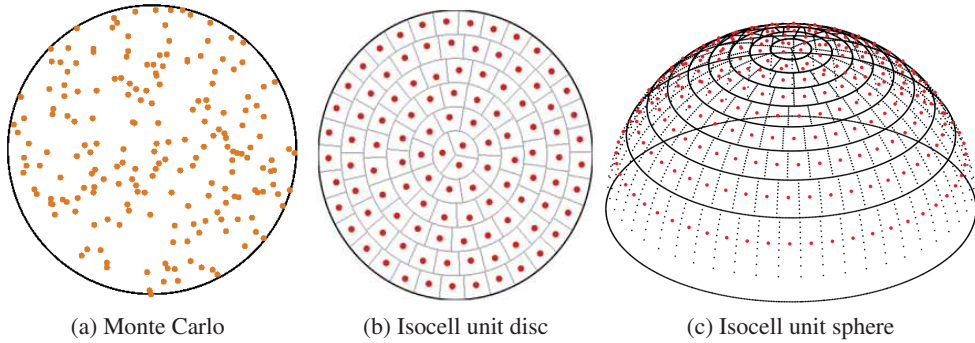| (a) Monte Carlo | (b) Isocell unit disc | (c) Isocell unit sphere |

FIGURE 4.2: Illustration of the two different ray distribution methods.

As shown in Fig. 4.2, there are different methods to distribute the rays *e.g.* the Monte Carlo method [CPC84; KAJI86], *i.e.* the ray orientations are sampled randomly (see Fig. 4.2a). This method always converges but requires a great number of rays to achieve a good precision, thus being computationally demanding. Instead we sample points in the unit disc (and therefore ray directions) by the approach of Masset *et al.* [BB16; MBK11], noted as "Isocell", due to its higher precision within a fixed computation time. This is based on a uniform discretization of the ray orientations (*c.f.* Fig. 4.2b). In our implementation, given our patch discretization and room size, we found that $1,000$ rays for each patch is a good compromise between accuracy and speed (*c.f.* Fig. 4.3c illustrates Isocell unit sphere rays, for a single patch on the office floor, within the CAD model). Thus, for computing the matrix F_{ij} we iterate through each patch and project the rays into the space from its center point and then we assign the corresponding form factor value according to the ratio described by Eq. (4.4).

Finally once we have populated the view factors matrix f_{ij} we need to ensure the reciprocity relation mentioned earlier. Here we adopt the iterative scheme of Van Leersum *et al.* [VAN89] that can be considered as a refinement of the naive form factors rectification.

4.2.3 Customizing radiosity for real environments

The radiosity model in Eq. (4.2) has two limitations: it assumes point light sources and it disregards the light perception. We address both aspects in our model extension, by introducing the LDC (for considering any light source, see Sec. 2.3) and the LSC (to model the light observer / sensor perception, see 2.2.2).
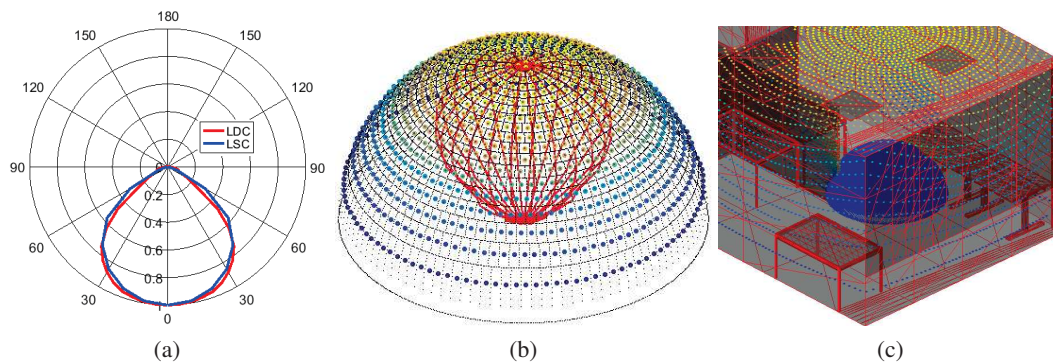

FIGURE 4.3: (*a*) Radial representation of LDC and LSC curves, illustrating how both quantities are attenuated (radius) with the growing light incident angle (radial angle); (*b*) weighted ray distribution of the Isocell unit sphere for LDC, heatmap-color-coded; (*c*) Isocell unit sphere (blue hemisphere) for a patch on the floor of the office CAD. The color-dots in (*c*) are taken from (*b*) but placed at the intersection of the cast ray with the other scene patches.

Light distribution curve (LDC). Let us recall Equations (4.1) and (4.2). The radiosity model uses a scalar variable e_i to set whether each patch is a light source ($e_i > 0$) or not ($e_i = 0$). The actual e_i value would relate to the radiant intensity or luminous flux of the light source, measured in lumens $[cd/sr]$. Since it uses scalar values, the radiosity formulation of Section 4.2.2 assumes isotropic light sources, *i.e.* active patches radiate with the same intensity in all directions.

But isotropy is hardly the case for real light sources which are normally represented by a radiation map, such as in Figure 4.3a (*red*). Their distribution is in general non-linear and

encoded in a light distribution curve which provides the radiant intensity with respect to distance and angle from the emitting source (*i.e.* how much light is emitted in each direction, as shown explicitly in Sec. 2.3). The LDC curves are obtained experimentally in lab environment conditions with the use of specific equipment, *i.e.* goniophotometer, and provided in datasheets by the luminaire manufacturer.

Thus, similarly to the existing solutions we include the LDC curve into the ray tracing procedure of our application and therefore we encode the non-linear radiation of light sources into the form factors F_{ij}. In more detail, we associate scalar values to each light source patch cast ray, which are proportional to the angle of emission, as illustrated in Figure 4.3b. These are then used to re-write Eq. (4.4) as a weighted mean. Since this procedure encodes the non-linear LDC into the form factors, our re-formulation of radiosity preserves the linearity of Eq. (4.2).

Note that, in this way, the non-linearity is embedded in the form factor computation process, which is already non-linear. This results therefore in a minimal impact on computation, since the system in Eq. (4.2) remains linear.

Light sensitivity curve (LSC).   People perceive light differently, depending on their orientation with respect to it and depending on the distance from it. In a similar fashion, luxmeter sensors have different sensitivity to lighting, depending on the lighting angle and distance as well as to manufacturing characteristics (see Sec. 2.2.2). The LSC plot in Figure 4.3a (*blue*) illustrates the perception characteristic of the luxmeter which we adopt in order to meet the measuring requirements of the collected ground truth data. Note the strong similarities to LDC as in Figure 4.3a (*red*).

Therefore, we correct the sensor's light perception by the LSC. Similarly to the LDC, we formulate the corresponding weighted Isocell unit sphere, as in Figure 4.3b, and integrate the weights into the form factor calculation by modifying the ray casting. In this way, we maintain the radiosity linear formulation and alter the computation time minimally. It is worth mentioning here that interestingly enough this functionality is not applicable in any commercial light planning software available in the market.

4.2.4   Method resume

Algorithm 2 resumes the steps of the RGBD2lux approach necessary that estimates the light intensity using the input from an RGBD sensor.

Algorithm 2: RGBD2lux method.

---
1 From depth data compute a 3D mesh, see Sec. 4.2.1 ;

2 From a time-lapse video sequence, identify a subset of images with single light activations and run photometric stereo to find the albedo values ρ, see Sec. 4.2.1 ;

3 Compute the form-factors F_{ij} given the 3D mesh model, see Sec. 4.2.2 ;

4 For the 3D patches related to light sources apply the LDC computation of form factors, see Sec. 4.2.3 ;

5 For the 3D patches where perceived light is measured, apply the LSC computation of form factors, see Sec. 4.2.3 ;

6 Solve for the radiosity matrix, Eq. 4.2, see Sec. 4.2.2 ;

Once the radiosity values are computed by the RGBD2lux approach, they can be converted to illuminance considering the Lambertian assumption [LA70] and the fact that reflections are scattered isotropically and thus the illumination intensity is not a function of the direction of the ray/beam:

$$L_{(x,y,\theta,\phi)} = L_{(x,y)} \neq f(\theta, \phi), \tag{4.5}$$

but instead will be constant for different sets of measuring equipment and directly proportional to the exitance intensity $E_{(x,y)}$ with a conversion factor of κ:

$$E_{(x,y)} = L_{(x,y)} \cdot \kappa = L_{(x,y)} \cdot \frac{\pi}{\rho} \tag{4.6}$$

where $E_{(x,y)}$ is the illuminance measured in *lux* and $L_{(x,y)}$ the radiance. This equation is directly related with the equation 2.10 that we have seen in Section 2.4.1.

4.3 Illumination evaluation

4.3.1 Light measurement benchmark

To the best of our knowledge, there was no dataset for benchmarking light measurements with ground truth in real scenes. Thus for the purposes of this study, we define a new benchmark for light measurement in real world scenes. We selected two different office rooms, namely room_1 and room_2 (see Figure 4.4). As shown, we provide a detailed CAD design for the rooms representing the 3D ground truth. This includes an accurate labelling of the object textures and reflectivities for each surface facet in the CAD.

Both rooms have been equipped with a controlled lighting environment, since the position of eight luminaries is known (Fig 4.4b), as well as their type (Siteco Mira), meaning that the LDC curve is given by the manufacturer. Furthermore, to achieve even more accurate ground truth, we measured in lab conditions the overall luminous flux and temperature of each luminaire being $7913lm$ and $4250k$ respectively, while the multiplicative parameter for the luminaire age[2] is assigned to value 1 since the installation is fairly new.
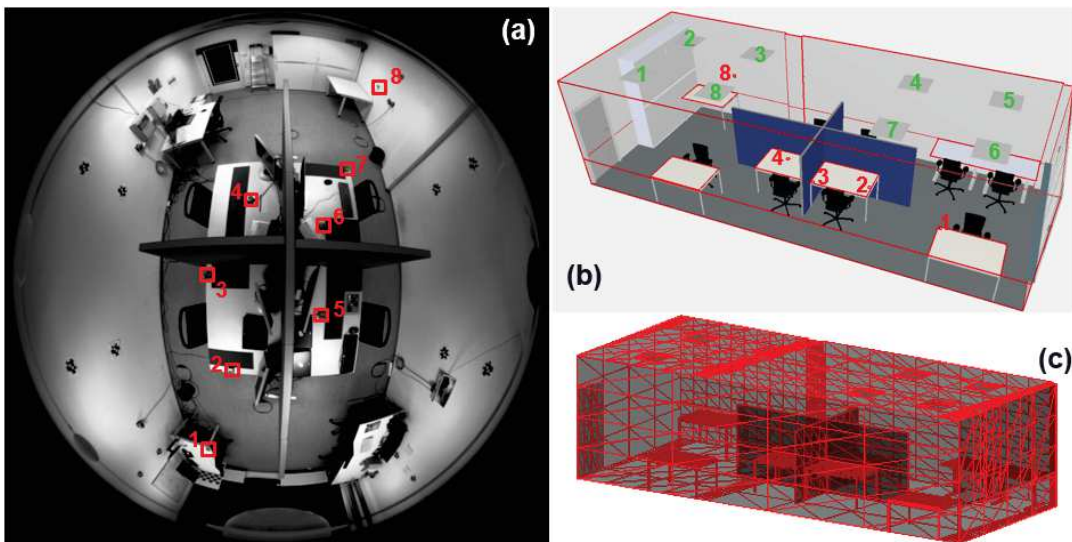

FIGURE 4.4: Room_1 full scene. *(a)* illustrate image as it looks from the camera, red bounding boxes are showing the location of luxmeters within the indoor space. Images *(b)* and *(c)* are detailed CAD design of the room showing the luminaire positioning and how the scene is subdivided into individual patches respectively.

---

[2]The efficiency of every luminaire degrades over time.

We provided for both rooms a number of sensory data. First, we set up in the ceiling of the rooms a RGBD calibrated and aligned sensory system which consists of an rgb camera with fish-eye lens with 180° FOV and a depth camera (Bluetechnix time-of-flight) with 90° FOV. However, in our experiments we considered only the part that is visible to both images after we have registered and undistorted them (as shown in Fig. 4.1, *left*).

The sensors are synchronized with luxmeters (also indicated in Fig. 4.4a and 4.4b), providing therefore the ground truth data for illumination intensity. The luxmeters however give localized (*i.e.* point-to-point) *lux* readings only, so we installed 8 of them in different areas so providing a reasonable sampling of the environment. We chose mainly locations over desk areas, since they are of major importance in the lighting field. For each luxmeter, we additionally report the type and their specific light sensitivity characteristic curve (LSC, see Fig. 4.3a), namely the sensor sensitivity across the incident light angles.

Thereafter, we evaluated 31 different luminaire activations (luminaires switched on or off) for each room, see Fig. 4.5 for a sample of images obtained from room_1. We target the use of rgb and depth input just for light measurement, the use of luxmeters as ground truth, and all other provided information for ablation studies.
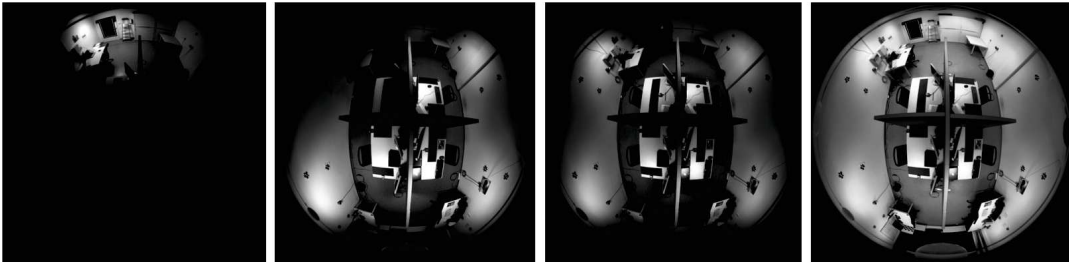

FIGURE 4.5: Illustration of 4 illumination variants within room_1. From the left to the right, the images illustrate the illumination provided by 1, 3, 4 and all 8 luminaires switched on in the scene.

4.3.2 Quantitative comparisons in *lux*

Table 4.1 summarizes the overall performance of the commercial light modeling software Relux, compared to ours and both to the ground truth under all the different experimental inputs. Beyond the proposed RGBD2lux, which does not require the CAD model and integrates the LDC and LSC curves, we additionally compare Relux to our own model using the CAD input provided (*i.e.* **Ours with CAD**). Results are reported for both room_1 and room_2.

Note that our radiosity model using the knowledge of the CAD model outperforms Relux, thanks to the induced LDC and LSC distribution curves. In fact, against ground truth, **Ours CAD (LDC_LSC)** achieves an average luxmeter error (across all 1-8 luxmeters) of 36 *lux* for room_1 and 70 *lux* for room_2[3]. These errors are 9.2% and 3.7% better than Relux, for room_1 and room_2 respectively.

Our proposed RGBD2lux (modelling LDC and LSC but not requiring the CAD) is still comparable with Relux (instead requiring the CAD). Here we only consider the luxmeters in the area visible from the depth camera (**Avg (2-7)**), *i.e.* those for which RGBD2lux may compute a *lux* value. For room_1 we achieve an error of 61 *lux*, and for room_2 we achieve an error of 99 *lux*. Thus, in this case the average *lux* error percentage is respectively worse than Relux by 4.6% for room_1, and by 8.1% for room_2.

---

[3]The material in Appendix B describes which are the structural differences between the two rooms together with the reason why room_2 is more complex in terms of light modelling than room_1

It is of interest to note that the inclusion of LSC provides a larger error reduction than LDC (rows 6 and 7 in the table), consistently across the two rooms. We interpret this as the way the form factors matrix is populated, *i.e.* based on the ratio between the rays arriving at a patch by the overall emitted rays, where there is a higher dependency on the emitted rays rather than the received rays, which would otherwise be equally strong from any angle.

	Room 1 Error (in *lux*)										Room 2 Error (in *lux*)									
	Luxmeters								Avg. (1-8)	Avg. (2-7)	Luxmeters								Avg. (1-8)	Avg. (2-7)
	1	2	3	4	5	6	7	8			1	2	3	4	5	6	7	8		
Relux	167	96	27	26	43	10	96	39	63 (21.4%)	50 (20.7%)	206	97	27	80	97	49	73	44	84 (22.2%)	71 (20.4%)
Ours with CAD (no_LDC_LSC)	188	150	33	45	43	34	91	65	81 (27.5%)	66 (27.3%)	207	114	99	148	105	117	93	81	120 (31.8%)	112 (32.2%)
Ours with CAD (LDC)	199	152	29	41	40	33	95	57	81 (27.5%)	65 (26.9%)	213	117	82	125	97	97	86	63	110 (29.1%)	100 (28.8%)
Ours with CAD (LSC)	73	45	24	32	40	34	46	52	43 (14.6%)	37 (15.3%)	69	80	98	136	70	84	56	62	82 (21.7%)	87 (25.0%)
Ours with CAD (LDC_LSC)	69	24	22	38	28	28	38	41	**36 (12.2%)**	**30 (12.4%)**	70	57	76	106	75	69	55	53	**70 (18.5%)**	73 (21.0%)
Ours with CAD Camera visible (LDC_LSC)	-	64	28	20	17	22	52	-	-	34 (14.1%)	-	54	36	59	101	69	54	-	-	**62 (17.8%)**
Ours RGB2Lux (LDC_LSC)	-	53	41	67	68	40	98	-	-	61 (25.3%)	-	98	90	85	136	108	77	-	-	99 (28.5%)

TABLE 4.1: Illumination estimation errors by the considered approaches, applied to both rooms. Values for individual luxmeters, cols 1-8, correspond to the average *lux* values estimated over the results of the 31 different lighting combinations. Avg. (1-8) corresponds to the total average values for all installed luxmeters. By contrast, Avg. (2-7) only considers those luxmeters which are visible from the RGBD camera, *i.e.* within its field-of-view. The percentage values correspond to the *lux* average percentage error in regards to the ground truth.

