UC Berkeley

UC Berkeley Previously Published Works

Title

Painting with light: An interactive evolutionary system for daylighting design

Permalink

<https://escholarship.org/uc/item/5361j7tk>

Authors

Caldas, Luisa
Santos, Luis

Publication Date

2016-11-01

DOI

10.1016/j.buildenv.2016.07.023

Peer reviewed

Painting with Light: an Interactive Evolutionary System for Daylighting Design

Luisa Caldas^a

Luis Santos

UC Berkeley
Department of Architecture
College of Environmental Design
lcaldas@berkeley.edu

UC Berkeley
College of Environmental Design
Center for the Built Environment
luis_sds82@berkeley.edu

^aCorresponding author. Current address: Department of Architecture. University of California, Berkeley, 384 Wurster Hall #1800, Berkeley, CA 94720-1800, USA.

Abstract

Painting with Light is a user-guided interactive evolutionary system for daylighting design. It allows architects to use color to specify desired light levels in spaces, and searches for solutions that bring building geometry and materials close to performance targets desired by the architect. The proposed interface directly addresses a main limitation of generative design systems based on building performance metrics, by allowing the user to specify daylight spatial patterns with a high degree of granularity. This development poses new challenges for both objective function and penalty function definitions, addressed in this paper. Painting with Light automatically computes and displays statistical indexes that inform the user on the deviation error between the performance of the solution found by the system and the desired targets. The system was implemented in Python on top of a popular Computer-Aided Design software, Rhinoceros, as an add-on to its Visual Programming Language, Grasshopper. This choice of implementation allows the use of Grasshopper's built-in functions and methods for 3D parametric modeling, of tools that provide direct access to Radiance, a lighting simulation software, and of different types of genetic algorithms. Five experiments were conducted on a freeform parametric model for progressive system calibration, which encompassed four steps: 1) adjustment of painted targets to fit the problem feasible solution space; 2) devise appropriate weights and penalty factors for the fitness function; 3) test two different evolutionary solvers; 4) test the system's capability to find a predefined solution where the optimal values were known. After calibration, the system was able to produce solutions that closely approximate the painted goals.

Keywords: performance-based generative design systems, goal-oriented design, daylighting design, digital interfaces, building performance optimization, sustainable architecture

1. Introduction

Painting with Light was developed to help architects deal with one of the most illusive and yet desirable aspects of architecture: daylight behavior in space. Although daylighting design dwells at the very core of architecture creation, light's high degree of variability in time, and the impact it suffers from spatial variables, both intrinsic and extrinsic to a building, makes light behavior in space a challenging aspect to predict and visualize.

Performance-based generative design systems (PGDS) are computational tools used in building design that invert the typical performance-driven design workflow, based on iterative trial-and-error cycles. With PGDS systems, the user first describes the design problem (for example, geometry, materials, variable parameters, and their possible domains and constraints), then establishes desired building performance goals. The system will subsequently search for a solution that best matches the desired performance criteria in the feasible solution space [1, 2]. These systems can be based on a single building performance targets (single-objective PGDSs), or on different performance goals (multiple-objective PGDSs). New programming tools for architectural parametric design [3, 4] that integrate interfaces to building performance simulation software [5, 6], such as Radiance [7] and EnergyPlus [8], with optimization algorithms, are facilitating both the use and deployment of PGDS.

In the field of PGDS, one of the most important factors to control is the ability to set adequate performance targets that closely approach architects intentions for a building. While the search mechanism used certainly represents an important, but solvable, technical aspect, the main challenge in this emerging field resides in establishing a set of goals that allows the search to become meaningful, by producing results that are close to the architect's design intentions. Only by achieving this can performance-based generative design systems become a useful part of the vast toolkit that architects resort to during early design stages workflows.

The new kind of workflow introduced by PGDS systems thus poses two major challenges to designers: (1) how to selected adequate performance targets, needed to compute the fitness function, in a way that actually brings the solutions close to their design intentions, and (2) how to explore and evaluate the solutions found by the system. Significant work has been done in the latter, namely on decision probing tasks for multi-objective optimization problems [9] and in the understanding of the optimal space [10, 11]. Regarding the first problem little has been done, namely in goal-oriented design workflows for sustainable building design. This is mostly because PGDS for environmental design are designed to optimize metrics based on a single or on a simple array of performance values over time (e.g.; end-use energy consumption, total energy consumption, energy use intensity (EUI), carbon dioxide equivalent emissions (CO₂-eq) [12] which can be considered as a relatively simple problem. However, in cases where the building performance metrics have high spatial resolution, such as daylight or acoustic metrics, setting performance targets is a complex task because it should add the description of spatial granularity and pattern variation for the metric in use. That is the main reason that even PGDS based on daylight performance goals [13] tend to oversimplify the fitness function and target distribution by summarizing the performance spatial resolution into a single number or target interval.

