



The Society of
Light and Lighting

An integrated light management system with real-time light measurement and human perception

T Tsesmelis PhD^{abc} , I Hasan PhD^{abc}, M Cristani Prof^c, A Del Bue PhD^b and F Galasso Prof^d 

^aVision Laboratory, Department of Innovation and Innoventures, OSRAM Licht AG, Garching, Germany

^bVisual Geometry and Modelling Laboratory, Istituto Italiano di Tecnologia, Genova, Italy

^cVision, Image Processing and Sound Laboratory, Department of Computer Science, University of Verona, Verona, Italy

^dPerception and Intelligence Laboratory, Department of Computer Science, Sapienza University of Rome, Rome, Italy

Received 8 December 2019; Revised 8 April 2020; Accepted 10 July 2020

Illumination is important for well-being, productivity and safety across several environments, including offices, retail shops and industrial warehouses. Current techniques for setting up lighting require extensive and expert support and need to be repeated if the scene changes. Here we propose the first fully automated light management system which measures lighting in real-time, leveraging an RGBD sensor and a radiosity-based light propagation model. Thanks to the integration of light distribution and perception curves into the radiosity, we outperform a commercial software (Relux) on a newly introduced dataset. Furthermore, our proposed light management system is the first to estimate both the presence and the attention of the people in the environment, as well as their light perception. Our new light management system adapts lighting to the scene and human activity and it is capable of energy savings of up to 66%, as we experimentally quantify, without compromising the lighting quality.

1. Introduction

Light is indispensable in our perception of the world and it affects our emotional and physiological responses.^{1,2} Well-lit workplaces provide visual comfort and improve productivity.^{3,4} However, lighting may reach 15% of the overall building electricity consumption, with peaks above 25%.⁵ To save

energy, a light management system (LMS) would need to measure lighting and to reason on the human perception of it in real-time, e.g. to dim down lighting where none sees it.

Current LMSs cannot automatically obtain dense measures of the spatial illumination in real-time. The former task is either accomplished manually with luxmeters (point-to-point measurements) or via offline CAD-based simulations. These require an operator to visit the scene and perform the measurements, or a detailed CAD model of the 3D scene, respectively. Both approaches need to be repeated upon any scene change and are impractical in the flexible open-plan

Address for correspondence: Theodore Tsesmelis, Visual Geometry and Modelling (VGM) Laboratory, Pattern Analysis and Computer Vision (PAVIS) Department, Istituto Italiano di Tecnologia (IIT), Via Enrico Melen 83, 16152 Genova, Italy.

E-mail: theodore.tsesmelis@iit.it

spaces of modern offices, designed to adapt to the daily work plan.

Modern LMSs cannot estimate the human light perception. Current implementations rely on daylight harvesting and occupancy sensing. The first adjusts the luminaires to maximize the use of daylight, when available.⁶ The second leverages thermal- or radar-based motion detectors⁷ to switch on/off all luminaires in the room when people enter/leave, no matter the size of the (open-plan) office.⁸

Here we introduce a new LMS which estimates both the scene lighting and the human perception of it, to optimize for the quality of illumination while saving energy. Light estimation leverages an RGBD camera and a radiosity model for light propagation, and distinguishes the scene 3D structure, the object reflectance and light positions. For the human perception of light, we locate people in the environment, estimate their visual frustum of attention (VFOA) and the incident light onto their VFOA, depending on their position and head pose. We also introduce a new labelled dataset, featuring a number of lit rooms, with and without human activity. We provide RGBD images and 3D meshes, labelled with material reflectance properties; luminaire positions, characteristics and dimming level; person locations and VFOA. Light intensity and human light perception are captured by illuminance meters, placed in the scene or worn by people, respectively, to provide ground truth measurements.

Finally, we introduce a new end-to-end system architecture and implement the autonomous system that we call the ‘Invisible Light Switch’ (ILS). ILS encompasses an RGBD camera, a processor, a light controller, a communication bus and the luminaires. Since the ILS estimates how much light each person receives, it may switch off or dim those luminaires which are not visible, e.g. on the other side of large open spaces or behind cubicle panels. This removes the need for manual switches and provides a

boost in energy efficiency, saving up to 66% without compromising the light quality.

Our main contributions are: 1. We propose a real-time light estimation process from an RGBD sensor as well as its perception by the scene occupants; 2. We collect a new benchmark for quantitative evaluation; 3. We propose an end-to-end autonomous LMS, to control lighting and save energy. This manuscript brings together two previous conference publications on the topics of light estimation and control,^{9,10} and extends the work with *a*) novel dataset and annotations, *b*) more experiments and *c*) the definition and implementation of the overall end-to-end system architecture.

We present the system and modelling in Section 3; we discuss the dataset and experiments in Section 4; finally, Section 5 concludes the manuscript. Next, we review related work.

2. Related work

Light measurement and management encompass different fields in science and engineering, which we review here.

2.1 Light measurement

According to Cuttle¹¹ the current status of lighting profession and lighting evaluation is on luminance-based light assessing procedures, i.e. estimating the light arriving at the eyes of a virtual observer. However, in most previous work light measurement and modelling refer mainly to research in image and visual computing. As we review in Tsesmelis *et al.*,¹⁰ previous research regarded image spatio-temporal and pixel-like approaches,^{12,13} and the creation of photorealistic renderings¹⁴ but not the actual spatial lighting measurement. By contrast, commercial light modelling software, e.g. Relux, DIALux, AGi32, focuses on measuring light. They mainly target offline measurements and the evaluation of lighting solutions in a simulated environment. We build our work on Tsesmelis *et al.*,¹⁰ adopting

computer vision techniques to measure lighting in real-time by means of a camera.

2.2 Light management systems

LMSs play a crucial role in reducing energy consumption in offices¹⁵ and prior work mainly addressed daylight harvesting¹⁶ and occupancy sensing systems (OSS).^{17,18} In particular, Ul Haq *et al.*¹⁵ emphasized the potential of OSSs for energy saving, surveying available technology and the performance compromise; Guo *et al.*⁷ additionally focused on the implementation of occupancy sensors; Pandharipande and Caicedo¹⁶ specifically considered an OSS based on ultrasonic and proposed improvements in user localization based on time-difference-of-arrival and transmission over multiple time slots. However, none of the above addressed energy saving, nor distinguished among private, single-user and open-plan offices as one would intuitively argue for.