4.3.3 Comparisons against Relux using CAD models

In this section, we analyze in more details our proposed system against the commercial Relux software [RELU10]. For comparison, since Relux requires the CAD model, we also provided this information to our method. These experiments are a sanity check for our proposed approach (since Relux also uses radiosity) and we provide a detailed analysis on the contribution of LDC/LSC curves and the effect of considering the visible scene only.

Figure 4.6 shows six box plots of the light measurement error (y-axis), as measured by each of the 8 installed luxmeter sensors (x-axis). In each column of the box plots, the 31 gray dots represent the measured error of each of the lighting scenarios. The pink box represents the central 50% of the data, while the upper and lower vertical lines indicate the extension of the remaining error points outside it. The central red line indicates the mean error.

In Figure 4.6, we note a similar performance of Relux (leftmost box plots) and ours without LDC and LSC corrections, for both rooms (top and bottom rows), in line with the findings of Table 4.1. Interestingly, highest errors occur at luxmeters 1, 2, and 7 which are closely localized under the luminaries and "looking" directly at the light source (check Figure 4.4 for the sensors geometrical distribution). At these positions, the larger errors are due to a lower illuminance estimation by the radiosity model[4] as it is also exemplified in Fig. 4.9.

The rightmost plots in Figure 4.6 clearly show that the _LDC_LSC corrections are highly beneficial. The error reduces across all luxmeters and most prominently for those luxmeters 1, 2 and 7, addressing the orthogonally-incident light. Of particular interest is the case of luxmeter 4 for room 2, whereby the error remains larger. Figure 4.7 shows the region where the luxmeter is located and this reveals the cause of the issue. The screen monitor shields (occludes) luxmeter 4 from the light of some of the luminaires. Since the simplified CAD model does not include the monitor, the radiosity cannot model the occlusion properly.

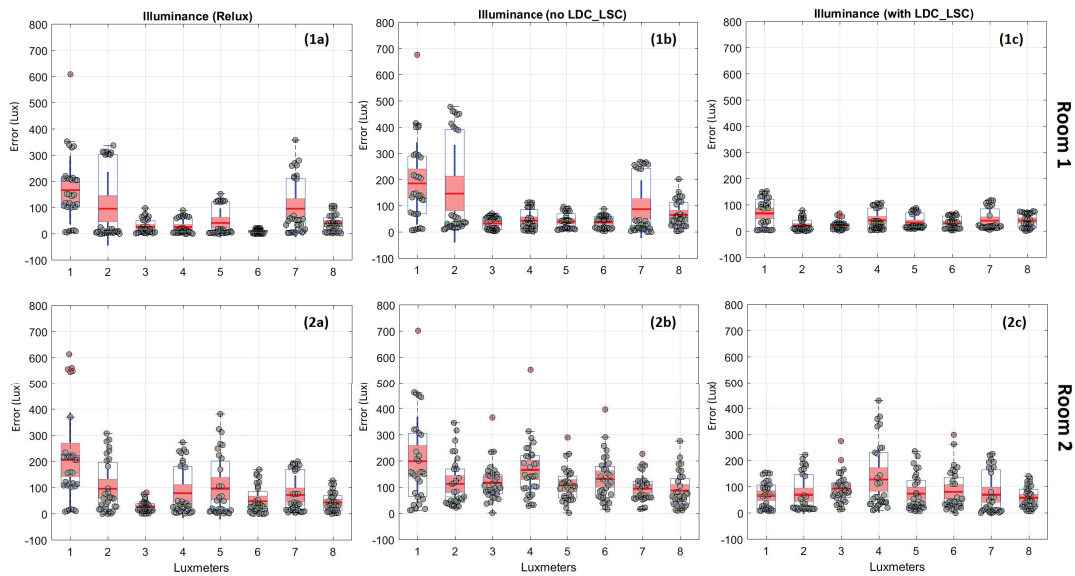[4]See the signed plots in the material in Appendix B.

FIGURE 4.6: Room_1 & 2 *lux* boxplot error evaluation using the full CAD information against *(a)* Relux software, *(b)* standard radiosity model and *(c)* radiosity with _LDC_LSC applied.
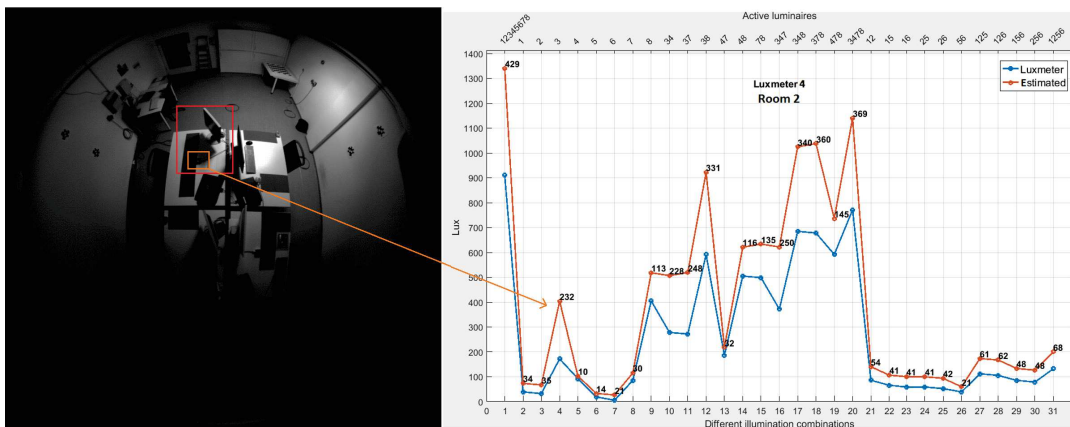

FIGURE 4.7: A detailed description of the problem for luxmeter 4 in room 2. There are higher errors due to shadows on the luxmeter sensor and inaccuracy in the 3D CAD model. The graph plot next shows the correlation of the sensor's measurement and which light activation the error is a result of.

Before we move to the depth based simulations we established another experiment where we simulated a partial part of the CAD model. Here we simulate that the only CAD information available is the one actually visible by the camera. Such information might be available for instance to a depth camera observing the scene. In such case almost 60% of the room is not visible so providing a real challenge for our light estimation system. Figure 4.8 shows the results for our radiosity with LDC_LSC model since i) it is the best performing against standard radiosity, ii) Relux software does not work with open surfaces (while we do).

Note the increment in the error for luxmeters 1, 2 and 8, especially in room 1, in comparison to Figure 4.6 *(1c)*, which are closer to the walls. Since part of the walls are not visible (because they are not included in the camera field of view) the form factors fail to grasp the light contribution of the wall reflections (see Figure 4.8).
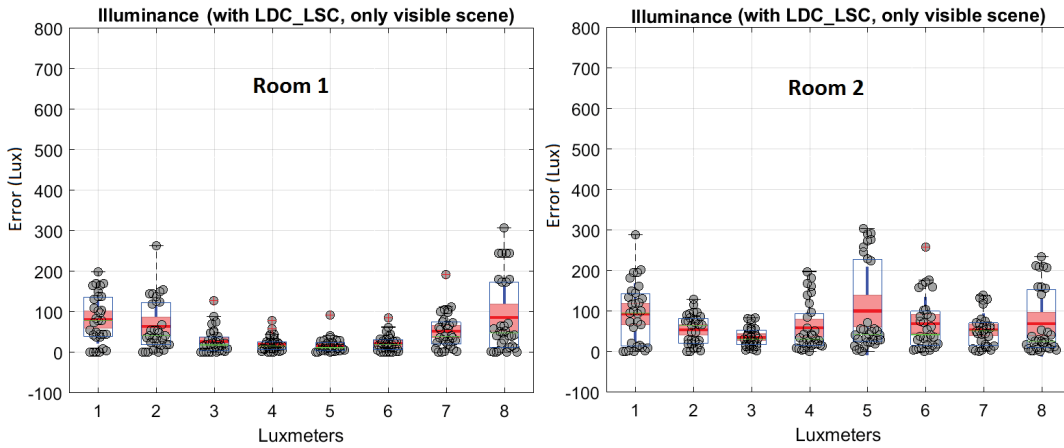
FIGURE 4.8: Room1 & 2 *lux* boxplot error evaluation using only the visible CAD information on the radiosity with _LDC_LSC.

4.3.4 Light measurements from RGBD data

Here we evaluate the performance of our proposed RGBD2lux. We use just the RGBD image and apply the correcting _LDC and _LSC curves to the radiosity model.
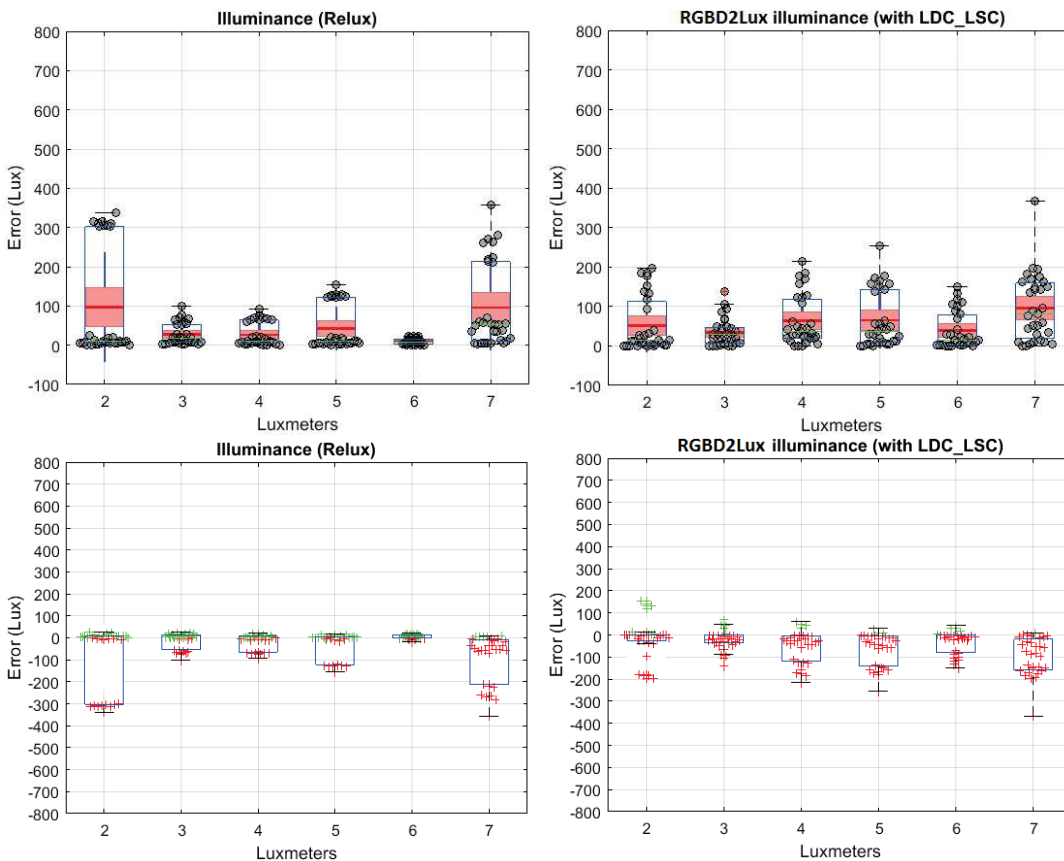

FIGURE 4.9: The figure shows the comparison between the Relux simulation with CAD model *(left)* and our RGBD2lux approach *(right)* for room_1. Even if our method do not use the CAD model, RGBD2lux achieves better or equivalent performance in some scenarios. Plots in the second row specify the type of the error, *i.e.* due to lower *(red marker)* or over *(green marker)* estimation.

The box plots in Figure 4.9 show that RBGD2lux achieves better or equivalent performance than the Relux simulation in some of the 31 lighting scenarios tested. Note that the results using the RGBD input have an error mainly due to lower estimation of the illumination levels within the room as a cause of the missing geometry. However, considering the incomplete geometry from the depth sensor, this experiment shows that our system can match or even overcome the CAD-based (with full geometry given) state of the art in challenging scenarios (similar results are obtained for room_2 and included in the supplemental material). To conclude, the computational time for our modelling is approx. 3-5 minutes (Matlab) while Relux software simulations requires approx. 15-20 minutes for the same scenes, meaning that the computational complexity of our method allows real-time performance

4.4   Conclusion

In this Chapter we presented the RGBD2lux pipeline and how we modified the radiosity model so that it could be used as a viable option for estimating light intensity in real environments. Interestingly, we showed that it is possible to challenge or even surpass current state-of-the-art light planning and modeling simulation software where the complete CAD model geometry is a necessary requirement. Moreover, we have shown that the RGBD sensor, even if providing partial geometry information, could be a more realistic expression of the 3D structure of the current scene. Even the most accurate CAD model might not be aware about changes on the room structure that happens after the original planning.

The output of this study if attached to dimmable light sources, these modules may represent a major improvement for smart lighting, where light level is continuously updated to ensure maximum comfort, well-being and power efficiency as we will show in the next Chapter 5.

The proposed lighting model can surely be improved and further experiments can be conducted, taking into account other material (*e.g.* specular components, caustics, *etc*.) and other scene (*e.g.* natural light) properties. Moreover, combined with light calibration approach it could drive towards fully automatic solution where the positioning and the light intensity of the light sources will also be estimated on the fly eliminating any need for on demand information. Moreover, considering that the major cause of the illumination estimation error is due to the fact of the limited geometry of the scene, the proposed system could be optimized by obtaining a boarder view of the scene with the use of and the combination of a fish-eye camera and a depth sensor in a similar way proposed in [PLG16; FPLG18; FFPD+18] and in that way addressing this problem.

In the next Chapter 5 we introduce the first method that makes use of the spatial illumination estimation in conjunction with the human occupancy and visual attention of the occupants within an indoor environment. In particular, by means of a state-of-the-art variant of a Mask R-CNN model [HGDG17] we integrate both aspects in order to estimate how much light each person perceives, depending on their position and gaze.

Chapter 5

Human-centric Light Sensing and Estimation

Contents

5.1 Introduction

Lighting design in indoor environments is of primary importance for at least two reasons: 1) people should perceive an adequate light; 2) an effective lighting design means consistent energy saving. Thus, we presented the *"Invisible Light Switch"* (ILS) to address both these aspects. The main idea of *ILS* was presented in Chapter 1.2, in practice *ILS* tries to dynamically adjust the room illumination level to save energy while maintaining constant the light level perception of the users. So the energy saving is invisible to them.

Our proposed *ILS* leverages a radiosity model to estimate the light level which is perceived by a person within an indoor environment, taking into account the person position and her/his viewing frustum (head pose). *ILS* may therefore dim those luminaires, which are not seen by the user, resulting in an effective energy saving, especially in large open offices (where light may otherwise be ON everywhere for a single person). To quantify the system performance, we have collected and evaluated a new dataset where people wear portable luxmeter devices while working in office rooms. The luxmeters measure the amount of light (in *lux*) reaching the people gaze, which we consider a proxy to their illumination level perception. The initial results are promising: since we show that in a room with 8 LED luminaires, the energy consumption in a day may be reduced from 446.05 to 148.95[1] *kWh* with *ILS*. While doing so, the drop in perceived lighting decreases by just 200 *lux*, a value considered negligible when the original illumination level is above 1200 *lux*, as is normally the case in offices.

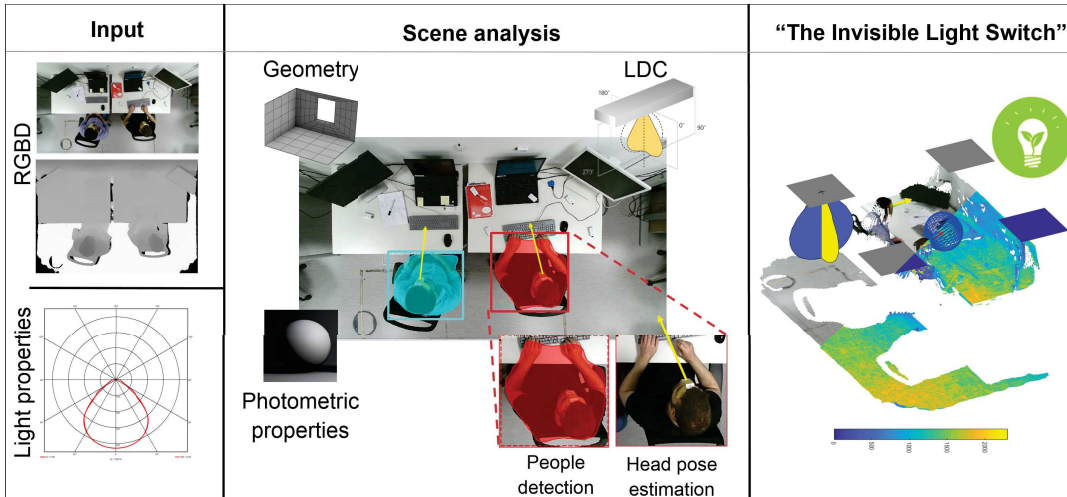[1]including the operational consumption of the *ILS* system

FIGURE 5.1: Overall pipeline of our system. We first acquire the RGBD input from the camera system *(left)* and together with the lighting system properties we use this information to create the *"Invisible Light Switch"* (ILS). That is, reconstructing the 3D geometry of the scene, extracting the photometric properties of the material and applying a human centric analysis from where we detect the human presence in the scene and extract the possible head poses. Lastly we utilize the output of the scene analysis as the "Invisible Light Switch" application targeting a power saving framework.

5.2 Ego light perception

The *ILS* is presented, as a smart lighting framework for dynamically adjusting the illumination level in an indoor environment. *ILS* takes into account the geometry of the scene, the presence of people and their light perception with the goals of maximizing the human comfort in terms of perceived light and, at the same time, with the lowest cost in terms of energy consumption. We do this by bringing together individual works into a unique pipeline as we show in Figure 5.1. The framework builds upon the light estimation system presented in Chapter 4, which is capable of estimating the light in a given 3D point of a multi luminaire indoor environment. As we have shown, the presented radiosity model has been customized to take into account a realistic model of light propagation, outclassing even industrial software in the task.