To address the complexity of PGDS for daylighting design, this paper proposes Painting with Light, a new, user-guided interactive evolutionary system that allows architects to color in a plan of a space, the desired light levels and patterns. This new interface also supports decision making processes in building design by automatically visualizing and compute different statistical metrics that informs the user on error deviation between the desired targets and the current design performance. To better match the workflow of the proposed PGDS with typical design processes the system was implemented in a popular 3D modeling software suite, Rhinoceros. [14]

2. Related Work

General reviews of the field include the work of Asadi and Geem [1], a technical review of relevant research applying simulation-based optimization methods to sustainable building design, including a summary of the application of common heuristic and meta-heuristic methods to different fields of sustainable building design. Machairas et al. [2] present an extensive review of algorithms for building design. Attia et al. [15] review current trends in simulation-based building performance optimization, and assess the gaps and needs for integrating optimization tools in net zero energy buildings (NZEB) design. Nguyen et al. [16] focus the discussion on discontinuous multi-modal building optimization problems, the performance and selection of optimization algorithms for building performance, multi-objective optimization, optimization under uncertainty, and the propagation of optimization techniques into real-world design challenges. Huang and Niu [17] analyze the history, current status and potential of optimal building design based on simulation performance.

In terms of architectural design strategies, Turrin et al. [18] discuss the benefits derived from combining parametric modeling and Genetic Algorithms (GAs) to achieve a performance oriented process in design, with specific focus on architectural design, with a focus on the key role played by geometry in architecture. Granadeiro et al. [19] propose a method to integrate shape grammars and building envelope design, using MATLAB and EnergyPlus. Caldas and Santos [20] had previously applied a combination of shape grammars and GAs, linked to energy simulations, to the design of energy efficient patio houses.

Regarding decision-making support, Basbagill et al. [21, 22] propose a multi-objective feedback approach for determining life-cycle environmental impact and cost performance of buildings at conceptual design stage, exploring the concept of multidisciplinary design optimization (MDO) already used by Flager et al. [23] in the multi-objective optimization of the structural and energy performance of a classroom. Brown and Mueller [24] also present a multidisciplinary optimization model for structural and energy performance at early stage design phases. In that work, the structural optimization objective was to minimize the amount of steel required, and the energy optimization objective was to reduce the annual energy of the building in terms of lighting, heating, and cooling. The Non-Dominated Sorting Genetic Algorithm II (NSGA-II) was used to iteratively approach the Pareto front over several generations of design alternatives of three types of long steel span structures: a trussed arch, a “PI” structure, which contains two columns and a spanning truss, and a “x-brace” structure, which consists of a three-hinged arch supporting cantilevering roof beams through a series of vertical struts. Lin and Gerber [25, 26] focus on early stage design decisions, proposing a framework for concept design based on MDO, using a custom GA for design and performance optimization. The work demonstrates the impact of applying this process to energy use intensity in early stage architectural design, and identifies performance feedback criteria for MDO design and implementation.

Welle et al. [27] propose to address the dependency of MDO processes on the designer’s ability to structure optimization problems for specific challenges, by developing a computer-aided design (CAD)-centric attribution methodology for multidisciplinary optimization environments (CAMMOE), that enables designers to improve the accuracy of their optimization processes by helping them develop and analyze alternative spaces that are of the right size and composition to meet their design intent. In a

previous work, Welle et al. [28] present ThermalOpt, a methodology for automated BIM-based multidisciplinary thermal simulation intended for use in MDO environments.

Diakaki et al. [29] propose a multi-objective mathematical programming method to select among different energy efficiency measures, considering both energy savings and costs. Ferrara et al. [30] describe a simulation-based optimization method for to find the cost-optimal level for a nearly Zero Energy French single-family building. Xu et al. [31] developed a systematic methodology to minimize building heating and cooling loads using a combination of experimental design, to create an initial ranked list, and non-sorting GA to select optimal sets of building design factors. Also for residential building design, Fesanghary et al. [32] developed a multi-objective optimization algorithm that tackles at the same time energy efficiency, financial cost, and different environmental impacts measured in CO₂-eq emissions. The optimization algorithm is based on harmony search algorithm and it was tested in a typical single-family house model with a parametric envelope.