Among related work considering open-plan offices, Kim and De Dear¹⁷ determined three layout categories: (a) cubicle layout with high partitions (~1.5 m or higher), (b) cubicle layout with low partitions (up to ~1.5 m) and (c) open layout with no or limited partitions. Wang *et al.*¹⁸ and Pandharipande and Caicedo¹⁹ proposed and demonstrated by simulation an optimization algorithm to compromise energy efficiency and user-comfort, by considering user presence and daylight conditions. However, they did not address any occupancy sensing algorithm, nor a real-time light propagation model, which we propose here for a real end-to-end light management system. Finally, the lack of a proper review study on LMSs in open-plan offices has been recently stressed by De Bakker *et al.*⁸ They envisioned modern LMSs, capable of dimming light according to how much of it is perceived by each person in the scene, by means of their visual frustum of attention. We realize this vision here with our 'Invisible Light Switch'.

3. Proposed end-to-end system

The proposed end-to-end system encompasses three functional tasks: the light-centric scene understanding, the human-centric scene understanding and the light management. These are depicted in Figure 1.

Light-centric scene understanding is responsible to estimate the scene 3D geometry and the light propagation within the scene. For the light propagation, we adopt and extend the radiosity model to additionally model the actual light source (office luminaires) as well as the receiving sensor (luxmeters or viewing people).

Human-centric scene understanding stands for detecting the people in the scene and estimating their view frustum of attention (VFOA). In other words, the given scene is understood in terms of how much light (from each light source) a person perceives in each part of it and from each viewpoint, whereby the latter is estimated via the VFOA.

Both the light-centric and human-centric scene understanding modules are necessary when it comes to open-plan offices, where multiple light sources create complex light patterns which act on moving employees. Finally, light management stands for automatically adjusting lighting in response to the people position and attention. This includes dimming or switching the luminaires off when people do not see them, as in our ILS system. The first two are sensing tasks, performed in first place by the sole use of an RGBD camera. The third task is about light control and leverages an established light-communication BUS (digital addressable lighting interface (DALI)), to read statuses and send commands.

3.1 Light-centric scene understanding

The ILS targets the estimation of the scene 6-DoF illumination, i.e. anywhere in the 3D space and from any 3D direction, by the sole use of RGBD images. Towards this goal, we define a procedure to retrieve the scene 3D

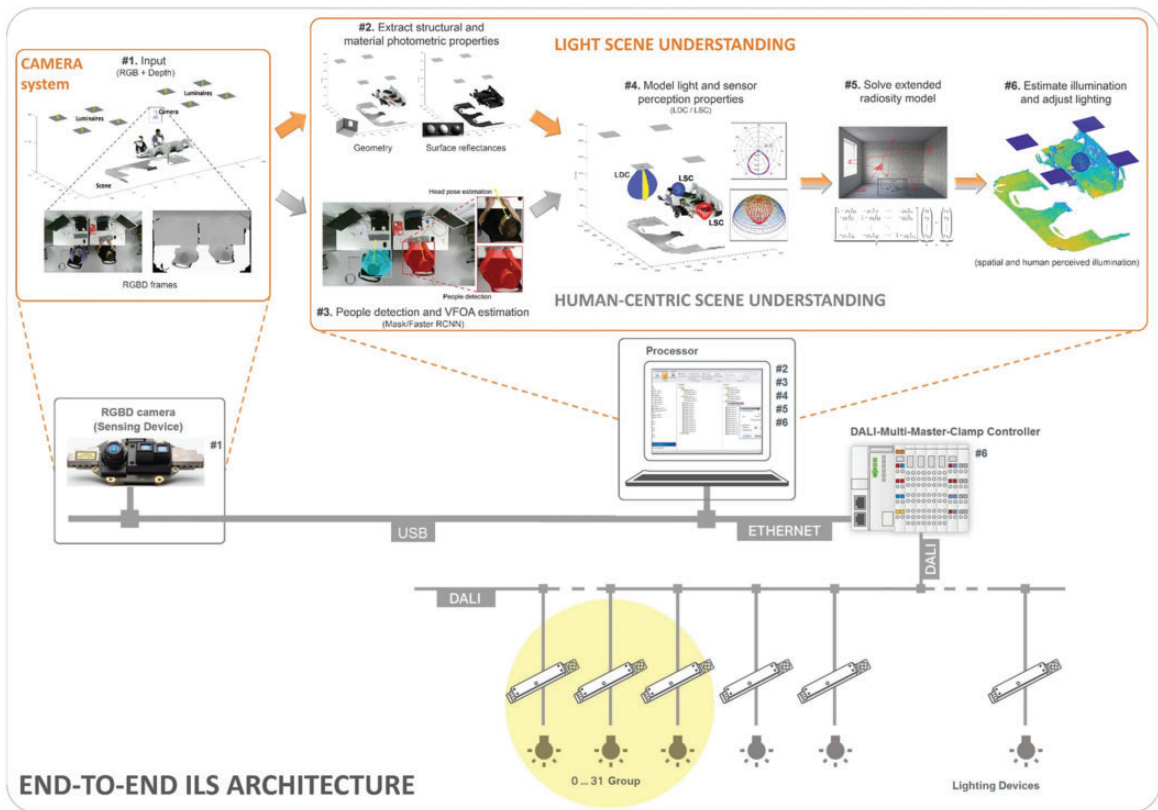


Figure 1 Overview of the invisible light switch (ILS). The architecture includes an RGBD sensor, installed in the room ceiling, which acquires top-view colour and depth images, as detailed in the camera system box (step #1 in the pipeline). The input RGBD images are processed in the processor. There it extracts the photometric properties of the surfaces and the geometric structure of the scene by leveraging the RGB and depth image pairs (step #2) while in parallel it detects the people and estimates the VFOA of each of them (Step #3). Thereafter, it models the actual light sources (luminaires) and light-perceiving devices (luxmeters or people) by using the corresponding distribution curves LDC/LSC (step #4). Finally, it solves for the extended radiosity model (Step #5) and estimates the spatial and human-perceived illumination over the visible 3D scene (Step #6). Based on the light estimation, we steer lighting in the room via the WAGO DALI ConFigureur software. This encodes luminaire commands to a DALI Master Light Controller via TCP/IP, which it relays to the luminaires via the DALI BUS. VFOA: visual frustum of attention; LDC: light distribution curve; LSC: light sensitivity curve; DALI: digital addressable lighting interface (available in colour in online version).

surface from the point cloud, as well as the surface reflectance. We then use and extend radiosity to model the light interaction among 3D scene parts by introducing the distribution characteristics of the actual lights (luminaires) and sensing elements (the illuminance meters or the people in the scene) c.f. illustration in Figure 1.