We further enriched that model by including the human aspect, and showing how the interplay between the light estimation system and the human activity may lead to a consistent energy saving framework. The invisible light switch summarises the idea: an individual has the feeling of an environment which is globally illuminated, while in reality an automated light switch dims the luminaires in a way which is invisible to the users. This was possible by estimating the position of a person in the sensed environment, its head orientation, and understanding the light which is perceived by him. In fact, the lighting sensed by a human can be assumed as the light contained in a conic volume departing from the mean point connecting the human's eyes in the direction of the nose. Given this, it is possible to determine which luminaries could be switched off/dimmed down while maintaining the level of perceived light unchanged. The head pose is provided by detecting the person first and then estimating the head orientation. The former is carried out by means the state-of-the-art detector Mask R-CNN [HGDG17] with ResNet [HZRS16] as a backbone architecture, while head pose estimation is done using Hasan's *et al*. method [HTGD$^+$17].

5.2.1 People detection and head-pose estimation

We aim at detecting people and estimating their head pose (their viewing angle). For the first task as we mentioned we adapted the Mask R-CNN [HGDG17] object detector, while for the second one the head pose estimator proposed in [HTGD+17].

The R-CNN [HGDG17] detector has the ResNet-101 [HZRS16] as a backbone architecture, trained on 80k images and 35k subset of evaluation images (trainval35k) of MS COCO dataset [LMBH+14]. We fine-tuned the detector on our top-view dataset (see Sec. 5.3.1), adopting a specific training portion of the data. We randomly partitioned the data into training and testing set, keeping 70% of the data for training and 30% for testing. Since the top-view images are different from the frontal-view images of the COCO dataset [LMBH+14], the fine-tuning had a crucial role. We adopted a similar procedure for training the head pose estimator as in [HTGD+17]. It is worth noting that the input for the head pose is the whole body detection bounding box: this is because [HTGD+17] has been specifically designed for managing small-sized head patches, exploiting the body as contextual cue for a better final head orientation classification. In particular, 4 and 8 classes related to angles have been taken into account.

During testing time, a cascaded approach is followed, first by applying the people detector and then feeding the detected body bounding box as input into the head orientation module.

5.2.2 Spatial light estimation

To obtain an estimate of a dense spatial illumination map, we adapted our pipeline presentedin Chapter 4. As we have presented there we make use of a radiosity model [CWH93] for estimating the spatial illumination over time by just using the input from an RGBD camera. Furthermore, we extract the information regarding the photometric properties of the material of the scene based on a photometric stereo baseline approach that is applied on the time-varying RGB images. This approach allows us to extract a scalar albedo at each pixel by using a set of images with different light sources that are switched on/off during the day. Having the light sources position and intensity, the scalar albedo under Lambertian assumptions, and the depth map from the sensor, our proposed method in Chapter 4 showed that it is possible to obtain a dense measurement of the light emitted by a 3D patch in the indoor environment. In order to provide more realistic estimates, we have shown how to model real lighting systems that, differently from point-like sources, emit light given a specific light distribution curve (LDC). The LDC is custom for each lighting system and their properties are considered to be known when estimating the light instensity. The proposed method shows that, even by accounting the non-linearities of LDC, it is possible to solve for the radiosity equation linearly and so obtain a more reliable measure of the light intensity, which we evaluated by using point-to-point sensory equipment *aka*. luxmeters installed across the scene.

5.2.3 Gaze-dependent light modelling

Light measurements are practically made using a luxmeter sensor. This sensor measures the perceived light that is in function of the distance to the light, the orientation and other manufacturing characteristics. These properties are resumed by the Luxmeter Sensitivity Curve (LSC) as in Figure 5.2a. The LSC illustrates the perception characteristic of every luxmeter sensor which in this work we adopt in order to meet the measuring requirements of the collected ground truth data and to simulate the human light perception. We have chosen this solution because this is the standard de facto in the lighting industry and it provides satisfactory solutions when doing light commissioning [NL11].
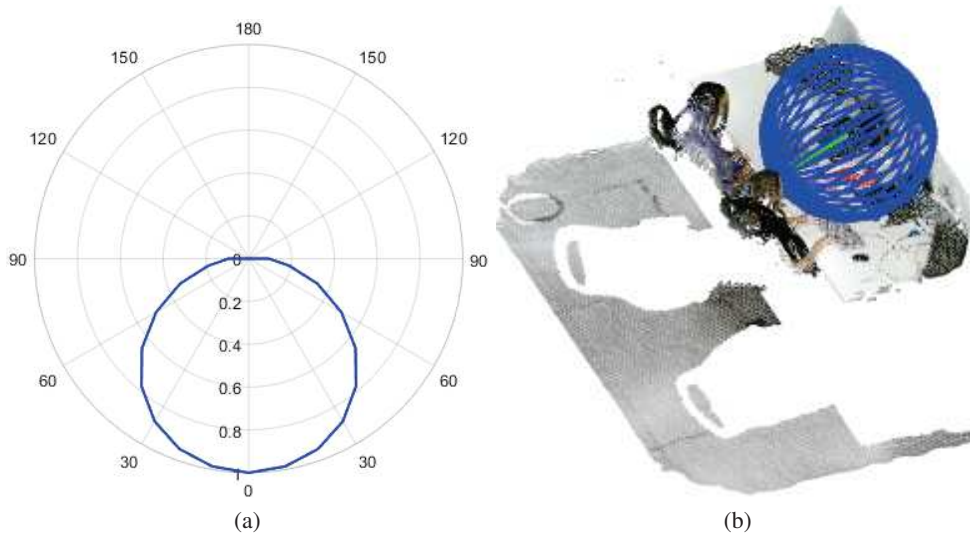
FIGURE 5.2: Modeling of the Luxmeter Sensitivity Curve (LSC) as a human light perception model.

The key idea in this procedure is that, once we have detected a person in the image and estimated his head positioning and orientation as described in Sec. 5.2.1, we extract his posture in the 3D space by mapping the 2D image coordinates of his detected head to the corresponding depth information. Thereafter, once we have the positioning of the head in the 3D space as well as its orientation (where the person looks at), we estimate the light that arrives to his/her face (or to the luxmeter as in our case) by applying a ray-casting procedure where we simulate the human field of view (FOV). Such view frustum is obtained by using emitted rays starting from the estimated head position towards the corresponding estimated head orientation. The total illumination arriving to the person is computed by adding the related spatial illumination (radiance) from the patches of the scene that are in the direct visibility of the person. The rays project in the space as a uniform generated sequence over the unit sphere and weighted accordingly, based on the modelled luxmeter's LSC, towards the visible patches from the FOV of the sensor. The contribution of each patch to the total amount of lighting perceived by the occupant, is computed by estimating the percentage of rays intersecting that patch.

5.3 Invisible light switch evaluation

5.3.1 Dataset overview

In the previous Chapter we introduced a dataset for benchmarking light measurements with ground truth sensory data in real scenes. In this study we extended this dataset by introducing two more scenes with human activity, one based on a normal office environment and a second one representing a relaxing area (see Figure 5.3).

Both new scenes comprehend different human activities *e.g.* watching TV, working on a desk area, chatting, *etc.*, as well as different head orientations (VFOA) and multiple light combinations. In this work, VFOA is a cone with vertex in the middle of a person's eyes, oriented as the gaze direction and an aperture angle of $\alpha = 30°$.

In both rooms there is a controlled light management installation, where the position, type and properties (*e.g.* luminous intensity, light distribution curve, *etc.*) of the luminaires (eight in total) are considered known, see Figure 5.4.
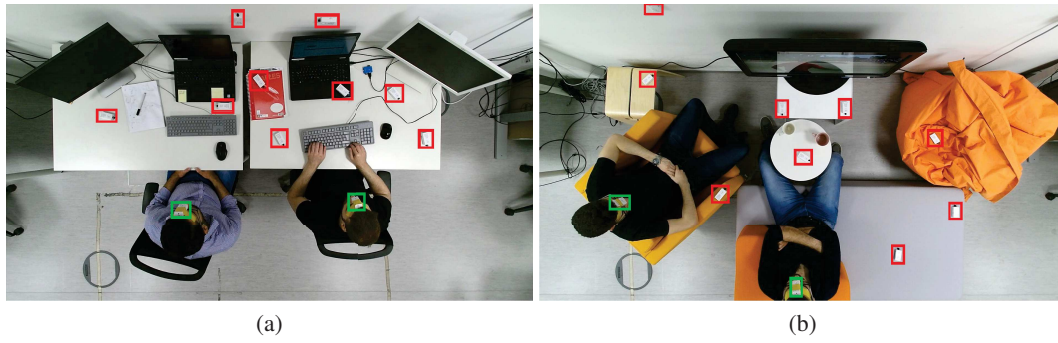
FIGURE 5.3: Illustration of the two indoor scenes used for evaluation: (a) illustrates a normal office environment and (b) shows a relaxing area. Red and green bounding boxes are showing the location of luxmeters within the space covering the spatial and gaze-gathered illumination ground truth measurements respectively.
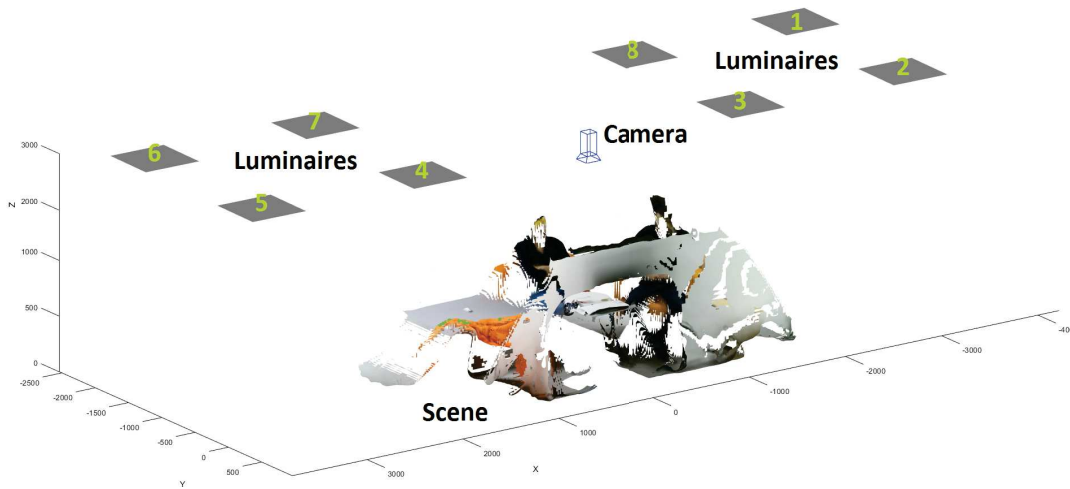

FIGURE 5.4: Illustration of the light management installation.

For obtaining the ground truth data similarly to the previous dataset (Sec. 4.3.1) we have installed and used a number of sensory equipment. A calibrated and aligned RGBD camera system (Kinect v2) is installed in the ceiling of the room providing a top-view perspective of the scene, see Fig. 5.3 and 5.4. Moreover, the camera is synchronized with a number of luxmeters (also indicated in Fig. 5.3) providing the light intensity ground truth data both for the spatial as well as for the gaze-gathered (attached to the forehead of the occupants) illumination. Considering the limitation (*i.e.* point-to-point) of *lux* readings that the luxmeters provide, we installed 11 sensors in different areas, thus providing a reasonable sampling of the scene. We use 9 luxemetes for evaluating the spatial illumination across the environment and 2 luxmeters for measuring the light intenisty that arrives to each one of the occupants appearing in the scenes. For each luxmeter, we additionally report the type and their specific light sensitivity characteristic curve, LSC (see Fig. 5.2) giving the sensor's sensitivity across the incident light angles.

Thereafter, we evaluate 24 and 30 different scenarios with different luminaire activations (luminaires switched on/off) for each room respectively (see Fig. 5.5). Our target was the use of RGB and depth input just for light measurement, the use of luxmeters as ground truth, and all other provided information for evaluation studies.
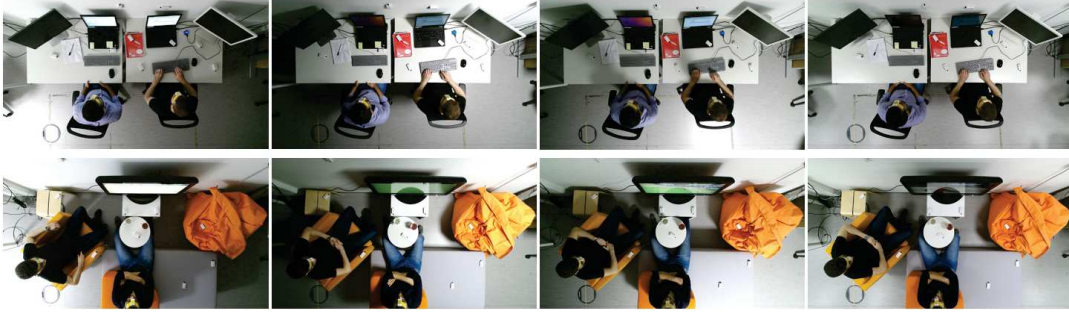
FIGURE 5.5: Illustration of 4 illumination variants within the two rooms. From left to right, the images illustrate the illumination provided by 1, 4, 7 and all 8 luminaires switched on in the two scenes.

5.3.2 Top-view detection and head-pose estimation

We fine tuned both the person detector and the head pose estimator on our top-view dataset. We report an average precision (AP) of 98% in terms of people detection. As mentioned previously we test our approach on the testing set of our top-view dataset. For the head pose orientation fine tuning on the whole body has been crucial for the performance, since using the sole head region produced definitely worst scores. In particular, we adopted two different class numbers for head pose, namely 4 and 8. The corresponding confusion matrices are reported in Fig. 5.7, showing an accuracy of 43.2% (8 classes) and 70.7% (4 classes) respectively. The scarce performance in the 8-class case was due to the mix among adjacent viewing angles: actually, the average size of the head region in the dataset is approx. 40x50 pixels. For these reasons, we used the 4-class version in the light perception studies.

(a)   (b)
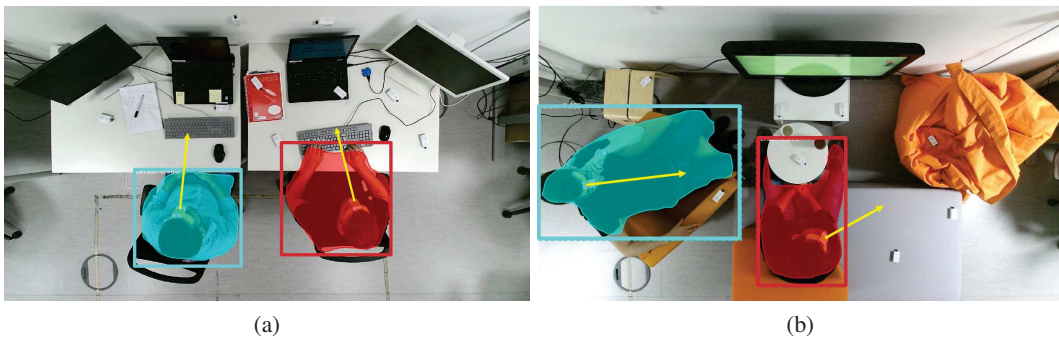
FIGURE 5.6: Illustration of people detection and head pose estimation. We detect people in the scene by using Mask R-CNN and then the detections are provided as input to the head pose estimator.

5.3.3 Person-perceived light estimation

Table 5.1 presents the quantitative results of our adopted light estimation approach. The table shows the average estimated error in *lux* values for both spatial (luxmeters 1-9) and gaze-gathered light estimation (luxmeters 10-11) cases. It can be easily noticed that the error, ε_{est}, for all luxmeters does not exceed the range of 100 *lux*, this yields an overall average light estimation error approx. 56 *lux* for Scene 1 and 36 *lux* for Scene 2. On the other hand, if we now consider only the luxmeters intended for evaluating the gaze-gathered light estimation, *i.e.* luxmeters 10 and 11, we notice that the error raises up to 94.7 *lux* and 55.4 *lux* for each scene respectively. This can be justified due to inaccuracies in the reconstruction
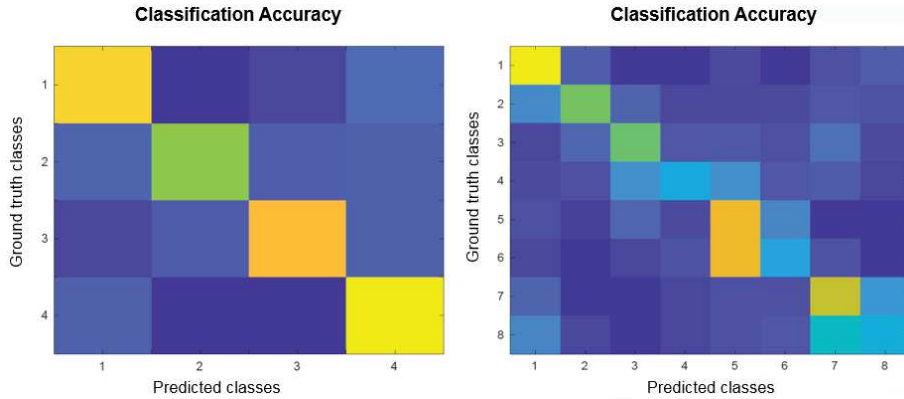
FIGURE 5.7: Confusion matrices of the head pose estimator. From left to right, the 4 and 8 classes confusion matrix respectively.

		Avg. error ε (in *lux*)												
		Luxmeters												
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**Avg. (1-9)**	**Avg. (11-10)**
Scene 1	ε_{est} (w.r.t. GT)	62.5	26.3	68.0	65.1	47.9	57.1	44.0	29.9	28.0	97.6	92.2	56.2	94.7
	ε_{est_d} (w.r.t. GT)	-	-	-	-	-	-	-	-	-	216.08	166.4	-	191.24
Scene 2	ε_{est} (w.r.t. GT)	35.3	33.8	44.0	20.1	31.5	39.6	23.6	27.9	27.3	41.7	69.2	35.8	55.4
	ε_{est_d} (w.r.t. GT)	-	-	-	-	-	-	-	-	-	55.42	151.93	-	103.68

TABLE 5.1: The values represent the average estimated illumination error over the different lighting activation w.r.t. the ground truth measurements, for both scenes. Columns 1-9 corresponds to the spatial average values for the corresponding installed luxmeters in the environment. By contrast, values in columns 10-11 consider those luxmeters for evaluating the human light perception.

of the 3D mesh areas corresponding to the head position and orientation of the occupants, as well as to the fact that the inter-reflections from the wall towards the sensors are limited due to incomplete reconstruction as an outcome of the limited FOV of the depth sensor. In any case, the fact that the average light estimation error does not exceed 100 *lux* indicates that the estimated illumination map can be considered reliable for describing the global illumination of the scene.