Asadi et al. [33] propose a multi-objective optimization model that combines the fast evaluation of Artificial Neural Networks (ANNs) with the optimization power of GAs. The optimization algorithm is able to quantitatively assess and select, from a set of predefined retrofit actions, the most suitable suite in a building retrofit project. The optimization method was applied in the retrofitting of a school building and it focused on optimizing energy consumption, retrofit cost, and thermal discomfort hours. Murray et al. [34] also propose a simulation-based optimization approach for the retrofitting of existing buildings. The method uses a simplified and quick method to predict the overall thermal energy consumption of a building based on degree-days, derived from the Chartered Institute Building Services Engineers (CIBSE) Guide TM41, Degree Days: Theory and Application. From a discrete list of variables the optimization model, based on GAs, minimizes carbon emissions, payback time, and overall energy consumption.

For façade design, Wright et al. [35] present a multi-objective optimized design of fenestration based on dividing a building facade into a number of small regularly spaced cells, using a multi-objective GA to minimize energy use and capital cost. C Kasinalis et al. [36] describe a framework for assessing the performance of seasonally adaptable facades, using multi-objective optimization. Schneider and Donath [37] present a generative method that simultaneously varies both topological and metric properties of a façade geometry. Wang et al. [38] use a combination of GAs and EnergyPlus to select glazing types that promote an energy efficient balance between thermal and lighting performance, based on a database of NFRC rated products from different glazing manufacturers.

In the field of daylighting simulations, more advanced lighting metrics are continuously being proposed. An unified visualization framework for daylighting simulation is advanced by Andersen [39], including both perceptual, visual (comfort) and non-visual physiological (health) approaches to daylighting. The paper describes a holistic view on daylighting performance based on five “good daylighting” goals, calling for goal-based design decision support with non-deterministic feedback loops. Andersen et al. [40] extended the Lightsolve expert system, by creating a comprehensive, goal-based design support focused on daylight dynamics, able to balance illumination, glare and solar gains over a year, and testing it in a user study focused on early design stages. Rakha and Nassar [41] used a combination of Radiance simulation and LUA scripting to create a generic tool for generation of curvilinear ceiling forms, according to a criteria of daylight uniformity ratio.

Futrell et al. [42] propose a method for optimizing building design to minimize lighting loads, by combining GenOpt with Radiance simulations, which determine the building daylighting performance using dynamic climate-based lighting simulations. Four optimization algorithms implemented in GenOpt are compared: Simplex Algorithm of Nelder and Mead with the Extension of O'Neill (SA), Hooke Jeeves (HJ), Particle Swarm Optimization using Inertia Weight (PSOIW), and a hybrid PSO Constriction/Hooke Jeeves (PSOC/HJ) algorithm. Futrell et al. [43] then propose a bi-objective optimization of building enclosures, both for thermal and lighting performance.

Despite the current variety and application of PGDSs, only a few goal-oriented design tools have used visual interfaces based on brushing and linking techniques [44], to set the goals for an optimization process. In the field of building performance-based design, there is very few goal-oriented design tools that use a similar approach to the one described in this paper. AudiOptimization [45], for acoustics, and LightSolve [40, 46], for lighting, will be shortly reviewed below.

1) AudiOptimization

AudiOptimization [45] is a PGDS for acoustic design. The workflow proposed by this system encompasses two stages. In the first stage the user describes a parametric model of the space to optimize, including both its fixed characteristics and discrete geometric and material variables to be optimized, and respective minimum and maximum ranges. This description defines the solution space of the design problem, i.e.; all the geometric and material combinations possible. In the second stage the user sets the goals for different acoustic metrics (interaural cross-correlation coefficient, early decay time, bass ratio, and sound strength/level) by painting them in different spatial elements. Figure 1 shows AudiOptimization's goal specification visual interface, and illustrates how objectives are painted through it.

After describing both the design problem and the desired acoustic targets, AudiOptimization searches the solution space for a high performance solution by combining two search procedures: Simulated Annealing (SA) for a more global search, and steepest descent (a gradient-based optimization algorithm) for a more refined search around the global minimum found by SA.

Figure 1. Painting objectives through AudiOptimization's visual interface. Left: sound level painted goals in different surfaces for 80 ms. Right: sound level specification editor for a specific surface. Top row shows the current sound level simulation, the middle one the painted targets for this particular metric, and the bottom row the difference heat map between the two. Image source: [45]

2) LightSolve

Light Solve [40, 46] is an interactive goal-oriented design tool that integrates in a single workflow several daylighting considerations such as daily and seasonal variations, illumination, and visual comfort. LightSolve design workflow is based on a bidirectional interaction between user inputs and system outputs through three main interfaces: (1) geometry and material editor; (2) a set of selection tools for specifying areas, views and/or times of interest; and finally (3) the performance target interface, which allows the user to specify or change the current design goals and/or constraints directly on the areas of interest of the selected views simulated in Radiance.