3.1.1. From RGBD images to surface and reflectance

Light propagation requires a scene described by patches, i.e. 3D facets characterized by a surface and an orientation. But a RGBD camera only provides a sparse and noisy point cloud. We set to recover the scene surfaces by first denoising (weighted median

and bilateral filtering) the point cloud by means of the colour and depth image pairs. Then we reconstruct the surface as a 3D mesh of multiple scattered patches by using the Open3D library. Given the surface normals, ILS estimates the albedo ρ of each surface element by the first-order spherical-harmonics method, the same as the one adapted in Tsesmelis *et al.*¹⁰ This requires multiple images of each surface patch under different illumination conditions (e.g. alternating the lights from each of the scene luminaires). We attain the image-set by time-lapse recordings and by selecting/synthesizing single-light-source-lit images with the method in Tsesmelis *et al.*²⁰

3.1.2 The Radiosity model

Given the scene 3D surface and the corresponding reflectance, ILS estimates the illuminance of each surface patch with the radiosity model. Radiosity is adopted by most commercial light simulation software (Relux, DIALux, AGi32) because it describes the physical light propagation phenomenon and provides light estimations which are close to those measured by illuminance meters, in lux.

Radiosity is a simple linear model

$$r_i = e_i + \rho_i \sum_{j=1}^{n-1} f_{ij} r_j \quad (1)$$

where i and j index the n patches of the room 3D mesh, 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 , f_{ij} is the form factor between patch i and j and r_i is the estimated radiance value on surface i related with the actual light intensity accordingly.

The form factors f_{ij} are defined as the fraction of energy reaching one patch j over the emitted by patch i , and in practice is a

square matrix $n \times n$. These describe the scene geometry and encode two main aspects:

- **Visibility:** whether two patches are visible from each other, being 0 if there is no line of sight between them;
- **Distance and orientation:** how well two patches see each other. Small values correspond to far away patches with an oblique line of sight, while large values mean close, fronto-parallel patches.

In this work we compute the form factors by adapting the ‘Isocell’ technique,²¹ (c.f. Figure 2), due to its higher precision at a given computation time (compared to analytic alternatives)

$$f_{ij} = \frac{\mu_j}{\mu_i} \quad (2)$$

where μ_j is the number of rays reaching patch j over the total number of rays μ_i emitted by facet i .

3.1.3 Light distribution and sensitivity curves

The radiosity model in equation (1) has two main limitations: it only models point light sources and it disregards the light perception. The latter limitation affects all current commercial light planning software. We address both aspects by extending the radiosity model and introducing the light distribution (for considering any non-ideal light source type) and the light sensitivity (to model the light observer / sensor perception) curves respectively.

Light distribution curve (LDC): The radiosity formulation, equation (1), assumes isotropic light sources, i.e. active patches radiate with the same luminous intensity in all directions. However, isotropy is hardly the case for real light sources which are normally represented by a radiation map, such as in Figure 2(a) (solid line graph). This distribution is in general non-linear and encoded as a

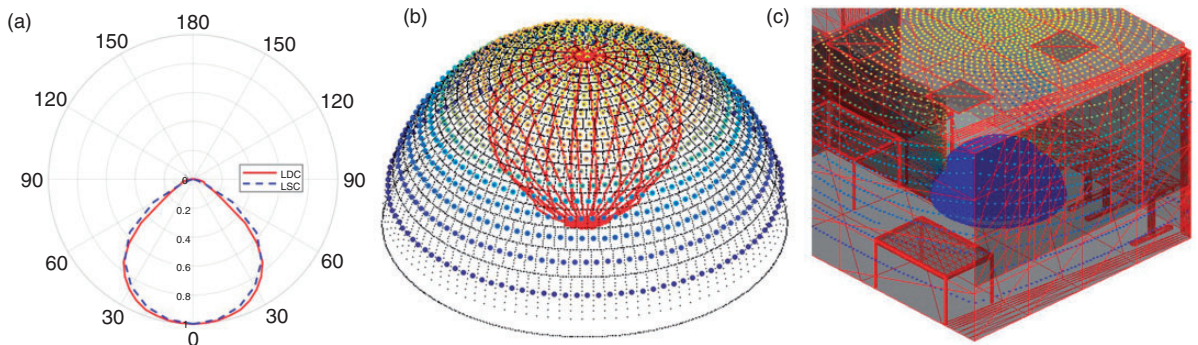


Figure 2 (a) Radial representation of LDC and LSC curves, illustrating how both quantities are attenuated (radius) with the growing light incident angle (radial angle); (b) weighted ray distribution of the Isocell unit sphere for LDC, dots gradient follows the curvature of the distribution; (c) Isocell unit sphere (blue hemisphere) for a patch on the floor of the office CAD. The gradient-dots in (c) are taken from (b) but placed at the intersection of the cast ray with the other scene patches. LDC: light distribution curve; LSC: light sensitivity curve (available in colour in online version).

curve which provides the radiant intensity with respect to the distance and angle from the emitting source (i.e. how much light is emitted in each direction). The LDC curves were obtained experimentally in laboratory environment conditions with the use of specific equipment, i.e. goniophotometer, and provided in datasheets by the luminaire manufacturer. We include the LDC curve into the radiosity model by encoding the non-linear radiation in the form factors f_{ij} . In more detail, we associate scalar values (i.e. weights) to each cast ray leaving a light source patch, proportionally to the angle of emission, as illustrated in Figure 2(b). These are then used to re-write equation (2) as a weighted mean. By encoding the non-linearity of the LDC curve into the form factors computation, our re-formulation of radiosity preserves the linearity thus having a minimal impact on the overall computation.