Furthermore, to demonstrate the applicability of our model, we used as explained a real person detector and a head pose estimator (making the pipeline completely automatic). In Table 5.1 the ε_{est_d} rows for column 10 and 11, illustrates the error based on the detectors output for both scene 1 and 2. It can be observed that while the average error w.r.t. the oracle is less than 100 *lux*, this error raises up to the range of 200 *lux* negative variation w.r.t. to the ground truth measurements. The last can be justified by erroneous head pose estimations, considering the large step size (90°) of the 4-class adapted classification problem. This further brings into discussion the fact that this error could further be substantially reduced by improving the head pose estimator.

Figure 5.8 shows in a graph analysis the values presented in Table 5.1. The left graphs show the absolute light estimation error (y-axis), as estimated for each of the 11 (9 for spatial and 2 for the human light perception) used luxmeter sensors (x-axis). The gray dots, forming each of the box plot boxes, represent the estimated error of each of the lighting scenarios for each scene while the pink box represents the central 50% of the data. The upper and lower
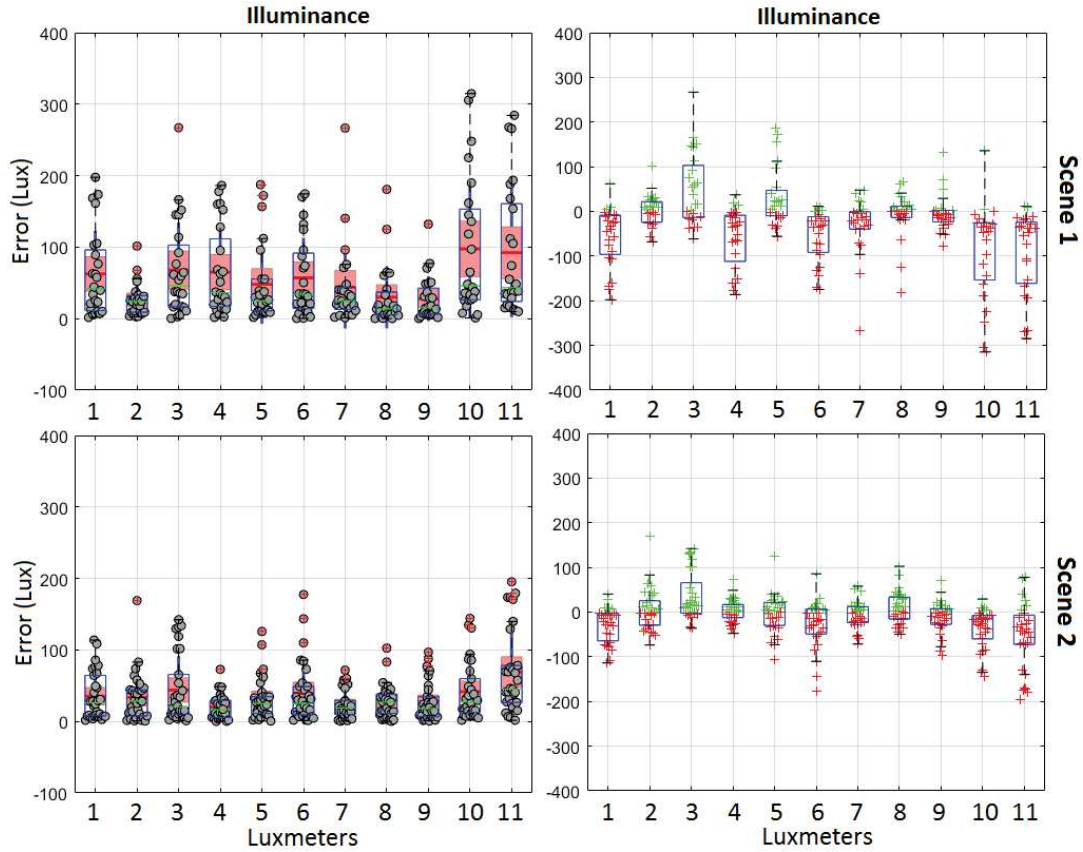
FIGURE 5.8: Scene 1 & 2 boxplot error evaluation (in *lux*) using based on the presented framework. The boxplots in the first and second columns show the absolute and signed illumination estimated error for each lighting scenario in each scene respectively.

vertical lines indicate the extension of the remaining error points outside it and the central red line indicates the mean error which comes in alignment with the values shown in Table 5.1. Similarly, the boxplots on the right present the signed illumination error accordingly. The green and red markers indicate whether the error is due to an over or under estimation of the illuminance at the sensor's location respectively. As it can be noticed in the most of the cases the error is a result of an under estimation of the illuminance which as explained earlier are a cause of the incomplete geometry of the scenes as we only consider the parts of the environment within the FOV of the camera sensors.

Finally, figures 5.9 and 5.10 visualise the illumination maps in the 3D space for one of the illumination scenarios in each of the scenes. As it can be seen the visualized illumination maps provide an accurate dense representation of the global illumination of the environment over time.

5.3.4 Applications of the invisible light switch

The idea of the Invisible Light Switch is straightforward as we have presented it in Chapter 1.2. Thus, in Table 5.2 we examine the applicability of the invisible light switch from the human perspective aspect (luxmeters 10-11) for different head orientation cases (VFOA) in the two scenes. The value Δ_{lux} provides the information regarding what is the impact to the light perceived from the occupants (based on the ground truth sensor measurements) on different light source combination scenarios. As it can be seen this gives us a range of 0-200 *lux* negative variation even to the most constraint scenario of having only two luminaires
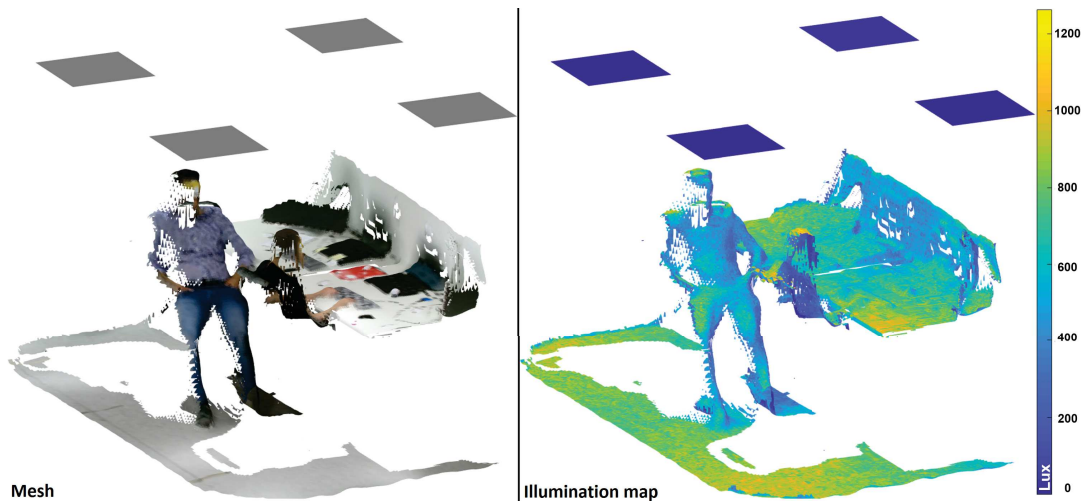
FIGURE 5.9: Illumination map of the full-lit scenario in scene 1 with a dense representation of the global illumination of the environment.
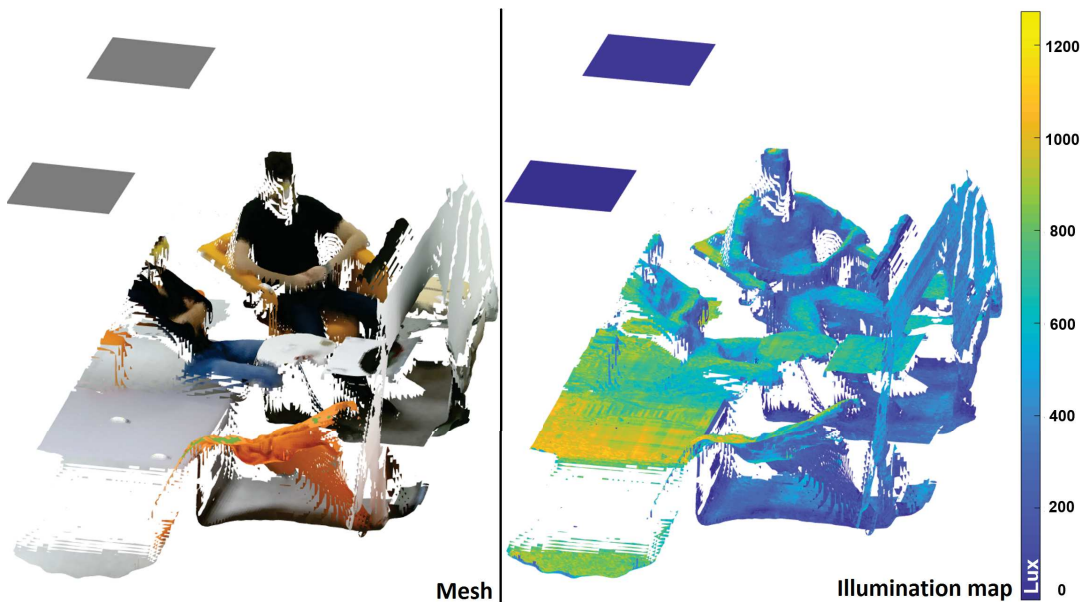

FIGURE 5.10: Illumination map of the full-lit scenario in scene 2. Notice the estimated illumination in the area in front of the occupants which is less bright in comparison to the one that are on their side. This is due to the body occlusion on the direct illumination coming from the luminaires from their back which is correctly estimated by the *ILS*.

active (the ones to the direct view of the occupants each time). If we connect this with the amount of *watt* that we can save for this corresponding lighting scenario, *i.e.* $\Delta_{watt} = 580.8$ *watt* w.r.t. to the full lit case, this can give us a total power efficiency of 297.1 *kWh* through a whole day. The value ε_{est} reports the light estimation error based on our framework, which as we can see again it settles within a range of 0-200 *lux* overall negative variation. This error shows us how our system aligns with the ground truth measurements, *i.e.* a lower ε_{est} error the better, and whether the same pattern described above could be followed. A visual example of the VFOA 1 case for scene 1 (see Table 5.2) can be seen in Figure 5.11. As it can be easily noticed the estimated illumination over the desk areas have the less affect as we switch off the peripheral light sources and still providing an optimally lit scenario while it is

		Scene 1						Scene 2						
		VFOA 1			VFOA 2			VFOA 1			VFOA 2			
Luminaire activations		3|4|7|8	2|3|4|5	3|4	3|4|7|8	2|3|4|5	3|4	3|4|7|8	2|3|4|5	3|4	1|2|3|4|5|6	2|3|4|5	1|3|4|6	3|4
Luxmeter 10	Δ_{lux} (w.r.t. full-lit)	116.15	123.77	189.01	85.4	123.8	163.85	84.23	93.69	151.92	106.52	148.12	157.07	191.15
	ε_{est} (w.r.t. GT)	167.2	144.09	102.73	235.3	200.1	163.28	85.85	94.1	43.76	22.94	12.97	13.59	25.69
	Δ_{watt} (w.r.t. full-lit)	387.2	387.2	580.8	387.2	387.2	580.8	387.2	387.2	580.8	193.6	387.2	387.2	580.8
Luxmeter 11	Δ_{lux} (w.r.t. full-lit)	97.68	125.15	169.72	167.4	86.34	194.37	62.67	118.21	153.02	99.17	154.28	167.93	194.85
	ε_{est} (w.r.t. GT)	194.63	171.74	131.55	91.14	128.7	70.21	15.26	67.87	5.39	9.4	241.12	2.81	203.69
	Δ_{watt} (w.r.t. full-lit)	387.2	387.2	580.8	387.2	387.2	580.8	387.2	387.2	580.8	193.6	387.2	387.2	580.8

TABLE 5.2: Quantitative analysis of four different head orientation class studies (VFOA), two for each scene. Δ_{lux} shows the discrepancy of different lighting scenarios w.r.t. the full lit scenario (reference). ε_{est} shows the corresponding average error of the estimated light in regards to the ground truth *lux* measurements and Δ_{watt} shows the discrepancy of the power consumption in *watt* considering the active/non active luminaires for each corresponding scenario.
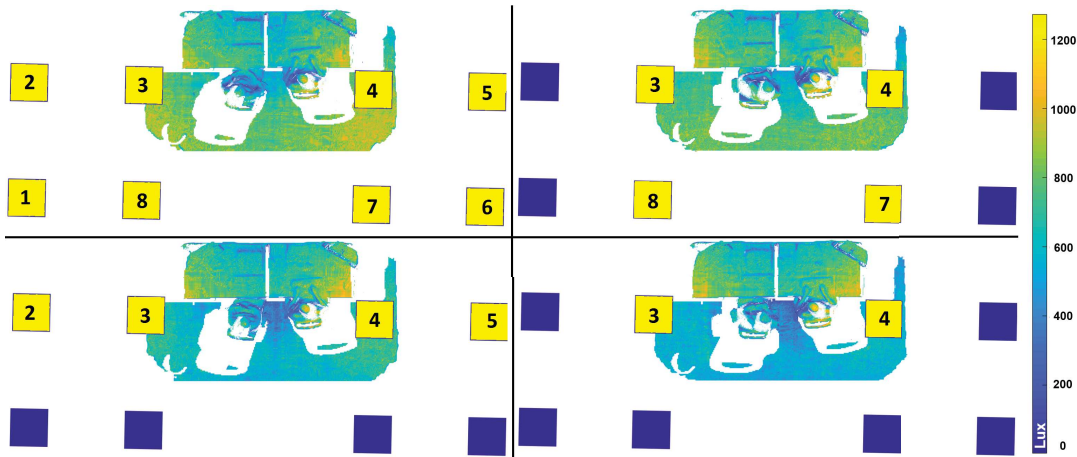
minimally lit.

FIGURE 5.11: Qualitative illustration of the VFOA 1 ablation study for Scene 1 presented in table 5.2. The top left corner shows the illumination map of the full lit case, in comparison to three other light scenarios. As it can be seen the estimated illumination over the desk area where the two occupants have their attention is less affected in comparison to the areas behind them. This show in practice how the invisible light switch application could be established.

5.4   Conclusion

In this Chapter we highlighted the importance of a human-centric aided lighting management system which targets productivity over a power saving framework. As a result, in this work we proposed and evaluated a practical (application-wise) system which tries to encapsulate three aspects, the ambient illumination, the human activity and the power efficiency. We also for the first time presented a complete system that estimates both the spatial and the individual gaze-dependent light intensity based on a camera-aided solution. It leverages light estimation and human perception for energy saving and advanced personalized lighting solutions and setting the baselines for a dynamic and interactive lighting management system. Therefore, we have shown that such a lighting model is suitable for the driving a new generation of

smart lighting management systems with main target power efficiency without compromising productivity and well-being.

Overall we illustrated a possible 66% of power saving by deploying our framework as the "Invisible Light Switch" application which can be also used to exploit an optimal illumination pattern for a given human activity. The remarkable aspect behind this achievement is not only the high percentage of the power efficiency but the point that this is done by managing not to compromise the gaze-gathering illumination of the occupants since the negative variation was shown to be within the range of 200 *lux* even in the most constraint scenario.

The proposed system could be further extended so that to be able to forecast the occupants future activities will help understand (in advance) which lights should be switched on, avoiding the user to continuously act on the illumination system, and predicting for the user the illumination scenario that he wants to have or is suitable for the related circumstance/activity.

Chapter 6

Conclusion, Perspectives and Future Work

In this thesis, we presented some novel contributions to the state of the art regarding the integration and the development of a smart lighting management with the use of advanced computer vision and machine learning. We have proposed and evaluated several methods and tools that attempt to extract useful cues from an indoor environment and utilize a camera sensor for integrating light calibration, light estimation and human light perception into a end-to-end light management application. To this extent we evaluated and extended well known computer vision and numerical algorithms that work with non-conventional camera systems, *i.e.* RGB-D cameras by taking into account features and aspects of the real world. To summarize our contributions and align with Section 1.4 we have divided our research in the following blocks:

- We have examined and shown whether factorization techniques could play a semantic role for light decomposition and light source identification, and in general as a novel approach to light source calibration. We have shown that by taking advantage of the physical features of light, *i.e.* additivity and non-negativity features that are explicitly exploited in the non-negative matrix factorization techniques is possible to separate mixed sources to their individual components and at the same time get an understanding about their intensity. We further deployed a novel benchmark dataset and metrics in order to evaluate this outcome and justify the numerical correctness of this approach.

- In order to reason about a real time dense light estimation of the whole scene and not only to a specific spot provided by a conventional luxmeter or from a timely simulation, we have extended a well known global illumination model (*i.e.* radiosity) with real light source properties and by eliminating well known limitations (*e.g.* timely operational and simulation sessions, point based light sources, isotropic light perception, known photometric properties and geometry, *etc.*). In this line, our two major contributions were: **i)** the introduction of a camera sensor as the input to a lighting management system for lighting estimation and **ii)** the numerical extension of the radiosity model by inducing the light propagation and perception distributions, *i.e.* LDC and LSC, of real life light sources and light sensing devices respectively. Consequently the major advantages of this solution is that it does not require any previous knowledge of the environment from the user and at the same time it can adapt to any dynamic changes due to the live feed from the camera. Thus, providing a reliable dense light estimation overall.