In the first iterations of LightSolve the user could draw areas for a minimum illuminance target, areas that should avoid glare problems, and finally areas for specific light effects. The current LightSolve interface is illustrated in Figure 2, where an interactive and linked relation between performance and renders are explored.

If the current design does not meet all of the designer's goal, LightSolve will propose effective changes to improve its daylight performance through a search mechanism based on a set of 'expert rules'. This 'expert system' is based on a Design of Experiments (DoE) approach. Thus, unlike other PGDS such as AudiOptimization [45], eifForm [47], genOpt [48], and GENE_ARCH [49, 50], LightSolve's goal-oriented design workflow does not aim to find a global or even a local optimum but to predict the effectiveness of certain design changes on the daylighting improvement of a current design status.

Figure 2. LightSolve interface. Left: time-varied daylight performance displayed in temporal maps interactively linked to renders of the space under study. Right: annual image map showing renderings over time. Image source: [40]

Audioptimization and LightSolve have some limitations in their approaches to target setting and output assessment tasks. Regarding Audioptimization, both the visual interface and the brushing and linking techniques available to paint objectives are advanced and complete. The generation of a difference heatmap between desired performance targets and the optimized solutions results is useful on design quality assessment tasks. This visualization in particular actively contributes to the understanding of the design problem on the following aspects: 1) to what degree spatial performance patterns have been met; 2) check the feasibility of painted goals; 3) constraint levels applied to the design problem, i.e., if the design problem is over or under-constrained.

However, this diagnostic technique is purely visual and only computes the difference between the desired targets and simulated performance. Advanced analytical statistical error metrics typically used in building performance simulation, such as percentage error (% error), Coefficient of Variation of the Root Mean Square Error (*CVRMSE*), and Normalized Mean Bias Error (*NMBE*), could be automatically provided by the system in order to better support better decision making processes.

Although LightSolve promotes the integration of different daylight considerations and metrics in a single design workflow, the tool, at its early stage of development [46], had two main limitations regarding performance target definition. The first one was the simplification of setting an illuminance target by assigning a minimum or a maximum value to a specific area. This did not allow the specification of spatial light patterns. The second limitation was related with view dependency on performance and constraint definition. The user could only set goals and constraints on a set of images/renderings that had been previously selected, thus promoting partial improvements rather than a global approach to the lighting quality of the space. More recently [39, 40], LightSolve overcome some of its early limitations. Currently, the tool supports a sensor point workplane grid, a goal-based illuminance temporal maps displayed alongside interactive renderings, and a triangular color scale specially developed to assess goal compliance. Different daylight metrics, such as Acceptable Illuminance Extent (AIE) [40], were also incorporated. Nevertheless, this new version seems to abandon goal assignment based on drawing

techniques, more suitable to describe spatial daylight patterns, focusing more on time-mapped visualizations of daylight performance. Target definition is still based on the definition of a minimum and maximum illuminance value, but now to several workplane sensor points.

Outside the building performance field, some computer graphics tools have been using interactive approaches to steer image-editing software solvers. An example is the Interactive Digital Photomontage tool [51]. From an initial set of color coded images of the same scene, the user defines a goal image by painting a series of strokes in one image using the designated color associated with the other source images. Then, the tool creates a composite image that approximate to the desired objective using two main techniques: graph-cut optimization, to find the best possible seams to cut the various source images; and gradient-domain fusion, to remove any visible artifacts that might remain after the image seams are joined. Figure 3 illustrates the Interactive Digital Photomontage workflow by exemplifying how the user selects and paints the objectives over an image, the relation between the targets set for each part of the goal-image and the initial set of images, and the final composite image generated by the tool.

Figure 3. Example of the Interactive Digital Photomontage workflow. Left: the initial source image set. The colors around each image represents an objective associated with each specific image to be painted on one of the source images. For example, red represents the maximum luminance objective, while blue is the minimum luminance objective. Middle: one of the source images painted with the different image-objective goals. Right: the final composite image found by the Interactive Digital Photomontage solver.
Image source: [51]

3. Research objectives

The main objective of this work is to address the problem of setting performance targets for daylighting-based PGDS by proposing an initial, proof-of-concept interactive interface in the Rhino environment. The integration of the proposed tool with Rhino allows the access and integration of the following components: (1) advanced 3D modeling tools; (2) Grasshopper's (a Visual Programming Language - VPL - for Rhino) built-in functions and methods for parametric/generative design; (3) dedicated interfaces for lighting simulation with Radiance and/or DAYSIM [52], such as DIVA [5]; and finally (4) evolutionary solvers for search and optimization, such as Galapagos, which uses a simple Genetic Algorithm (GA) or Octopus, which is able to perform multi-criteria optimization using Pareto Genetic Algorithms (SPEA2).