Light sensitivity curve (LSC): In a similar fashion as with light sources, the light-measuring sensors have different sensitivity to lighting, depending on the lighting angle and distance as well as to manufacturing characteristics. The LSC plot in Figure 2(a) (dotted line graph) illustrates the perception

characteristic of the illuminance meter sensor, by which we measure the spatial and human perceived illumination in the gathered data. As for LDC, we integrate the angular weighting of the cast rays by altering the cast rays, c.f. Figure 2(b). Also, this extension maintains the model linearity and the fast computation time.

3.2 Human-centric scene understanding

The main human subjective features of a light management system are the localization of people, the estimation of their visual attention and the consequent estimation of their lighting perception in the scene.

We cast the localization of people as a human detection task from top-view imagery based on the Mask R-CNN model with the ResNet-101 backbone.

We denote the visual attention of people by their visual frustum of attention (VFOA), which we estimate with the model of Hasan *et al.*²² This consists of a Faster R-CNN architecture with VGG16 backbone, extended by a VFOA branch for gaze estimation, taking as input the detected person bounding-box. As noted in Hasan *et al.*,²² the whole-person bounding box provides an

important contextual cue, to complement the tiny heads from the top-view imagery.

The VFOA is cast as classifying the person viewing angle into quantized direction bins. We experimented with 4-quantized viewing directions (North, West, South, East), as well as with eight (yielding a granularity of 45°). Moreover, attempting to estimate the VFOA by regression, underperformed the classification approach, based on the fact that it is easier to estimate a class label than the exact angle. Once we extract the detected 2D people positions and viewing angles, we map them onto the 3D space by means of depth to 3D mapping.

Finally, we model the light perceived by the people as the illumination reaching the person's field-of-view (FOV), as described by the 6-DoF head position and VFOA orientation in the scene. Each FOV is assumed to be a conic reception field. The arriving light follows the radiosity model of Section 3.1, the ray-casting simulation as described for the illuminance meter sensitivity curve (LSC) in Tsesmelis *et al.*,⁹ and the consequent integration across the human FOV. Overall, each person's light perception is approximated as a light sensor like an illuminance meter, positioned between the eyes.

3.3 Light management system

The ILS is a camera-aided smart LMS to control lighting in response to the currently spatial and human perceived estimated light as well as the people position and attention. The ILS reads the luminaire status and switches them on and off based on the peoples' presence. Additionally, it dims luminaires down when partially visible by the people, therefore saving energy '*in the invisible*', while maintaining the desirable scene illumination.

As illustrated in Figure 1, the proposed LMS consists of the sensing RGBD camera, a computing device to estimate the light and human factors in the scene (c.f. Sections 3.1.

and 3.2.), a Master Light Controller to communicate commands to luminaires, and the luminaires themselves, interconnected via a suitable protocol BUS. The ILS is implemented as a computer program running on the computing device, based on the peoples' presence and attention, and on the luminaire status readouts. Overall, the proposed LMS system is autonomous, end-to-end and real-time.

We report results based on images acquired via a Kinect v2 RGBD camera. The computing device is a laptop, running the sensing algorithms given in the sections on Light-centric scene understanding, Human-centric scene understanding and the controlling invisible light switch program. The Kinect v2 is connected to a laptop via a USB2 port.

The Master Light Controller interfaces the computer with the luminaires, by forwarding switching and dimming commands and reporting the luminaire statuses (on, off, dimming level etc.). We adopt the WAGO-I/O-SYSTEM 750 through a DALI Multi-Master Module 753-647.²³ As with most such controllers, it connects to the computer via Ethernet TCP/IP, enabling therefore IoT and cloud-based intelligence. The connection is accommodated via proprietary software, i.e. the WAGO DALI Configurator,²⁴ running on the computer. The WAGO DALI Configurator allows easy commissioning of the devices connected on the DALI network. This includes the offline configuration of the entire DALI network, including the electromagnetic control gears (ECGs), the sensors and the saving/repeating device configurations.

As a communication BUS between the Master Light Controller and the luminaires, ILS adopts the established DALI BUS. DALI is one of the simplest duplex-communication protocols, which allows to flexibly connect up to 64 devices in series, grouping them into up to 16 clusters.

4. Experiments

Here we evaluate the performance of the LMS system and its parts on a novel benchmark.

4.1 Dataset and metrics

In addition to the data of Tsesmelis *et al.*,^{9,10} we provide a dataset with RGBD images and light measurements for three new and diverse scenes with various new artificial-based (natural daylight is out of scope of this research) illumination scenarios (combinations of on/off light sources for each scene). The new dataset includes offices, meeting rooms and resting areas as depicted in Figure 3. For two of the scenes (Figure 3(b,c)), detailed 3D CAD models are available, as well as an accurate labelling of the object textures and reflectance properties. This allows the comparison to Relux, a state-of-the-art commercial software for light modelling. Moreover, for two other scenes, we provide annotated human activity from a group of two occupants (watching TV, working at the desk, chatting etc.), detailing the people's position, their visual frustum of attention (VFOA) – (Figure 3(g,h)) – as well as the illumination reaching their sight (i.e. within in their VFOA). This allows us to benchmark lighting efficiency, in relation to human perception.

For each frame, the ground truth illumination maps are collected by a number (8–11) of synchronized illuminance meters (depicted in Figure 3 as red boxes), to measure the light intensity reaching the specific spot, in lux. We measure the light estimation quality by the absolute error in lux, compared to the corresponding illuminance meter readings.

To assess the quality of the human-centric system, we benchmark people detection and (quantized) pose estimation accuracy by the established metrics of mean average precision (mAP) and classification accuracy. Both measures are quantified on the external

larger dataset of Demirkus *et al.*,²⁵ where people and their head pose are labelled. On the other hand, we assess the accuracy in estimating the perceived illumination by each person with additional illuminance meters, worn by the users on their forehead (see green boxes in Figure 3(g,h)). This is also quantified by the absolute errors in lux.