- Following the idea of taking advantage of a camera sensor for measuring the ambient illumination, we proposed a novel hybrid application which takes into account the aspect of the human occupancy and head pose estimation in order to retrieve information regarding the gaze-gathering light and how this could be used for ego-centric lighting

and power saving light management solutions. By means of a state-of-the-art variant of a Mask R-CNN model the system was able to identify the location and the posture of the individual occupants in the scene and based on the estimated gaze-gathering light within their visual attention to take decisions on lighting over time.

- Finally we have accommodated all these pipelines together into a complete end-to-end system and we have set the baselines for adapting numerical computing, computer vision and visual computing solutions into real life smart lighting applications. Thus, we for the first time with this study we bring together light estimation and human perception with the light management, and propose the idea of a novel autonomous *"invisible light switch"*.

More specifically in Chapter 3 we have shown that diverse factorization approaches and more specifically the idea of non-negative matrix factorization is well suited for the modeling of lights (*i.e.* estimating the visual appearance of each light source, including the natural light, and of identifying the active lights) in indoor environments in the 2D image space. Moreover, we illustrated some preliminary application scenarios, which despite adopting a simple approach could be used in several application such as smart lighting and ambient intelligence. The main problem with such techniques is that they demand an extended computation time as well as that they are dependent to the actual scene that are operated, limiting in that way their usage as generic solution. However, these are problems that we wish to address in the future.

In Chapter 4 we proposed and evaluated the proposed RGBD2lux light modeling and ambient light estimation method as a viable option for estimating light intensity in real environments. We provided a new direction for the research in the lighting field and in general for the lighting industry based on imaging and visual computing. On one side, we have shown that continuous and reliable camera-aided measurements of light on real environments is now viable, enabling the study of how light dynamics do influence human health, habits, social activities, and not only, to quote a few. Thus, to this end we answered to the question whether 3D modeling of a scene with a single camera is sufficient for reliable real scene light modeling and we set the baselines for adapting computer vision solutions for smart lighting applications. The latter though lead us to one of the main problems that arose and which is the limited view of the scene and we would like to address as a future work. The limited view of the scene contributes into a higher average lighting estimation error due to the fact that ignores the inter-reflections from the non visible surfaces of the scene. Therefore, we believe that the orthogonal combination of an RGB-D system with omnidirectional (fish-eye) cameras, as introduced in [PLG16], could be a viable solution to obtain a more complete 3D reconstruction of the scene geometry and therefore to a more accurate light intensity estimation. Another aspect that could be considered for future improvement would be the estimation of the photometric properties of the material in a more direct way, *e.g.* with the use of a real time intrinsic decomposition algorithm or similar approach [KGTM+17], which in principle will allow our system to be able to address other material and scene properties (*e.g.* specular components, caustics, *etc.*).

Finally in Chapter 5 we have studied the extension of the RGBD2Lux pipeline to handle in practice how the light dynamics can be adjusted, by taking into account the human occupancy and posture along with the spatial illumination, and providing in practice a desired power friendly framework without affecting the gaze-gathering light in the visual attention area of the occupants in the scene. We have shown through the key idea of the invisible light switch that indeed the proposed camera-aided lighting model is suitable for the driving the new generation of smart lighting management systems to more power efficient solutions without compromising the illumination levels within the frustum of attention of the occupants.

As stated already lighting can be used for more than just to illuminate, good-quality lighting can improve well-being, visual performance and interpersonal communications while well-lit workplaces provide visual comfort and improvement of intellectual productivity and effectiveness, however smart lighting is still a novel field of research. Therefore, this research has been proposed within the framework of combining smart lighting and computer vision in order to provide a new generation of lighting management systems which targets high light quality and power efficient solutions. These both aspects could open new horizons in the research field of lighting for different use cases, *e.g.* healthcare (surgical rooms), productivity and good mood of the employees, increased safety of the workers in working environments, green building and further deployment of daylight harvesting to name a few.

Overall the fact that the proposed system can achieve a 66% of power consumption in the most constraint lighting scenario within only a 200 *lux* range of average illumination estimation error for the gaze-gathering case and even less for the spatial light intensity estimation shows the high impact and possibilities that such a system can have in the future advancement of the lighting field.

In the future we would like to explore novel techniques that would optimize our lighting estimation procedures as well as to be possible to address scenes with natural light. The idea of creating a fully automatic solution where the detection of the light source positioning and intensity in the 3D space, in conjunction to the work shown in Chapter 3, could be another aspect for future research. Moreover, it would be interesting to try some neural-network-based approaches for 3D layout estimation [ZCSH18; FPLG18; FFPD+18] of the indoor environments. This arises from the fact that neural-networks and especially deep learning is lately almost taking over the world of computer vision and visual computing with extremely remarkable results. The latter could be further extended with some data driven scene semantic analysis and understanding procedures [DRBR+18; JDN19; ADDS+19] providing some interesting solutions into the field of light design and modeling.

To conclude, we could say that the study of light modeling it is still an open research problem with many considerations to take into account. This thesis proposes a collection of new methods that try to address some interesting tasks related to light modeling, but we believe that we just scratched the top of the iceberg and certainly there are more aspects to be reviewed. The current status of profound interest in computer vision and lighting along with the progressive enhancements in technology to be expected in the near future are very promising and encouraging to keep working in this topic.

Appendix A

Light Source Calibration Additional Material

Contents

In this appendix we complement Sec. 3.7 of Chapter 3. We provide some insights regarding the ablation data-subsets as well as the corresponding results for Scene 2.

A.1 Ablation data-subsets

In figures A.1, A.2, A.3, A.4 and A.5 we visualize the different ablation studies presented in Sec. 3.6.1.

LIT-All.

(a) Scene 1

(b) Scene 2

FIGURE A.1: LIT, full complex subset.

LIT-Artificial.

(a) Scene 1

(b) Scene 2

FIGURE A.2: LIT-Artificial, natural light is excluded.

LIT-Static.

(a) Scene 1

(b) Scene 2

FIGURE A.3: LIT-Static, frames with artificial motion are excluded.

LIT-Artificial-Static.

(a) Scene 1

(b) Scene 2

FIGURE A.4: LIT-Artificial-Static, frames with artificial motion and natural light are excluded.

LIT-Sources.

(a) Scene 1
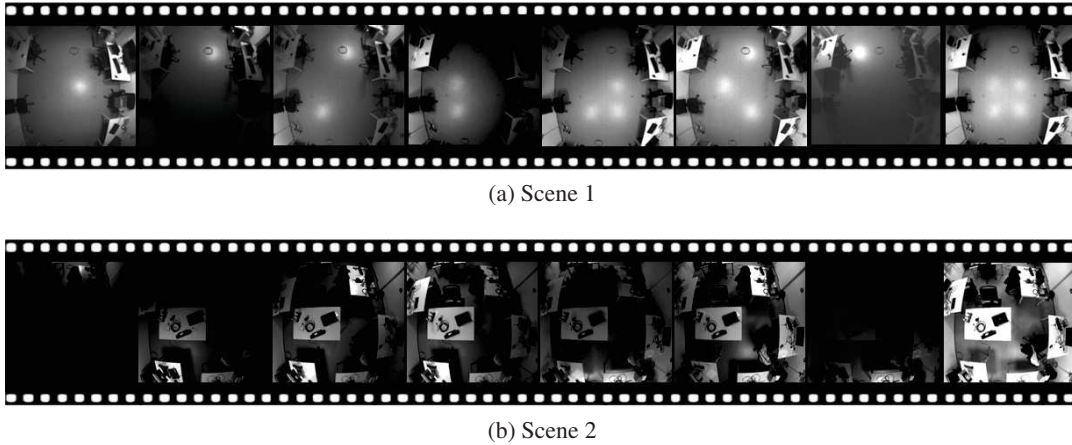
(b) Scene 2

FIGURE A.5: LIT-Sources, frames without artificial motion and natural light.

A.2 Benchmark evaluation for Scene 2

Here we analyze the methods introduced in Chapter 3 on the tasks of LIT-EST (Sec. 3.7.1), LID-ID (Sec. 3.7.2) and RE-LID (Sec. 3.7.3) for the second scene.
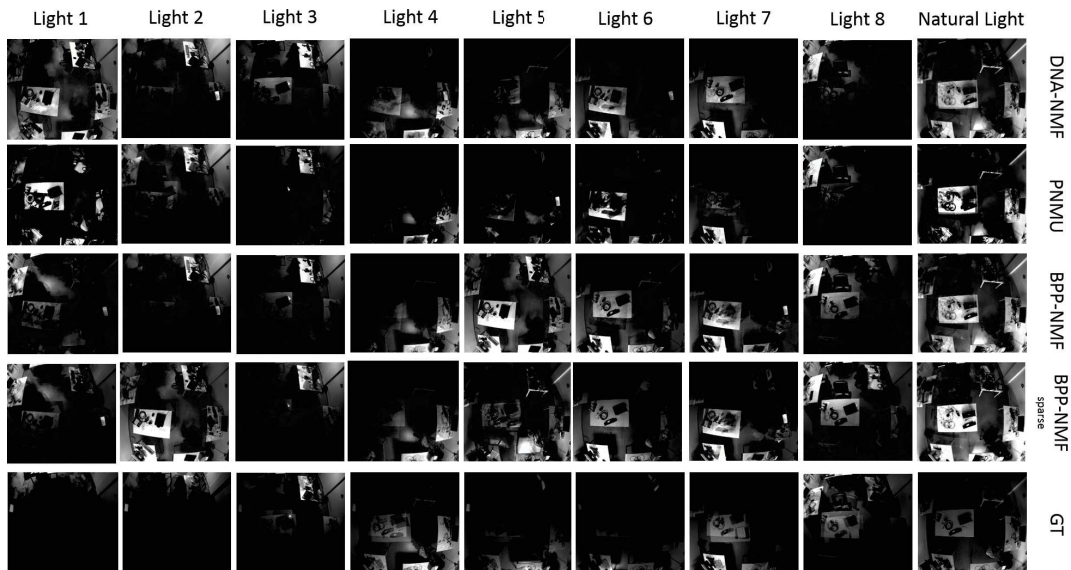
FIGURE A.6: Light sources estimates for each selected technique on the LIT-All set and subsets, Scene 2. For the PNMU and BPP-NMF techniques we show only the best performing alternatives, see Tables A.1, A.2 for all details. Column-wise, correct light source images should look as close as possible to the ground truth in the last row.

A.2.1 Light source estimation results

Table A.1 reports the light source estimation errors ε_{EST}, measuring the quality of the computed light source set $\tilde{\mathcal{L}}$; the source images for the Scene 2 are also visualized in Figure A.6.

In the first row of the table (Scene 2, LIT-full), one could see the superiority of NMF based techniques, DNA-NMF and NMF-BPP. DNA-NMF comes out as top performing with the reconstruction error of 56.61%, shadowed by BPP-NMF at a distance of $\sim 0.2\%$. However, BPP-NMF manages to reconstruct 6 out of 9 light sources, whereas DNA-NMF reconstructs 7 out of 9 light sources. PNMU is underperformed significantly by a margin of $\sim 11\%$, this can be validated by the number light sources constructed (5 out of 9) by PNMU which are in the mid-range in comparison to the best method. Sparsity degrades the performance of both approaches(PNMU and BPP-NMF), although the degradation of performance is more pronounced in PNMU as compared to BPP-NMF. The worse performers are again the PCA, ICA, PNMU-sparse methods, confirming once more the importance of non-negativity and of a correct sparsity prior.

Finally we should also note that considering LIT-All in Scene 1 and Scene 2, it is evident from the magnitude of reconstruction error and visual inspection that Scene 2 is more challenging, due to movements, physical configuration of luminaires, *etc*.

Ablation studies

As an ablation studies, we started by studying the influence of natural light by eliminating it from the LIT benchmark (LIT-Artificial). The order of the techniques remains the same, DNA-NMF being the top performer followed by BPP-NMF. However, the degradation in performance of BPP-NMF (56.8% to 58.8%) is surprising, since due to the dynamic nature of natural light(change over the course of day) it should be harder to model. However, upon visual inspection it is observed that due to the movement of objects and people, the scene becomes more challenging. Secondly, it is observed that sparsity brings in improvement of 2% in the case of BPP-NMF. However, as discussed in the Sec. 3.7 that sparsity needs to be carefully designed (l_1-norm), as PNMU get worse after the induction of the sparsity.

Building upon the previous findings, we removed motion from the dataset (LIT-static), the reconstruction error in the case of BPP-NMF drops by $\sim 27\%$, confirming that most errors are due to motion in the scene, with its sparse variant providing again better result.

Similar to Scene 1, in Scene 2 it can be seen that most approaches perform well in LIT-Artificial-Static (without movement and natural light) BPP-NMF-sparse has the lowest reconstruction error $\sim 27.7\%$. However as discussed the magnitude of the error in comparison to scene 1, illustrates that most approaches find it hard to decompose Scene 2 due to the challenging nature of the scene.

Finally, as expected it can be seen that the exposure of individual light sources (LIT-Sources) results in a marginal gain compared to LIT-All.

A.2.2   Light identification results

Testing on Scene 2 (Table 3.2, brings to somewhat similar conclusions as in Scene 1.

Table A.2 (LIT-All row), all techniques perform quite similarly(within 1%) of each other with PNMU being slightly better with the activation error of $\sim 22\%$. Importantly, the type of sparsity is an important factor and the superiority of l_1-norm can be further validated in the case of BPP-NMF. Again the LIT-Artificial appears more challenging for all techniques due to the after-sunset motion (*c.f*. second row); LIT-Static is in fact an easier task (BPP-NMF reducing its error to 15.03%); LIT-Static-Artificial further points out the fact that Scene 2 is a challenging scenario (luminaires topology)(15.7% for BPP-NMF-sparse).

We further compare the techniques to the Oracle (final column), which "knows" the true light sources.

Similar to Scene 1, we notice that the Oracle error on LIT-All is 10.5%. This further stresses the fact that even after a perfect light calibration, the light scene understanding still needs further research.

ε_{EST}	Scene 2						
	PCA	**ICA**	**PNMU**	**PNMU sparse**	**DNA-NMF**	**BPP-NMF**	**BPP-NMF sparse**
LIT All	0.8317 (3/9)	0.6571 (5/9)	0.6725 (5/9)	0.7896 (4/9)	**0.5661** **(7/9)**	0.5683 (6/9)	0.5872 (6/9)
LIT Artifcial	0.8091 (4/8)	0.7772 (4/8)	0.6715 (4/8)	0.6827 (5/8)	0.5589 (5/8)	0.5883 (5/8)	**0.5584** **(5/8)**
LIT Static	0.6886 (5/9)	0.7801 (4/9)	0.5811 (5/9)	0.7293 (5/9)	0.3392 (7/9)	0.3172 (7/9)	**0.2928** **(7/9)**
LIT Artificial-Static	0.8410 (3/8)	0.7618 (4/8)	0.5166 (5/8)	0.6866 (5/8)	0.3436 (6/8)	0.3217 (6/8)	**0.2775** **(7/8)**
LIT Sources	0.8246 (3/9)	0.7774 (3/9)	0.6901 (5/9)	0.7839 (4/9)	0.5584 (6/9)	0.5894 (6/9)	**0.5333** **(7/9)**

TABLE A.1: Comparative evaluation of light source reconstruction techniques over the LIT-All set and subsets for Scene 2. The value $\varepsilon_{EST} \in [0, 1]$ is the error measures (lower is better). Number in parentheses indicate the successfully matched light sources, see Section 3.7.1.

ε_{ID}	Scene 2							
	PCA	**ICA**	**PNMU**	**PNMU sparse**	**DNA-NMF**	**BPP-NMF**	**BPP-NMF sparse**	**Oracle**
LIT All	0.2874	0.2568	**0.2230**	0.2819	0.2395	0.2384	0.2346	0.1056
LIT Artifcial	0.3082	0.2988	0.2461	**0.2443**	0.2506	0.2497	0.2527	0.0829
LIT Static	0.2446	0.2719	0.2052	0.2187	0.1694	0.1592	**0.1503**	0.0413
LIT Artificial-Static	0.2906	0.2926	0.2121	0.2146	0.1750	0.1679	**0.1574**	0.0316
LIT Sources	0.2889	0.2937	0.2371	0.2889	**0.2311**	0.2443	0.2318	0.1056

TABLE A.2: Comparative evaluation of selected techniques in terms of light activation error ε_{ID} over the LIT-All set and subsets. $\varepsilon_{ID} \in [0, 1]$ (lower is better), see Section 3.7.2.

Second, the natural light emerges as the largest challenge (the Oracle error on LIT-Artificial drops to 8.2%). Removal of movement and natural light drops the Oracle error to mere 3.1%. Finally, it is clear that modelling of natural light and movement in the scene is the most challenging task in the proposed benchmark.

Appendix B

RGBD2Lux Additional Material

Contents

In this appendix, we provide the corresponding results for room_2 likewise to the results shown in Chapter 4 for room_1. Moreover, we further present some extended results that we have conducted to multiple different rooms in order to prove the generalization of our solution.

B.1 Illustration of room illumination variants
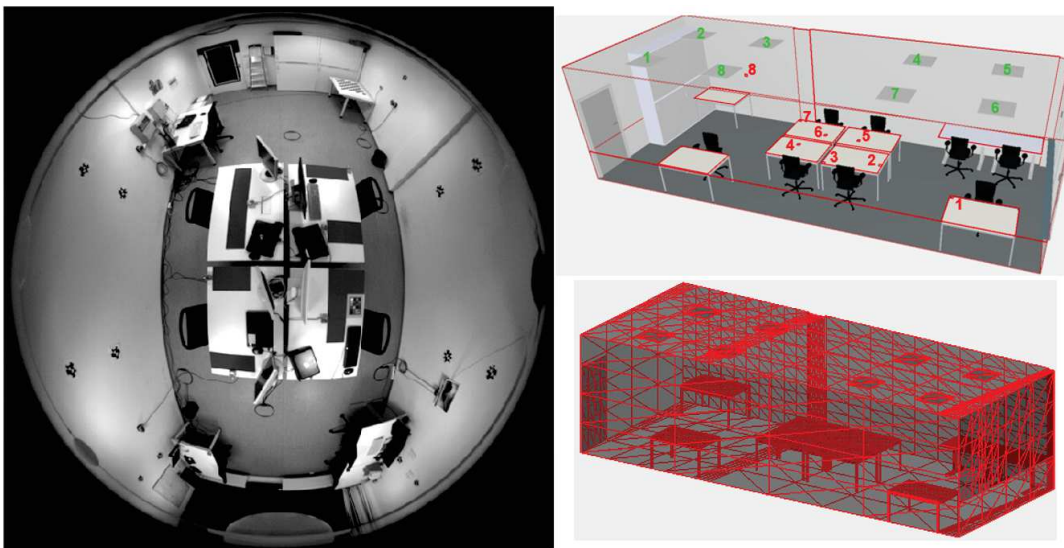
First we illustrate the geometric setup for room_2.

FIGURE B.1: room_2, as it appears from top view perspective (*left*) and corresponding CAD designs (*right*). Note the *red* numbers in the CADs indicating the luxmeter positions and the *green* numbers indicating the luminaire positions. The room measures 10×4 m^2.

The room_1 (shown in Figure 4.4) and room_2 (shown in Fig. B.1) have differences that impact the overall lighting condition in the experimental setups. The position of laptops, chairs, monitors and other objects in the scene vary, but the desks are almost in the same position, with respect to the camera. This allowed to compare similar setups and the effects

of objects in the scene. For example a monitor close to a luxmeter may shadow some of the luminaire, as it happens with luxmeter 4 and luminaire 3.

More importantly, note the difference of the **panels** between the central desks. By making a room_2 similar to room_1 but without panels, we could study the error increase due to a larger number of light inter-reflections. In fact, without the panels, luminaires reach most part of the room with their light, both in terms of direct and indirect lighting, *e.g.* wall, ceiling reflections.
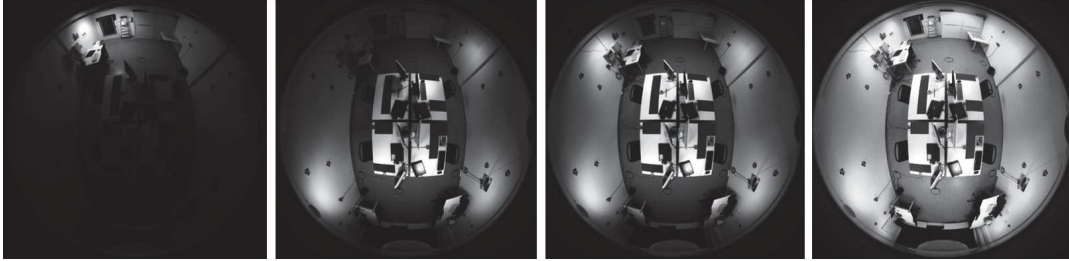
FIGURE B.2: The images show 4 illumination variants within room_2. From the left, the images illustrate the illumination provided by 1, 3, 4 and all 8 luminaires switched on in the scene. From left to right, the same luminaires were on as in Figure 5 for room_1 in the paper.
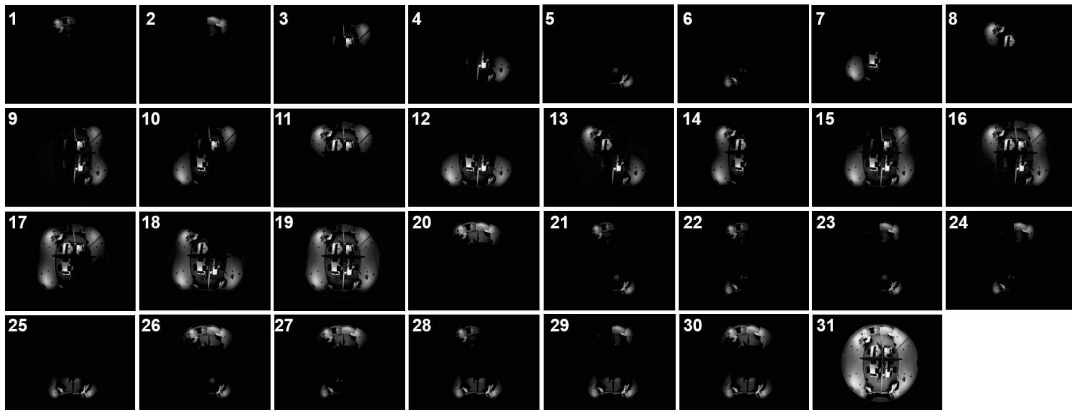
FIGURE B.3: Illustration of all 31 illumination variants within room_1.

Figures B.2 and B.3 show how different the illumination of the two rooms may be, when alternating the activation of the eight luminaires. We select for visualisation purposes 31 light configurations which are the most representative ones among the all possible configurations that can be observed in a time-lapse sequence. In Figure B.2, we show 4 illumination variants for room_2 similarly to Figure 4.5 for room_1. Note the strong illumination difference, mainly due to the absence of cubicle panels in room_2.

Figure B.3 shows the 31 light combinations for room_1. A likewise pattern is applied for room_2.

B.2 Further evaluation comparisons

For the sake of completeness, we report in Figure B.4 the boxplot errors and comparison between Relux and RGBD2Lux for room_2 (similarly to Figure 4.9 reported results for room_1). Results confirm the performance of our approach in coincidence with the discussion in Sec. 4.3.4.
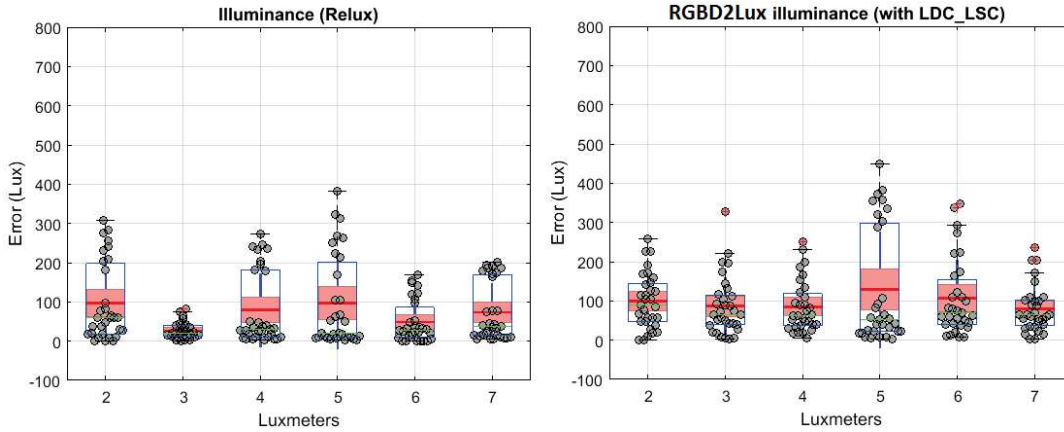
FIGURE B.4: Scatter errors with sign for room_2, comparing Relux vs RGBD2Lux.

Figures B.5, B.6, B.7 illustrate the same scatter plot results as Figures 4.6, 4.8 and B.4, but reporting errors **with sign** (*green* or *red* whether over- or under-estimating the true illumination) respectively.

The discussion from Sec. 4.3 applies to the general error analysis. Furthermore, Figure B.5 allows conclusions on a general error bias. In fact, when LDC and LSC are not considered, the simulations under-estimate the illumination more, *c.f.* Relux estimates (*1a, 2a*) and **Ours with CAD no_LDC_LSC** (*1b, 2b*).

By contrast, including LDC and LSC results in more over-estimation errors, *c.f.* **Ours with CAD LDC_LSC** (*1c, 2c*), further to resulting in a lower overall error.
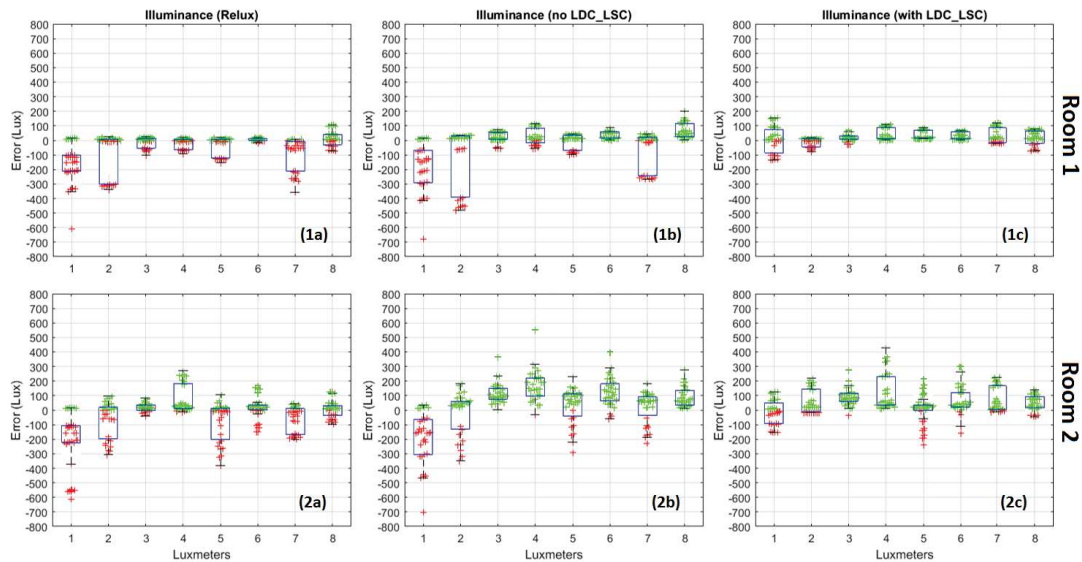
FIGURE B.5: Scatter error plots **with sign**. Green and red crosses indicate over- and under-estimation errors. See corresponding boxplot in Fig. 4.6

The scatter signed error plots in Figures B.6 and B.7 as also explained in Fig. 4.9 are comparing Relux to our proposed RGBD2Lux for the two rooms. This helps us to understand better the discrepancy and the cause of the error in our estimations. While the general performance is comparable (average absolute error), ours seems to under-estimate illumination, more than over-estimating. We explain this as due to using only the visible part of the scene,

i.e. the depth-camera view. Such as half-view is missing inter-reflections from the ceiling and from the non-visible portions of walls, thus naturally under-estimating illumination.
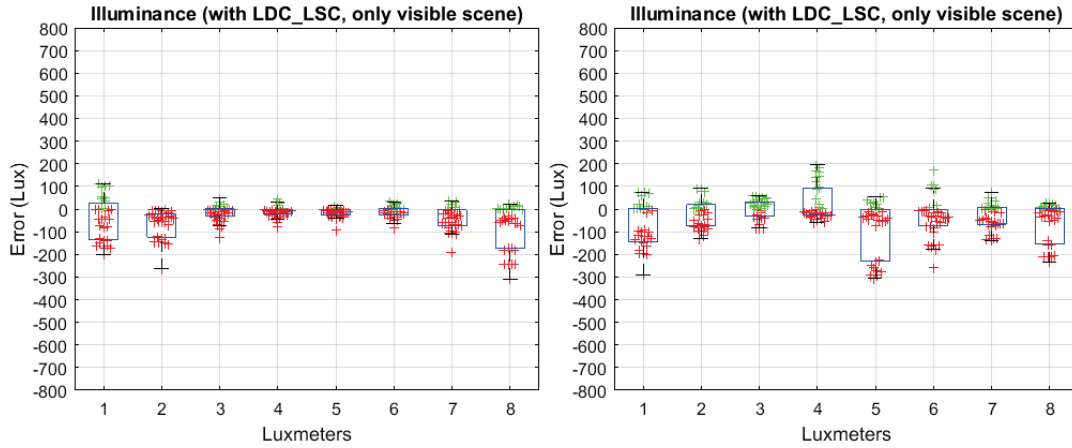

FIGURE B.6: Room1 & 2 signed *lux* boxplot error evaluation using only the visible CAD information on the radiosity with _LDC_LSC. See corresponding boxplot in Fig. 4.8 in the paper.
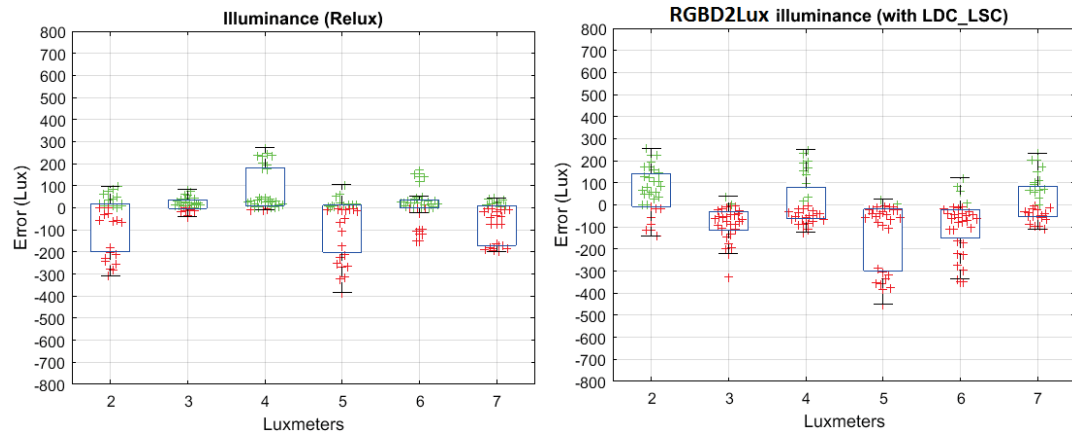

FIGURE B.7: Scatter errors with sign for room_2, comparing Relux vs RGBD2Lux. See corresponding boxplot in Fig. B.4.

B.3 Illustration of typical failure cases

Here we present some typical errors of the Relux in regards to our approach as shown in Figure 4.7.

Figure B.8 illustrates the error of Relux for the estimation of the illumination corresponding to luxmeter 1 in room_1, across all 31 lighting conditions. Note the consistent underestimation errors. We explain this due to the lack of the LDC and LSC modelling within Relux. In particular, since Relux misses the LSC modelling, it assumes that light comes equally strong from each direction. By contrast, direct light (orthogonal to the observing sensor) should be stronger then light perceived under an angle. In order to achieve a good average light estimation, Relux must therefore under-estimate the more orthogonal lights. This shows in Figure B.8, whereby the luminaires closer to the luxmeter 1 (6 - directly above- and 4,5,7 - just around it; *c.f.* Figure 4.4 for the geometry) are consistently under-estimated. This
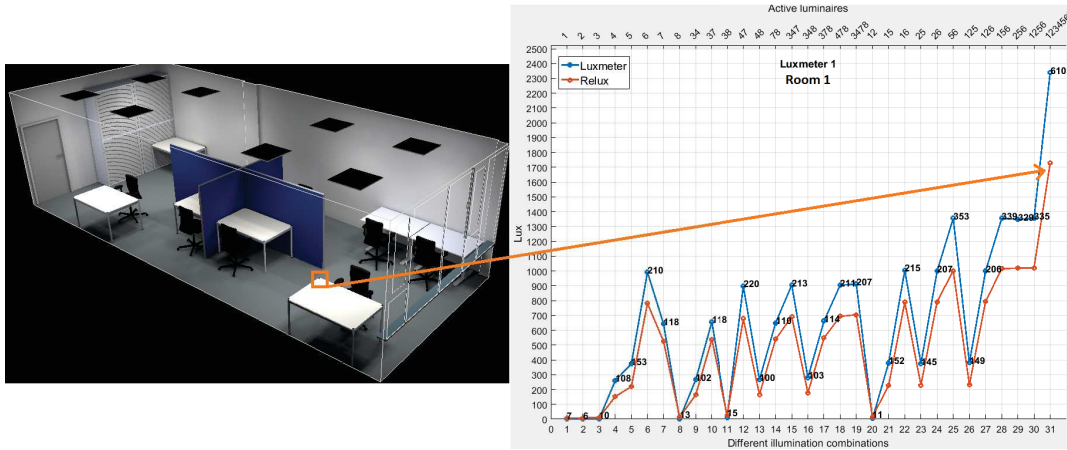
FIGURE B.8: A failure case of luxmeter 1 in room_1 with the Relux simulation. The two x-axes show the 31 different illumination cases (bottom) and which light source was active for each case (top) respectively.)

is also the case when all of them are switched on simultaneously (which we highlight with the orange arrow), where all errors sum up.

B.4 Extension of the RGBD2Lux pipeline to multiple rooms

A final experiment that we have applied was to test and evaluate the generalization of our proposal to different indoor environments with different structural design, lighting setup (different amount, type and positioning of the light sources) and of course under different lighting activations of the installed light sources.