4.2 Ambient light estimation

Table 1 first compares our proposed light estimation approach ('Ours w/o CAD') to the commercial software Relux, within the scenes of Rooms 1 and 2. Given the CAD model, our method ('Ours w/ CAD') outperforms the Relux software ('Relux w/ CAD') on both rooms. We achieve an average error of 36 lx and 70 lx for Rooms 1 and 2 across the installed illuminance meter sensors ('Avg. 1–8'), versus the Relux errors of 63 lx and 84 lx, respectively. Note also that these errors occur over an illuminance range of [0, 2000] lx. In both cases, these are good estimates for commissioning, since errors below 200 lx are generally acceptable in the lighting industry.

Table 1 also contains ablation studies 'w/ CAD', given by removing the LSC and LDC distribution curves ('no _LDC LSC', 'no LDC', 'no LSC'). Results confirm that both LDC and LSC are key features for best performance. It is of interest that LSC yields a larger error reduction than the LDC for both rooms. In our view, this happens because LSC properly considers the angle of impacting rays, especially down weighting the rays coming to the sensor from the side, reflecting the sensor absorption characteristic. When only using the RGBD image (i.e. 'Ours w/o CAD'), we may only estimate light propagation to those scene parts which are within the camera FOV. Since the camera is placed around the room centre, we leave illuminance meters 1 and 8 at the scene corners out (c.f. Figure 3(a)), and only provide the average errors for illuminance meters 2–7 (61 lx and 99 lx for Rooms 1 and 2, respectively). This

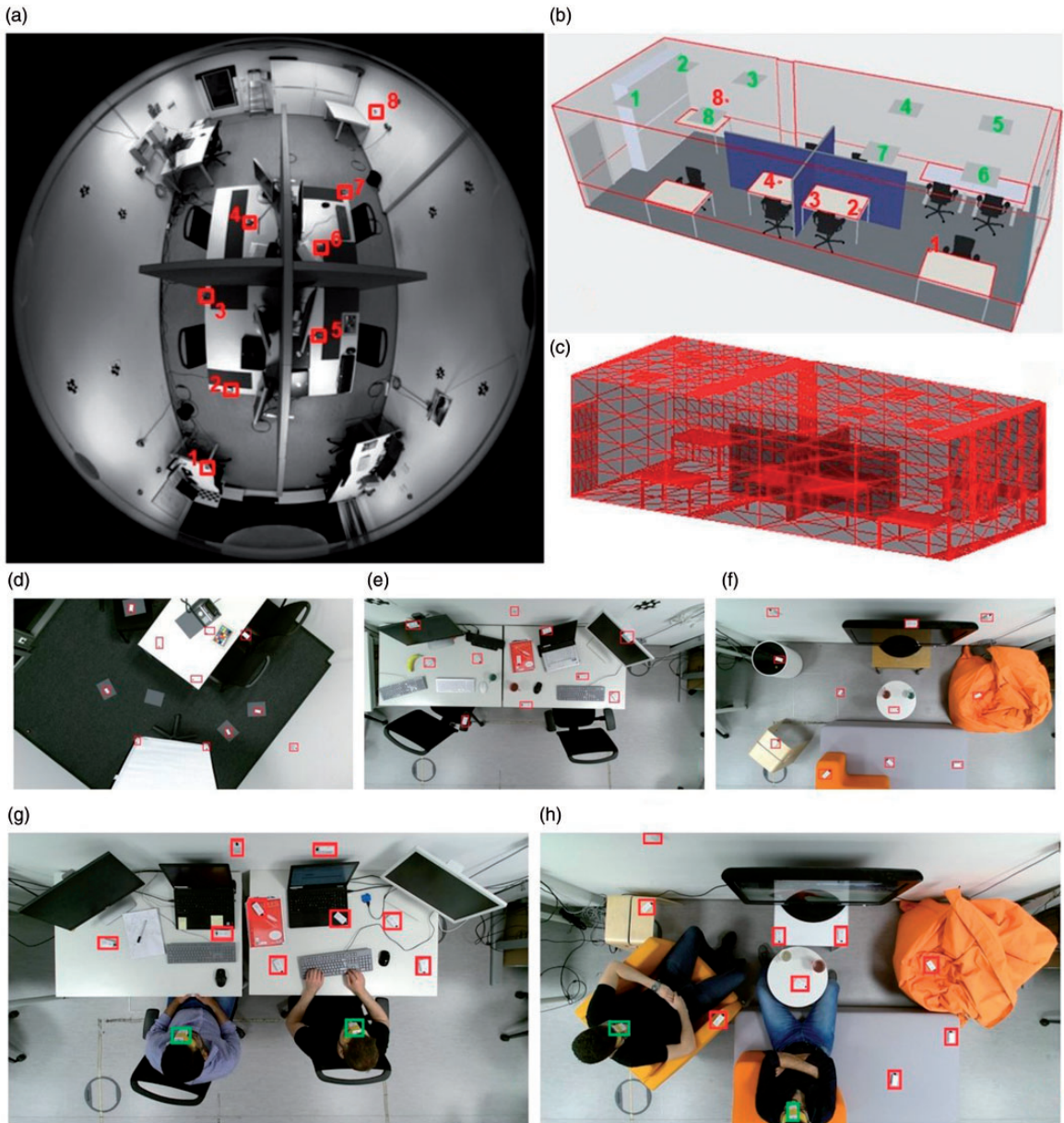


Figure 3 Illustration of the five dataset scenarios (5 rooms). (a) shows Room 1, from the top-view camera; (b) and (c) are detailed CAD models of Room 1, illustrating the luminaire positions (numbered boxes on the ceiling) and the scene subdivision into patches (c.f. light-centric scene understanding), respectively. Room 2 (not shown) is similar to Room 1, but it has no panels across the central four desks. (d), (e) and (f) Depict Rooms 3, 4, 5, also used to evaluate the illumination estimation quality (note the marked boxes within the scene, highlighting the positions of the illuminance meters for light measurement). (g) and (h) Illustrate Rooms 4 and 5 additionally featuring human activity (note the marked boxes on the occupants' forehead, indicating the worn illuminance meters) (available in colour in online version).