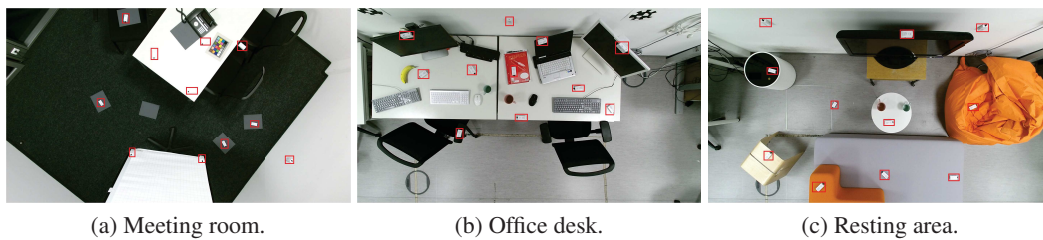| (a) Meeting room. | (b) Office desk. | (c) Resting area. |

FIGURE B.9: Different rooms that we applied our *RGBD2Lux* pipeline. The red boxes show the scattering of the luxmeters within the scene. The scenes were equipped with different amount and type of light sources. The scene in picture *(a)* was equipped with two (2) fluorescent luminaires while scenes *(b)* and *(c)* with eight (8) LED luminaires

Figures B.10, B.11, and B.12 show the corresponding overall spatial light estimation outputs for each of the scenes in Figure B.9a, B.9b, and B.9c respectively. If we look on the boxplots we can clearly notice that the discussion from Sec. 4.3 applies to the general error analysis in these cases as well, while the qualitative representation of the estimated illumination map over the 3D space (top right corner of each figure) show us visually the correctness of the our estimated output. To this end the evaluation of our pipeline in multiple rooms, with different structural design and content and with different amount and type of luminaires validates the ability of our method to address and provide a reliable solution for light estimation to different challenging indoor environments.
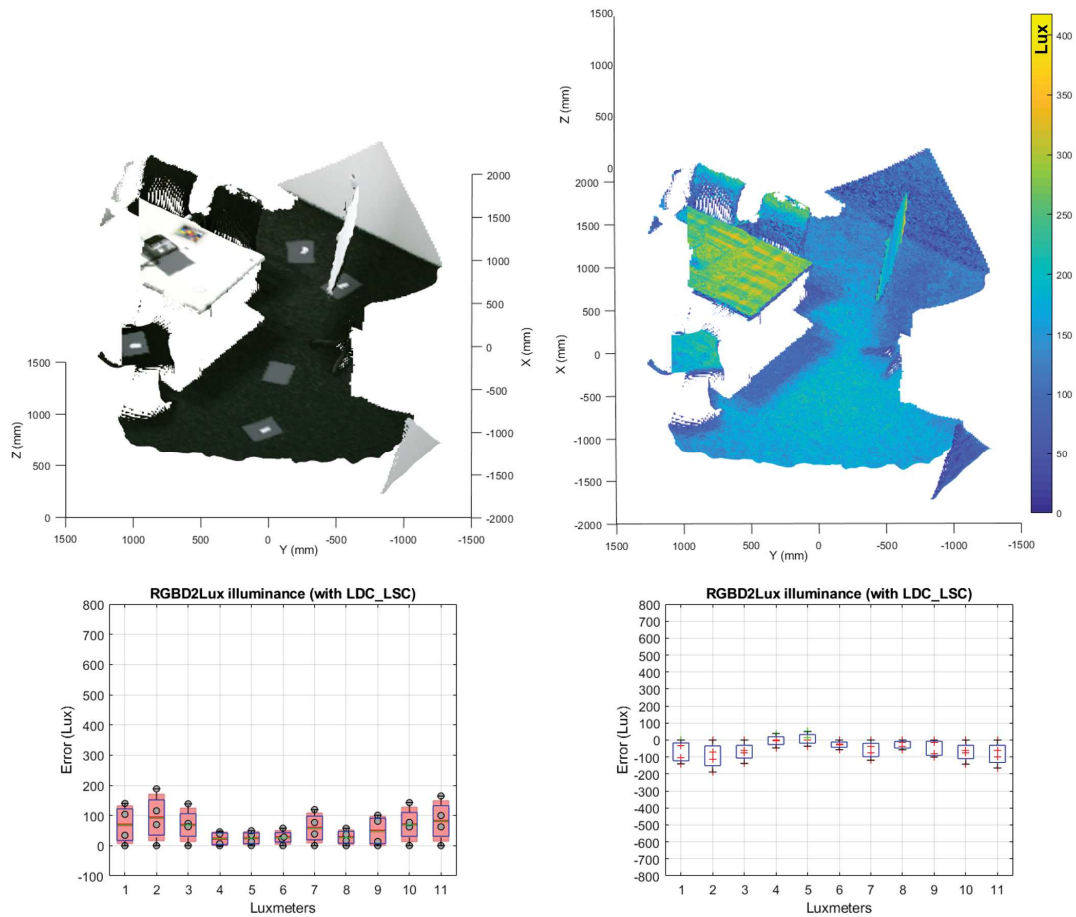
FIGURE B.10: Light estimation output for scene B.9a. On the top left we can see the normal output of the RGBD sensor, on the top right the color pixel intensities were replaced from the corresponding estimated *lux* values providing us with the illumination map in the 3D space. The boxplots in the bottom row show us the absolute *(bottom left)* and the signed *(bottom right)* average light estimation error in *lux* respectively.
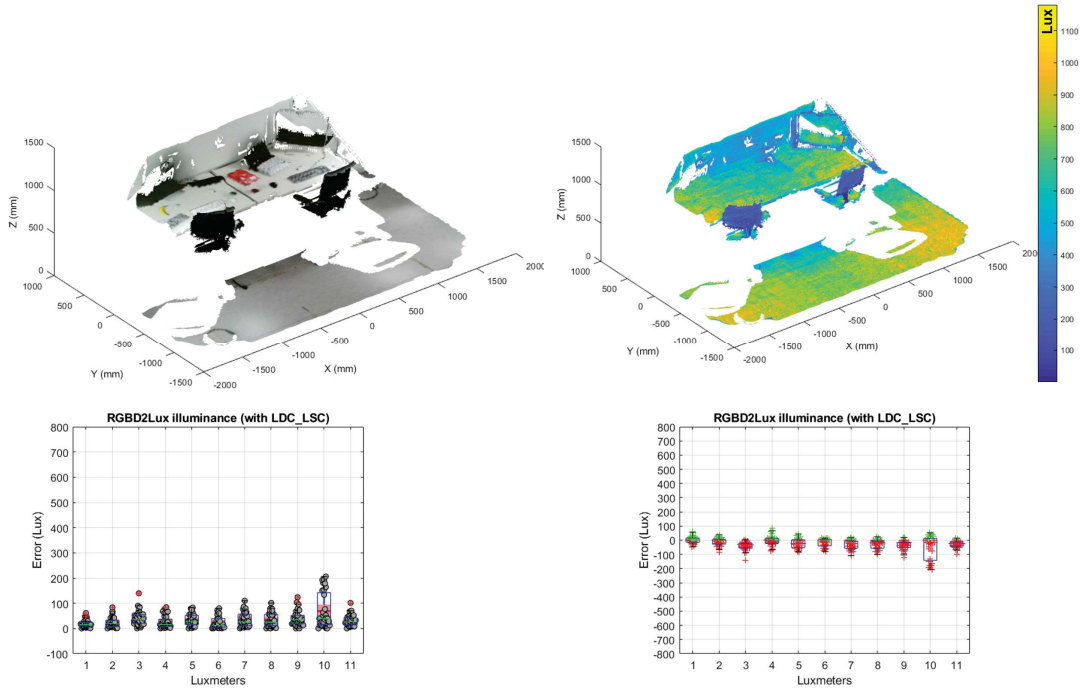
FIGURE B.11: Light estimation output for scene B.9b. On the top left we can see the normal output of the RGBD sensor, on the top right the color pixel intensities were replaced from the corresponding estimated *lux* values providing us with the illumination map in the 3D space. The boxplots in the bottom row show us the absolute *(bottom left)* and the signed *(bottom right)* average light estimation error in *lux* respectively.
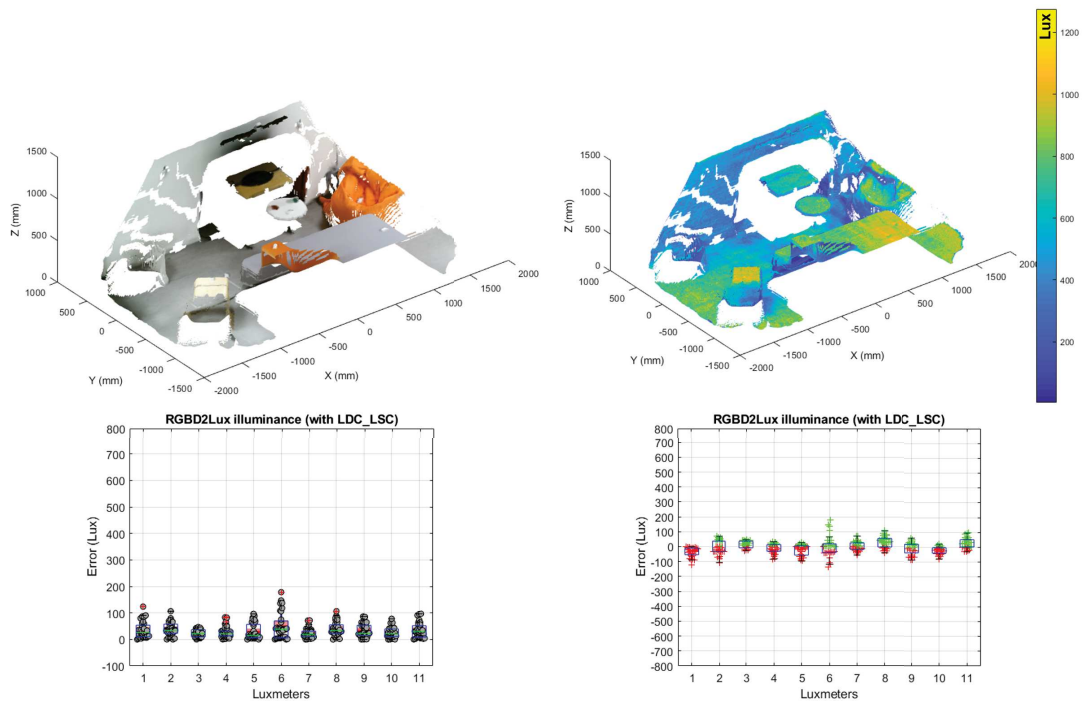

FIGURE B.12: Light estimation output for scene B.9c. On the top left we can see the normal output of the RGBD sensor, on the top right the color pixel intensities were replaced from the corresponding estimated *lux* values providing us with the illumination map in the 3D space. The boxplots in the bottom row show us the absolute *(bottom left)* and the signed *(bottom right)* average light estimation error in *lux* respectively.

Bibliography

[ADDS⁺19] Armen Avetisyan et al. "Scan2CAD: Learning CAD Model Alignment in RGB-D Scans". In: *Proc. Computer Vision and Pattern Recognition (CVPR), IEEE*. 2019.

[AFG13] Jens Ackermann, Simon Fuhrmann, and Michael Goesele. "Geometric Point Light Source Calibration". In: *VMV*. 2013.

[AZ91] Leslie Adams and David Zuckerman. "The effect of lighting conditions on personal space requirements". In: *The journal of general psychology* 118.4 (1991), pp. 335–340.

[BB16] Benoit Beckers and Pierre Beckers. "Fast and accurate view factor generation". In: *FICUP, An International Conference on Urban Physics* (Sept. 2016), p. 1.

[BERN17] Craig Bernecker. "Square Law Dimming: Presumed Perception or Reality?" In: Leducation, New York, 2017.

[BJK07] Ronen Basri, David Jacobs, and Ira Kemelmacher. "Photometric stereo with general, unknown lighting". In: *International Journal of Computer Vision* 72.3 (2007), pp. 239–257.

[BONM⁺15] Bastiaan J. Boom et al. "Interactive light source position estimation for augmented reality with an RGB-D camera". In: *Computer Animation and Virtual Worlds* (2015), n/a–n/a. ISSN: 1546-427X. DOI: 10.1002/cav.1686. URL: http://dx.doi.org/10.1002/cav.1686.

[BOYC04] Peter R Boyce. "Lighting research for interiors: the beginning of the end or the end of the beginning". In: *Lighting Research & Technology* 36.4 (2004), pp. 283–293.

[CAI16] H Cai. "Luminance gradient for evaluating lighting". In: *Lighting Research & Technology* 48.2 (2016), pp. 155–175.

[CBNR11] Manmohan Chandraker et al. "On the Duality of Forward and Inverse Light Transport". In: *IEEE Trans. Pattern Anal. Mach. Intell.* 33.10 (Oct. 2011), pp. 2122–2128. ISSN: 0162-8828. DOI: 10.1109/TPAMI.2011.124. URL: http://dx.doi.org/10.1109/TPAMI.2011.124.

[CD74] Suzanne J Carr and James M Dabbs Jr. "The effects of lighting, distance and intimacy of topic on verbal and visual behavior". In: *Sociometry* (1974), pp. 592–600.

[CG15] Gabriella Casalino and Nicolas Gillis. "Sequential Dimensionality Reduction for Extracting Localized Features". In: *CoRR* (2015). URL: http://arxiv.org/abs/1505.06957.

[CG85] Michael F. Cohen and Donald P. Greenberg. "The Hemi-cube: A Radiosity Solution for Complex Environments". In: *SIGGRAPH Comput. Graph.* 19.3 (July 1985), pp. 31–40. ISSN: 0097-8930. DOI: 10.1145/325165.325171. URL: http://doi.acm.org/10.1145/325165.325171.

[CPC84]   Robert L. Cook, Thomas Porter, and Loren Carpenter. "Distributed Ray Tracing". In: *SIGGRAPH Comput. Graph.* 18.3 (Jan. 1984), pp. 137–145. ISSN: 0097-8930. DOI: 10.1145/964965.808590. URL: http://doi.acm.org/10.1145/964965.808590.

[CS15]   Tay Lee Choo and Ong Hang See. *Illuminance Calibration of A Digital Camera based on Image Colour*. Putrajaya Campus: The 3rd National Graduate Conference (NatGrad2015), Universiti Tenaga Nasional, 2015. ISBN: 978-967-5770-63-0.

[CUTT10]   C. Cuttle. "Towards the third stage of the lighting profession". In: *Lighting Research & Technology* 42.1 (2010), pp. 73–93. DOI: 10.1177/1477153509104013. eprint: https://doi.org/10.1177/1477153509104013. URL: https://doi.org/10.1177/1477153509104013.

[CWH93]   Michael F. Cohen, John Wallace, and Pat Hanrahan. *Radiosity and Realistic Image Synthesis*. San Diego, CA, USA: Academic Press Professional, Inc., 1993. ISBN: 0-12-178270-0.

[DBBS06]   Philip Dutre et al. *Advanced Global Illumination*. AK Peters Ltd, 2006. ISBN: 1568813074.

[DIAL94]   DIAL GmbH: DIALux. *DIAL GmbH: DIALux*. Accessed: 2017-11-16. 1994.

[DMHT18]   Rada Deeb et al. "Interreflections in Computer Vision: A Survey and an Introduction to Spectral Infinite-Bounce Model". In: *Journal of Mathematical Imaging and Vision* 60.5 (June 2018), pp. 661–680. ISSN: 1573-7683. DOI: 10.1007/s10851−017−0781−x. URL: https://doi.org/10.1007/s10851−017−0781−x.

[DRBR+18]   Angela Dai et al. "ScanComplete: Large-Scale Scene Completion and Semantic Segmentation for 3D Scans". In: *Proc. Computer Vision and Pattern Recognition (CVPR), IEEE*. 2018.

[FFPD+18]   Clara Fernandez-Labrador et al. "Panoroom: From the sphere to the 3d layout". In: *arXiv preprint arXiv:1808.09879* (2018).

[FHSM79]   John E Flynn et al. "A guide to methodology procedures for measuring subjective impressions in lighting". In: *Journal of the Illuminating Engineering Society* 8.2 (1979), pp. 95–110.

[FPLG18]   Clara Fernandez-Labrador et al. "Layouts from panoramic images with geometry and deep learning". In: *IEEE Robotics and Automation Letters* 3.4 (2018), pp. 3153–3160.

[GCGL+11]   Felipe Gil-Castineira et al. "Experiences inside the ubiquitous oulu smart city". In: *Computer* 44.6 (2011), pp. 48–55.

[GG10]   Nicolas Gillis and François Glineur. "Using Underapproximations for Sparse Nonnegative Matrix Factorization". In: *Pattern Recogn.* 43.4 (Apr. 2010), pp. 1676–1687. ISSN: 0031-3203. DOI: 10.1016/j.patcog.2009.11.013. URL: http://dx.doi.org/10.1016/j.patcog.2009.11.013.

[GGO06]   A Gligor, H Grif, and S Oltean. "Considerations on an intelligent buildings management system for an optimized energy consumption". In: *Automation, Quality and Testing, Robotics, 2006 IEEE International Conference on*. Vol. 1. IEEE. 2006, pp. 280–284.

[GIFF88]   Robert Gifford. "Light, decor, arousal, comfort and communication". In: *Journal of Environmental Psychology* 8.3 (1988), pp. 177–189.

[GSYS+17]   Marc-André Gardner et al. "Learning to Predict Indoor Illumination from a Single Image". In: (2017). URL: http://arxiv.org/abs/1704.00090.

[GTGB84]   Cindy M. Goral et al. "Modeling the Interaction of Light Between Diffuse Surfaces". In: *SIGGRAPH Comput. Graph.* 18.3 (Jan. 1984), pp. 213–222. ISSN: 0097-8930. DOI: 10.1145/964965.808601. URL: http://doi.acm.org/10.1145/964965.808601.

[GV06]   Anca D Galasiu and Jennifer A Veitch. "Occupant preferences and satisfaction with the luminous environment and control systems in daylit offices: a literature review". In: *Energy and Buildings* 38.7 (2006), pp. 728–742.

[HAMM13]   Hugo Van hamme. "The Diagonalized Newton Algorithm for Nonnegative Matrix Factorization". In: *CoRR* (2013). URL: http://arxiv.org/abs/1301.3389.

[HAYA94]   Hideki Hayakawa. "Photometric stereo under a light source with arbitrary motion". In: *JOSA A* 11.11 (1994), pp. 3079–3089.

[HE14]   Peter D Hiscocks and P Eng. *Measuring Luminance with a Digital Camera*. Tech. rep. Syscomp Electronic Design Limited, 2014.

[HF14]   John F Hughes and James D Foley. *Computer graphics: principles and practice*. Pearson Education, 2014.

[HGDG17]   Kaiming He et al. "Mask r-cnn". In: *Computer Vision (ICCV), 2017 IEEE International Conference on*. IEEE. 2017, pp. 2980–2988.

[HILI16]   VRVIS Research Center: HILITE. *VRVIS Research Center: HILITE*. Accessed: 2018-05-23. 2016.

[HNL05]   K. Hara, K. Nishino, and K. lkeuchi. "Light source position and reflectance estimation from a single view without the distant illumination assumption". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 27.4 (Apr. 2005), pp. 493–505. ISSN: 0162-8828. DOI: 10.1109/TPAMI.2005.82.

[HOJ08]   Toshiya Hachisuka, Shinji Ogaki, and Henrik Wann Jensen. "Progressive Photon Mapping". In: *ACM Trans. Graph.* 27.5 (Dec. 2008), 130:1–130:8. ISSN: 0730-0301. DOI: 10.1145/1409060.1409083. URL: http://doi.acm.org/10.1145/1409060.1409083.