Table 1 Comparative evaluation of our approach w/ and w/o CAD including the light distribution curve (LDC) and luxmeter sensitivity curve (LSC) vs the Relux commercial software, alongside all the different ablation studies referred to in Section 4 for two of the evaluated rooms. The values represent the average absolute spatial light estimated error in lux in comparison to the ground truth values collected from the total eight illuminance meter sensors, over the different applied illumination scenarios. ‘Avg. 1–8’ and ‘Avg. 2–7’ show the average value based on the estimated error across the denoted illuminance meter sensors accordingly

Ablation Studies	Room 1 Avg. spatial light estimation error in lux										Room 2 Avg. light estimation error in lux									
	Luxmeters										Luxmeters									
	1	2	3	4	5	6	7	8	Avg. 1–8	Avg. 2–7	1	2	3	4	5	6	7	8	Avg. 1–8	Avg. 2–7
Relux w/ CAD	167	96	27	26	43	10	96	39	63	50	206	97	27	80	97	49	73	44	84	71
Ours w/ CAD	69	24	22	38	28	28	38	41	36	30	70	57	76	106	75	69	55	53	70	73
Ours w/ CAD (no LDC/LSC)	188	150	33	45	43	34	91	65	81	66	207	114	99	148	105	117	93	81	120	112
Ours w/ CAD (no LSC)	199	152	29	41	40	33	95	57	81	65	213	117	82	125	97	97	86	63	110	100
Ours w/ CAD (no LDC)	73	45	24	32	40	34	46	52	43	37	69	80	98	136	70	84	56	62	82	87
Ours w/ CAD (within camera FOV)	–	64	28	20	17	22	52	–	–	34	–	54	36	59	101	69	54	–	–	62
Ours w/o CAD (within camera FOV)	–	53	41	67	68	40	98	–	–	61	–	98	90	85	136	108	77	–	–	99

FOV: field-of-view

case is clearly more challenging, partly because of the camera FOV, and partly because the room reconstruction is effectively a noisy 2.5D. The camera FOV only covers ~40% of the actual scene, as it misses the ceiling and most of the walls. Moreover, the camera-based scene reconstruction only allows a noisy estimation of those surfaces which are visible from the camera viewpoint. We set to discern which factor matters the most by the ablation method ‘Ours w/ CAD (within camera FOV)’, which considers a manual input CAD, cropped according to the camera FOV. As shown in Table 1, this performs closely to ‘Ours w/ CAD’ (the slightly lower error for Room 2 seems due to the challenging scene with entangled lights, where removing walls reduces some of the possibly misestimated light rays). We conclude that most of the added challenge comes from the noise in the point-cloud and in the effective 2.5D estimate, resulting in a geometry with holes. However, in contrast to all

other techniques in the literature, the proposed ‘Ours w/o CAD’ is the sole fully-automatic one.

The left section of Table 2 illustrates similar performance of our fully automatic technique on the Rooms 3, 4 and 5, depicted in Figure 3. Additionally, we introduce two sets of dynamic experiments in Rooms 4 and 5. By dynamic we mean the inclusion of people activities with different VFOAs and interactions within the scenes. As it can be seen results are in line with previous experiments.

4.3 People detection and head-pose estimation

We attain best person detection performance (higher detection accuracy in terms of mean Average Precision (mAP))²⁶ by pre-training the detector on the large MS-COCO dataset (80 k training + 35 k validation images) and then fine-tuning on a selection of diverse top-view images from Demirkus *et al.*²⁵ Following Demirkus *et al.*,²⁵ we select

Table 2 Values on the left present the spatial light estimation errors across Rooms 3–5, obtained with the proposed approach (w/o CAD). Dynamic scenes are additional image sets, including people, activities and scene interactions (see Section 4 for details). Values on the right side of the table are the error in estimating the light perception of people with our proposed method, compared to the ground-truthed values provided by head-worn illuminance meters. We also report error values for the cases where the head pose is provided (oracle)

	Avg. spatial light estimation error in lux												Avg. human light perception error in lux		
	Illuminance meters												Illuminance meters (head-worn)		
	1	2	3	4	5	6	7	8	9	10	11	Avg.	1	2	Avg.
Static Scene															
Room 3	70	93	69	23	25	28	59	28	49	71	82	54	–	–	–
Room 4	18	23	41	26	76	23	35	34	40	69	31	38	–	–	–
Room 5	35	38	23	27	29	49	23	40	33	25	34	32	–	–	–
Dynamic Scene															
Room 4	62	26	68	65	48	57	44	30	28	–	–	48	Est. head pose	216	166
(dynamic)													Oracle head pose	98	92
Room 5	35	34	44	20	32	40	24	28	27	–	–	31	Est. head pose	55	152
(dynamic)													Oracle head pose	42	69

diverse frames by sampling one of every 20 frames, which yields 4459 training + 1736 test images. On this test set, we achieve 98% mAP (IoU = 0.5).

For estimating the VFOA, we train the tiny head-pose estimator of Hasan *et al.*²² on the above-named selection of training images from Demirkus *et al.*,²⁵ after having labelled all datasets. We test on head-pose angles quantized into 4 and 8 classes and obtained accuracy of 70.7% and 43.2%, respectively, on our test set. The degraded performance in the 8-class case is justified due to adjacent viewing angle confusion. This can be further explained by the tiny head region of people in the images, of just 40×50 pixels. Thus, we adopted the 4-quantized-angle head-pose estimator.

4.4 Human-centric light estimation

On the right part of Table 2, we illustrate the error of our method in estimating the illuminance arriving at the peoples' sight. We assume as ground truth the illumination estimation of an illuminance meter which the occupants wear on their forehead. Average errors for Rooms 4 and 5 are 191 lx and 104 lx, respectively. These values are

relatively large, compared to the average light estimation errors reported for the spatial light estimation. We justify this due to the challenges in the light propagation as well as to the error in the head-pose estimation. We therefore set to estimate errors when the head pose is given by an oracle.

Light estimation errors for the case of an 'oracle' head-pose estimator are also reported in the same section of Table 2 and are significantly lower ($\sim 50\%$). This shows that there is still much progress needed in head-pose estimation. Still, the residual error remains larger in average, compared to the scene illumination estimates (95 lx vs. 48 lx for Room 4; 55 lx vs. 31 lx for Room 5). We explain the discrepancy by two main factors: a) the reconstructed 3D mesh is less accurate for the peoples' heads than on the desks; and b) the light estimation arriving at the people suffers from the limited FOV of the depth sensor, since it excludes parts of walls, an important factor for the head-worn illuminance meter facing them.

4.5 Human-centric light management system

Table 3 evaluates whether we can effectively adopt our proposed end-to-end system

Table 3 Quantitative evaluation of how much power may be saved by the invisible light switch system by switching off luminaires that are not directly affecting the perception of light by the occupants, in Rooms 4 and 5. Δ_{watt} and Δ_{lux} quantify how much power can be saved in each setup of active luminaires and the perceived illumination in sight changes, respectively. The light estimation error w.r.t. to the ground truth values is also reported for the selected illumination scenarios

Active luminaires (out of 8 available, 1 2 3 4 5 6 7 8)			Room 4 (dynamic)			Room 5 (dynamic)		
			3 4 7 8	2 3 4 5	3 4	1 2 3 4 5 6	2 3 4 5	1 3 4 6
Δ_{watt} (w.r.t. full-lit)			387.2	387.2	580.8	193.6	387.2	387.2
Luxmeters (head – worn)	1	Δ_{lux} (w.r.t. full-lit)	116.15	123.77	189.01	106.52	148.12	157.07
		Light est. error (w.r.t. GT)	167.2	144.09	102.73	22.94	12.97	13.59
	2	Δ_{lux} (w.r.t. full-lit)	97.68	125.15	169.72	99.17	154.28	167.93
		Light est. error (w.r.t. GT)	194.63	171.74	131.55	9.4	241.12	2.81

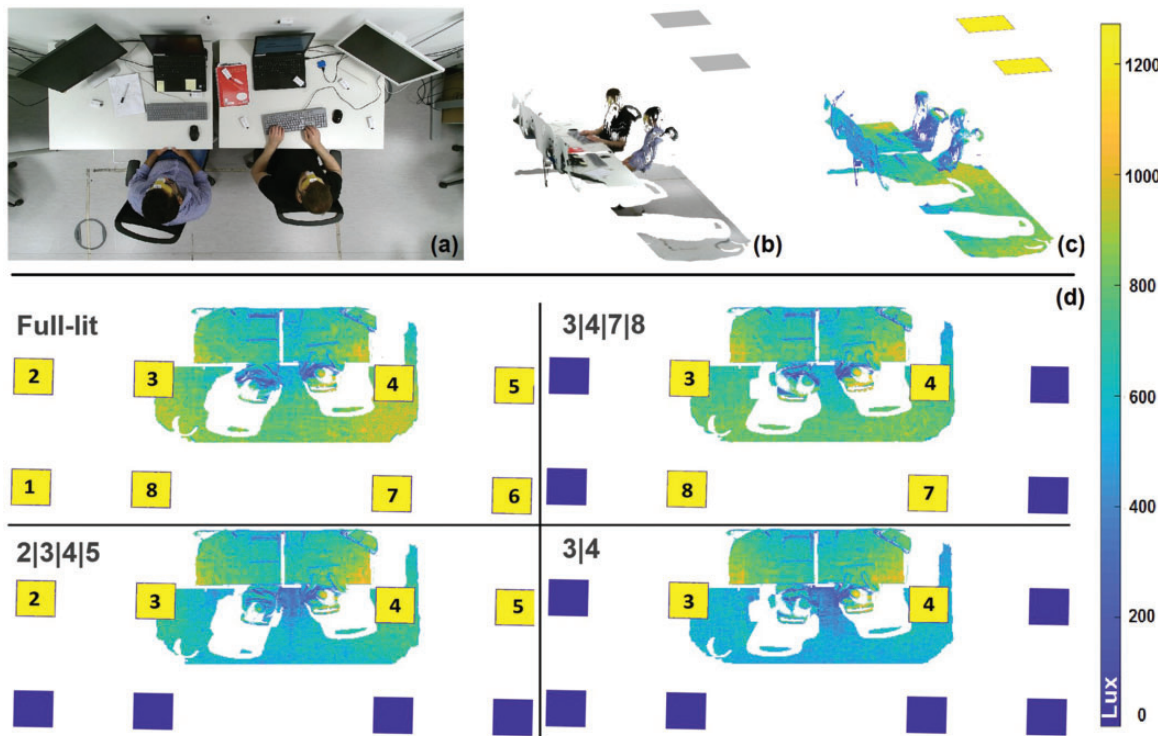


Figure 4 Illustration of the scene, person activity and illumination map for room 4. (a) The top-view image, corresponding to the 3D in (b) with mapped textures. (c) and (d) Present gradient illumination maps, from 3D- and top-view, respectively. In more detail, (d) shows four lighting setups, from the full lit (1|2|3|4|5|6|7|8) to the most energy-saving scenario (3|4) (available in colour in online version).

to implement the 'Invisible Light Switch' (ILS) for light management. The table reports experiments for Rooms 4 and 5, whereby occupants engage in activities, while we change the room illumination and switch off some of the eight luminaires (setup 1|2|3|4|5|6|7|8 refers to 'full-lit' where all luminaires are on, while 3|4 means keeping only luminaires 3 and 4 on. Figure 4 illustrates the setup of Room 4 and the resulting illumination maps, under different lighting setups.

The main performance measure in Table 3 is Δ_{watt} , which quantifies how much power can be saved when turning some luminaires off. For example, when keeping only luminaires 3 and 4 on (setup 3|4), one may save up to 580.8 watt. Over a full working day (8 hours), this setup allows to save up to 99 KWh/m² (including the power consumption of the processing unit), meaning 66% of energy efficiency.

The second important measure in Table 3 is Δ_{lux} , which quantifies how much the illumination in sight does change. We measure this by means of illuminance meters which the occupants wear on their forehead, as previously described in Section 4. For example, in Room 4 and setup 3|4, the person on the right (c.f. Figure 4) wearing the head-worn illuminance meter 1 perceives a difference of 189.01 lx (over the ambient lighting of 1200 lx), the largest in the most-energy saving scenario.

Finally, we report the light estimation error. In the above example (Room 4, occupant 1, setup 3|4) the error is 102.73 lx. This remains comparable to the actual light variation Δ_{lux} . Figure 4 supports the results qualitatively.