[HTDG+17]   Irtiza Hasan et al. ""Don't Turn Off the Lights": Modelling of Human Light Interaction in Indoor Environments". In: *New Trends in Image Analysis and Processing – ICIAP 2017*. Ed. by Sebastiano Battiato et al. Cham: Springer International Publishing, 2017, pp. 143–151. ISBN: 978-3-319-70742-6.

[HTGD+17]   Irtiza Hasan et al. "Tiny head pose classification by bodily cues". In: *Image Processing (ICIP), 2017 IEEE International Conference on*. IEEE. 2017, pp. 2662–2666.

[HZRS16]   Kaiming He et al. "Deep residual learning for image recognition". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016, pp. 770–778.

[IALD18]   IALD. *International Association of Lighting Designers (IALD)*. Accessed: 2018-09-11. 2018.

[IN06]   Mehlika N Inanici and Mojtaba Navvab. "The virtual lighting laboratory: per-pixel luminance data analysis". In: *Leukos* 3.2 (2006), pp. 89–104.

[JDN19]   Hou Ji, Angela Dai, and Matthias Nießner. "3D-SIS: 3D Semantic Instance Segmentation of RGB-D Scans". In: *Proc. Computer Vision and Pattern Recognition (CVPR), IEEE*. 2019.

[JENS96]   Henrik Wann Jensen. "Global Illumination Using Photon Maps". In: *Proceedings of the Eurographics Workshop on Rendering Techniques '96*. Porto, Portugal: Springer-Verlag, 1996, pp. 21–30. ISBN: 3-211-82883-4. URL: http://dl.acm.org/citation.cfm?id=275458.275461.

[KAJI86]   James T. Kajiya. "The Rendering Equation". In: *SIGGRAPH Comput. Graph.* 20.4 (Aug. 1986), pp. 143–150. ISSN: 0097-8930. DOI: 10.1145/15886.15902. URL: http://doi.acm.org/10.1145/15886.15902.

[KAW15]   Ruzena Kralikova, Miriam Andrejiova, and Emil Wessely. "Energy saving techniques and strategies for illumination in industry". In: *Procedia Engineering* 100 (2015), pp. 187–195.

[KELL97]   Alexander Keller. "Instant Radiosity". In: *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques*. SIGGRAPH '97. New York, NY, USA: ACM Press/Addison-Wesley Publishing Co., 1997, pp. 49–56. ISBN: 0-89791-896-7. DOI: 10.1145/258734.258769. URL: https://doi.org/10.1145/258734.258769.

[KGTM+17] Kihwan Kim et al. "A Lightweight Approach for On-The-Fly Reflectance Estimation". In: *The IEEE International Conference on Computer Vision (ICCV)*. Oct. 2017.

[KH13]   Michael Kazhdan and Hugues Hoppe. "Screened poisson surface reconstruction". In: *ACM Transactions on Graphics (ToG)* 32.3 (2013), p. 29.

[KHFH11]   Kevin Karsch et al. "Rendering Synthetic Objects into Legacy Photographs". In: *ACM Trans. Graph.* 30.6 (Dec. 2011), 157:1–157:12. ISSN: 0730-0301. DOI: 10.1145/2070781.2024191. URL: http://doi.acm.org/10.1145/2070781.2024191.

[KHP14]   Jingu Kim, Yunlong He, and Haesun Park. "Algorithms for nonnegative matrix and tensor factorizations: a unified view based on block coordinate descent framework". In: *Journal of Global Optimization* 58.2 (2014), pp. 285–319. ISSN: 1573-2916. DOI: 10.1007/s10898-013-0035-4. URL: http://dx.doi.org/10.1007/s10898-013-0035-4.

[KKGK+14] Jaroslav Křivánek et al. "Recent Advances in Light Transport Simulation: Some Theory and a Lot of Practice". In: *ACM SIGGRAPH 2014 Courses*. SIGGRAPH '14. Vancouver, Canada: ACM, 2014, 17:1–17:6. ISBN: 978-1-4503-2962-0. DOI: 10.1145/2614028.2615438. URL: http://doi.acm.org/10.1145/2614028.2615438.

[KKSH17]   Mike Kasper et al. "Light Source Estimation with Analytical Path-tracing". In: *CoRR* abs/1701.04101 (2017). URL: http://arxiv.org/abs/1701.04101.

[KLSW17]   Katharina Krösl et al. "LiteMaker: Interactive Luminaire Development using Progressive Photon Tracing and Multi-Resolution Upsampling". In: *Vision, Modeling & Visualization*. Ed. by Matthias Hullin et al. The Eurographics Association, 2017. ISBN: 978-3-03868-049-9. DOI: 10.2312/vmv.20171253.

[KPH16]   Ruzena Kralikova, Miriama Pinosova, and Beata Hricova. "Lighting Quality and its Effects on Productivity and Human Healts". In: *Int. J. Interdiscip. Theory Pract* 10 (2016).

[LA70]   J.H. Lambert and E. Anding. *Photometrie: Photometria, sive De mensura et gradibus luminis, colorum et umbrae (1760)*. Ostwalds Klassiker der exakten Wissenschaften. W. Engelmann, 1870.

[LEN12]   Jean-François Lalonde, Alexei A. Efros, and Srinivasa G. Narasimhan. "Estimating the Natural Illumination Conditions from a Single Outdoor Image". In: *International Journal of Computer Vision* 98.2 (2012), pp. 123–145. DOI: 10.1007/s11263-011-0501-8. URL: http://dx.doi.org/10.1007/s11263-011-0501-8.

[LHHC+13]   Wen-Chieh Lin et al. "Interactive Lighting Design with Hierarchical Light Representation". In: *Proceedings of the Eurographics Symposium on Rendering*. EGSR '13. Zaragoza, Spain: Eurographics Association, 2013, pp. 133–142. DOI: 10.1111/cgf.12159. URL: http://dx.doi.org/10.1111/cgf.12159.

[LHRG10]   Jorge Lopez-Moreno et al. "Compositing images through light source detection". In: *Computers & Graphics* 34.6 (2010). Graphics for Serious GamesComputer Graphics in Spain: a Selection of Papers from {CEIG} 2009Selected Papers from the {SIGGRAPH} Asia Education Program, pp. 698–707. ISSN: 0097-8493. DOI: http://dx.doi.org/10.1016/j.cag.2010.08.004. URL: http://www.sciencedirect.com/science/article/pii/S0097849310001299.

[LIGH90]   INC Lighting Analysts. *Lighting Analysts, INC*. Accessed: 2017-11-16. 1990.

[LMBH+14]   Tsung-Yi Lin et al. "Microsoft coco: Common objects in context". In: *European conference on computer vision*. Springer. 2014, pp. 740–755.

[LN16]   Stephen Lombardi and Ko Nishino. "Reflectance and illumination recovery in the wild". In: *IEEE transactions on pattern analysis and machine intelligence* 38.1 (2016), pp. 129–141.

[LS99]   Daniel D Lee and H Sebastian Seung. "Learning the parts of objects by non-negative matrix factorization". In: *Nature* 401.6755 (1999), pp. 788–791.

[LTHS+13]   Christian Luksch et al. "Fast Light-map Computation with Virtual Polygon Lights". In: *Proceedings of the ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games*. I3D '13. Orlando, Florida: ACM, 2013, pp. 87–94. ISBN: 978-1-4503-1956-0. DOI: 10.1145/2448196.2448210. URL: http://doi.acm.org/10.1145/2448196.2448210.

[LTMS+14]   Christian Luksch et al. "Real-time Rendering of Glossy Materials with Regular Sampling". In: *Vis. Comput.* 30.6-8 (June 2014), pp. 717–727. ISSN: 0178-2789. DOI: 10.1007/s00371-014-0958-x. URL: http://dx.doi.org/10.1007/s00371-014-0958-x.

[MAIT15]   Henri Maitre. *From Photon to Pixel: The Digital Camera Handbook*. 2015.

[MALL88]   TJ Malley. "A shading method for computer generated images". In: *Master's thesis, Dept. of Computer Science, University of Utah* (1988).

[MBK11]   Luc Masset, Olivier Brüls, and Gaëtan Kerschen. *Partition of the circle in cells of equal area and shape*. Tech. rep. Structural Dynamics Research Group, Aerospace and Mechanical Engineering Department, University of Liege, Institut de Mecanique et Genie Civil (B52/3), 2011.

[MG97]   Stephen R Marschner and Donald P Greenberg. "Inverse lighting for photography". In: *Color and Imaging Conference*. Vol. 1997. 1. Society for Imaging Science and Technology. 1997, pp. 262–265.

[MHF13]   Remco Magielse, Bart J Hengeveld, and Joep W Frens. "Designing a light controller for a multi-user lighting environment". In: (2013).

[MS15]   Steve Marschner and Peter Shirley. *Fundamentals of computer graphics*. CRC Press, 2015.

[NL11]   Illuminating Society of North America and IES Testing Procedures Committee. Photometry of Light Sources Subcommittee. *The commissioning process applied to lighting and control systems*. English. "IES DG-29-11." New York, N.Y. Illuminating Engineering Society of North America, 2011. ISBN: 9780879952556.

[NMT12]   Natalia Neverova, Damien Muselet, and Alain Trémeau. "Lighting Estimation in Indoor Environments from Low-Quality Images". In: *Computer Vision – ECCV 2012. Workshops and Demonstrations: Florence, Italy, October 7-13, 2012, Proceedings, Part II*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 380–389. ISBN: 978-3-642-33868-7. DOI: 10.1007/978-3-642-33868-7_38. URL: http://dx.doi.org/10.1007/978-3-642-33868-7_38.

[OIDS15]   Yohei Ogura et al. "Illumination estimation and relighting using an RGB-D camera". In: *VISAPP 2015 - 10th International Conference on Computer Vision Theory and Applications; VISIGRAPP, Proceedings*. Vol. 2. SciTePress, 2015, pp. 305–312. ISBN: 9789897580901.

[PF92]   Pierre Poulin and Alain Fournier. "Lights from Highlights and Shadows". In: *Proceedings of the 1992 Symposium on Interactive 3D Graphics*. I3D '92. Cambridge, Massachusetts, USA: ACM, 1992, pp. 31–38. ISBN: 0-89791-467-8. DOI: 10.1145/147156.147160. URL: http://doi.acm.org/10.1145/147156.147160.

[PL00]   Timo Partonen and Jouko Lönnqvist. "Bright light improves vitality and alleviates distress in healthy people". In: *Journal of Affective disorders* 57.1 (2000), pp. 55–61.

[PLG16]   Alejandro Perez-Yus, Gonzalo Lopez-Nicolas, and Jose J. Guerrero. "Peripheral Expansion of Depth Information via Layout Estimation with Fisheye Camera". In: *Computer Vision – ECCV 2016*. Ed. by Bastian Leibe et al. Cham: Springer International Publishing, 2016, pp. 396–412. ISBN: 978-3-319-46484-8.

[PMAJ13]   Esben Skouboe Poulsen et al. "Responsive lighting: the city becomes alive". In: *Proceedings of the 15th international conference on Human-computer interaction with mobile devices and services*. ACM. 2013, pp. 217–226.

[PP03]   Gustavo Patow and Xavier Pueyo. "A Survey of Inverse Rendering Problems". In: *Computer Graphics Forum* 22.4 (2003), pp. 663–687. DOI: 10.1111/j.1467-8659.2003.00716.x. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-8659.2003.00716.x. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-8659.2003.00716.x.

[PRDR03]   Jo Phipps-Nelson et al. "Daytime exposure to bright light, as compared to dim light, decreases sleepiness and improves psychomotor vigilance performance." In: *Sleep* 26.6 (2003), pp. 695–700.

[PSG01]   Mark W. Powell, Sudeep Sarkar, and Dmitry B. Goldgof. "A Simple Strategy for Calibrating the Geometry of Light Sources". In: *IEEE Trans. Pattern Anal. Mach. Intell.* 23 (2001), pp. 1022–1027.

[REA00]   M. S. REA. "The IESNA Lighting Handbook: Reference & application." In: *The IESNA Lighting Handbook: Reference & application.* Illuminating Engineering Society of North America. New York, NY: Illuminating Engineering Society of North America., 2000.

[RELU10]   Relux Informatik AG: ReluxDesktop. *Relux Informatik AG: ReluxDesktop.* Accessed: 2017-11-16. 2010.

[REND06]   WARD G.: Radiance Renderer. *WARD G.: Radiance Renderer.* Accessed: 2017-11-16. 2006.

[RH01]   Ravi Ramamoorthi and Pat Hanrahan. "A signal-processing framework for inverse rendering". In: *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*. ACM. 2001, pp. 117–128.

[SCK15]   DK Serghides, CK Chatzinikola, and MC Katafygiotou. "Comparative studies of the occupants' behaviour in a university building during winter and summer time". In: *International Journal of Sustainable Energy* 34.8 (2015), pp. 528–551.

[SF08]   G. Strang and G. Fix. *An Analysis of the Finite Element Method*. Wellesley-Cambridge Press, 2008. ISBN: 9780980232707. URL: https://books.google.de/books?id=K5MAOwAACAAJ.

[SKTK12]   Karin CHJ Smolders et al. "Need for recovery in offices: Behavior-based assessment". In: *Journal of Environmental Psychology* 32.2 (2012), pp. 126–134.

[SMAR04]   Paris Smaragdis. "Non-negative Matrix Factor Deconvolution; Extraction of Multiple Sound Sources from Monophonic Inputs". In: *Independent Component Analysis and Blind Signal Separation: Fifth International Conference, ICA 2004, Granada, Spain, September 22-24, 2004. Proceedings*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, pp. 494–499. ISBN: 978-3-540-30110-3. DOI: 10.1007/978-3-540-30110-3_63. URL: http://dx.doi.org/10.1007/978-3-540-30110-3_63.

[SOLS+16]   Johannes Sorger et al. "Litevis: integrated visualization for simulation-based decision support in lighting design". In: *IEEE Transactions on Visualization and Computer Graphics* 22.1 (2016), pp. 290–299.

[SSK17]   M. Sheinin, Y. Schechner, and K. N. Kutulakos. "Computational Imaging on the Electric Grid". In: *IEEE CVPR* (2017).

[SW14]   Michael Schwarz and Peter Wonka. "Procedural Design of Exterior Lighting for Buildings with Complex Constraints". In: *ACM Trans. Graph.* 33.5 (Sept. 2014), 166:1–166:16. ISSN: 0730-0301. DOI: 10.1145/2629573. URL: http://doi.acm.org/10.1145/2629573.

[SWSS+18]   Hiroaki Santo et al. "Light Structure from Pin Motion: Simple and Accurate Point Light Calibration for Physics-based Modeling". In: *The European Conference on Computer Vision (ECCV)*. Sept. 2018.

[TAK09]   P. TAK. "Lighting Design and Energy Savings." In: *Light, FCC Public s. r. o., Prague* Vol.10, No.2. (2009).

[TDRH+16]   Jiandong Tian et al. "Simple and effective calculations about spectral power distributions of outdoor light sources for computer vision". In: *Optics express* 24.7 (2016), pp. 7266–7286.

[TMNM09]   Takeshi Takai et al. "Difference Sphere: An Approach to Near Light Source Estimation". In: *Comput. Vis. Image Underst.* (Sept. 2009). ISSN: 1077-3142. DOI: 10.1016/j.cviu.2009.03.017. URL: http://dx.doi.org/10.1016/j.cviu.2009.03.017.

[TS74]   Lyle H Taylor and Eugene W Socov. "The movement of people toward lights". In: *Journal of the Illuminating Engineering Society* 3.3 (1974), pp. 237–241.

[TW13]   Peter Tregenza and Michael Wilson. *Daylighting: architecture and lighting design*. Routledge, 2013.

[VAN 89]   J Van Leersum. "A method for determining a consistent set of radiation view factors from a set generated by a nonexact method". In: *International journal of heat and fluid flow* 10.1 (1989), pp. 83–85.

[VAVA07]   Stephen A. Vavasis. "On the complexity of nonnegative matrix factorization". In: *CoRR* abs/0708.4149 (2007). URL: http://arxiv.org/abs/0708.4149.

[VEIT00]   JA Veitch. "Lighting guidelines from lighting quality research". In: (2000).

[VG95]   Eric Veach and Leonidas J Guibas. "Optimally combining sampling techniques for Monte Carlo rendering". In: *Proceedings of the 22nd annual conference on Computer graphics and interactive techniques*. ACM. 1995, pp. 419–428.

[WBSS04]   Zhou Wang et al. "Image quality assessment: from error visibility to structural similarity". In: *IEEE transactions on image processing* 13.4 (2004), pp. 600–612.

[WS03]   Yang Wang and Dimitris Samaras. "Estimation of Multiple Directional Light Sources for Synthesis of Augmented Reality Images". In: *Graph. Models* 65.4 (July 2003), pp. 185–205. ISSN: 1524-0703. DOI: 10.1016/S1524-0703(03)00043-2. URL: http://dx.doi.org/10.1016/S1524-0703(03)00043-2.

[XL13]   Alison Jing Xu and Aparna Labroo. "Incandescent affect: Turning on the hot emotional system with bright light". In: *ACR North American Advances* (2013).

[XLG03]   Wei Xu, Xin Liu, and Yihong Gong. "Document Clustering Based on Nonnegative Matrix Factorization". In: *Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Informaion Retrieval*. SIGIR '03. Toronto, Canada: ACM, 2003, pp. 267–273. ISBN: 1-58113-646-3. DOI: 10.1145/860435.860485. URL: http://doi.acm.org/10.1145/860435.860485.

[ZCSH18]   Chuhang Zou et al. "LayoutNet: Reconstructing the 3D Room Layout from a Single RGB Image". In: *CoRR* abs/1803.08999 (2018). arXiv: 1803.08999. URL: http://arxiv.org/abs/1803.08999.

[ZYHR15]   Xin Zhou et al. "Data analysis and stochastic modeling of lighting energy use in large office buildings in China". In: *Energy and Buildings* 86 (2015), pp. 275–287.