5. Conclusion

We have created, introduced and benchmarked the first real-time, end-to-end and human-centric light management system. ILS is based on models for estimating the scene lighting as well as the illumination in sight of the scene occupants. Building upon these

findings, we have defined a system which copes at the same time with human light perception and power efficiency by switching off or dimming luminaires while the changes cannot be perceived by the occupants. Thus, creating the impression of 'all lit' while the scene is optimally lit. Key aspects of this proposition are the relatively small light estimation errors, compared to lighting industrial standards.


Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement No. 676455 and has been partially supported by the project of the Italian Ministry of Education, Universities and Research (MIUR) 'Dipartimenti di Eccellenza 2018–2022'.

ORCID iDs

T Tsismelis  <https://orcid.org/0000-0001-9290-2383>

F Galasso  <https://orcid.org/0000-0003-1875-7813>

References

- 1 Partonen T, Lonnqvist J. Bright light improves vitality and alleviates distress in healthy people. *Journal of Affective Disorders* 2000; 57: 55–61.

- 2 Kuller R, Wetterberg L. Melatonin, cortisol, EEG, ECG and subjective comfort in healthy humans: impact of two fluorescent lamp types at two light intensities. *Lighting Research and Technology* 1993; 25: 71–80.
- 3 Boyce PR. Lighting research for interiors: the beginning of the end or the end of the beginning. *Lighting Research and Technology* 2004; 36: 283–294.
- 4 Kralikova R, Pinosova M, Hricova B. Lighting quality and its effects on productivity and human health. *International Journal of Interdisciplinarity in Theory and Practice* 2016; 10: 8–12.
- 5 Kralikova R, Andrejiova M, Wessely E. Energy saving techniques and strategies for illumination in industry. *Procedia Engineering* 2015; 100: 187–195.
- 6 Kaminska A, Ozadowicz A. Lighting control including daylight and energy efficiency improvements analysis. *Energies* 2018; 11: 2166.
- 7 Guo X, Tiller D, Henze G, Waters C. The performance of occupancy-based lighting control systems: A review. *Lighting Research and Technology* 2010; 42: 415–431.
- 8 De Bakker C, Aries M, Kort H, Rosemann A. Occupancy-based lighting control in open-plan office spaces: A state-of-the-art review. *Building and Environment* 2017; 112: 308–321.
- 9 Tsesmelis T, Hasan I, Cristani M, Del Bue A, Galasso F. Human-centric light sensing and estimation from RGBD images: The invisible light switch. In: *IEEE Winter Conference on Applications of Computer Vision (WACV)*, Waikoloa Village, HI, USA, January 2019: pp. 416–425.
- 10 Tsesmelis T, Hasan I, Cristani M, Galasso F, Del Bue A. *RGBD2lux: Dense light intensity estimation with an RGBD sensor: IEEE Winter Conference on Applications of Computer Vision (WACV)*, Waikoloa Village, HI, USA, January 2019: pp. 501–510.
- 11 Cuttle C. Towards the third stage of the lighting profession. *Lighting Research and Technology* 2010; 42: 73–93.
- 12 Cai H. Luminance gradient for evaluating lighting. *Lighting Research and Technology* 2016; 48: 155–175.
- 13 Hiscocks PD. *Measuring luminance with a digital camera*. Technical report, Syscomp Electronic Design Limited, 2014.
- 14 Marschner S, Shirley P. *Fundamentals of Computer Graphics*. New York: A K Peters/CRC Press, 2015.
- 15 Ul Haq MA, Hassan MY, Abdullah H, Rahman HA, Abdullah MP, Hussin F, Said DM. A review on lighting control technologies in commercial buildings, their performance and affecting factors. *Renewable and Sustainable Energy Reviews* 2014; 33: 268–279.
- 16 Pandharipande A, Caicedo D. *User localization using ultrasonic presence sensing systems: IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, Seoul, South Korea, October 2012, pp. 3191–3196.
- 17 Kim J, De Dear R. Workspace satisfaction: The privacy-communication trade-off in open-plan offices. *Journal of Environmental Psychology* 2013; 36: 18–26.
- 18 Wang H, Pandharipande A, Caicedo D, van den Bosch PPJ. *Distributed lighting control of locally intelligent luminaire systems: IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, Seoul, South Korea, October 2012: pp. 3167–3172.
- 19 Pandharipande A, Caicedo D. Adaptive illumination rendering in LED lighting systems. *IEEE Transactions on Systems, Man and Cybernetics (SMC). Part A, Journal of Systems and Humans* 2013; 43: 1052–1062.
- 20 Tsesmelis T, Hasan I, Cristani M, Del Bue A, Galasso F. *LIT: A system and benchmark for light understanding: IEEE International Conference on Computer Vision Workshops (ICCVW)*, Venice, Italy, October 2017: pp. 2953–2960.
- 21 Masset L, Bruls O, Kerschen G. *Partition of the Circle in Cells of Equal Area and Shape*. Structural Dynamics Research Group, Aerospace and Mechanical Engineering Department, University of Liege, Institute de Mecanique et Genie Civil (B52/3), Technical Report, 2011.
- 22 Hasan I, Tsesmelis T, Galasso F, Del Bue A, Cristani M. *Tiny head pose classification by bodily cues: IEEE International Conference on*

- Image Processing (ICIP)*, Beijing, China, September 2017: pp. 2662–2666.
- 23 WAGO. *Wago Dali Module*. Retrieved 22 June 2020, from <https://www.wago.com/us/controllers-bus-couplers-i-o/dali-multi-master/p/753-647>.
- 24 WAGO. *Wago Dali Configurator*. Retrieved 22 June 2020, from <https://www.wago.com/global/d/6866>.
- 25 Demirkus M, Wang L, Eschey M, Kaestle H, Galasso F. *People detection in fish-eye top-views: VISIGRAPP (VISAPP)*, Porto, Portugal, February 2017: pp. 141–148.
- 26 Lin T, Maire M, Belongie S, Hays J, Perona P, Ramanan D, Dollár P, Zitnick L. *Microsoft coco: Common objects in context: IEEE European Conference on Computer Vision (ECCV)*, Zürich, Switzerland, September 2014: pp. 740–755.