The proposed interface aims to facilitate target description through the use of painting-style processes based on brushing and linking techniques. A painting-style interface will allow designers to easily set performance goals and define their spatial distribution to building performance-based problems. By

promoting the draw of colored regions, the proposed interface better approximates the methods commonly used by designers in their workflows, thus being more intuitive and easy to use.

Finally, this work aims to extend and address some of the limitations found in previous research [40, 45, 46] by implementing a visual performance target editor that: 1) fully addresses the spatial resolution and pattern variation of the daylight problem; 2) supports any grid-based daylight metric, namely, Daylight Factor, point-in-time illuminance and climate-based metrics such as Daylight Autonomy, Spatial Daylight Autonomy, and Useful Daylight Illuminance.

4. Methods

The development of the Painting with Light encompassed three main stages: (1) interface development and implementation; (2) design of experiments (DoE) and system calibration; (3) results and discussion. The methods for each stage are shortly described below:

4.1 Interface development and implementation

Painting with Light was implemented in Rhino using Grasshopper VPL and integrated development environment (IDE), and GHPython - a Python library and interpreter for Grasshopper. The interface allows designers to use typical Computer Aided Design (CAD) drawing techniques to color the performance targets for the space under study. This is done in two steps: (1) define areas with a specific color/performance target by drawing enclosed areas with closed polylines or Non-Uniform Ratio Basis Splines (NURBS) curves; and (2) use sliders to control the degree and strength of light decay from the drawn areas. Figure 4 shows how by drawing closed NURBS curves the user paints the space and links the sensor nodes to a specific performance target.

Figure 4. Painting performance targets by drawing closed NURBS curves over a sensor node grid.

Figure 5 shows how the user can redraw the NURBS curves that represent lighting boundaries, by manipulating their control points to repaint the goals.

Figure 5. Adjusting the painted performance targets by manually manipulating curve control points.

Finally, Figure 6 illustrates how the desired light decay can be controlled through the use of sliders.

Figure 6. Controlling the light decay pattern through dedicated sliders.

The system automatically groups the sensor nodes by color, and consequently by performance target value. It is also possible to change the colors' performance target by changing the numeric range of the color scale, by using dedicated sliders for that purpose. Figure 7 illustrates how performance targets can be changed.

Figure 7. Changing performance targets by resetting the color scale values using sliders.

Regarding the visualization of results, Painting with Light supports both visual and analytical tools. Visually, the interface pairs in the same viewport three heat maps: (1) the painted performance target sensor node grid mesh, (2) the simulation result of the solution found by the PGDS's search mechanism, and (3) a difference map (Δ map) between simulation results and desired targets. The difference map is automatically generated with a color scale that ranges from blue to white, to red. Blue nodes indicate that the solution has a performance below the target, white nodes shows that the design achieved the target, and red nodes report that the daylight performance is higher than the painted targets. Figure 8 illustrates the three painted meshes, in a Rhino viewport.

Figure 8. Visualization of results in Painting with Light.

Finally, the system automatically computes and displays the different statistical metrics mentioned in section 3. Figure 9 shows how the computed statistics indexes are presented. More details about the applied statistical methods can be found in section 4.2.2.

Figure 9. Statistic results dashboard in Painting with Light.

4.2 Design of Experiments and system calibration

Several experiments were conducted to assess the accuracy of the system and to progressively calibrate it. The method used to conduct those experiments contemplated three different steps: (1) parametric implementation of the building geometry used to conduct the several optimization runs; (2) definition of different metrics used, both for simulation and for comparison between experiments; and (3) Design of Experiments (DoE). Each of these steps is described below.

4.2.1 Implementation of the parametric model

A parametric model of a free form open plan building was developed to test the system a parametric model of a free form open plan building was developed. The building assumes a curve bobbly form with an envelope that alternates between concave and convex shapes, inducing therefore self shading surfaces. The floor plan is fixed and north-south oriented in the images. A deep floor plan is used, particularly in the north side, to promote skylight use, as side lighting becomes insufficient to appropriately daylight such deep spaces. The non-conventional building form challenges the designer's light prediction skills, while promoting the generation of a wide variety of light distribution patterns (which are more difficult to obtain with more narrow and conventional layouts).

Roof shape, skylight radius and skylight positioning are the model's parametric component. Roof shape is generated by four parametric interpolated NURBS curves, which define the edges of a surface that is trimmed by the floor plan surface projection. Each curve has seven control points, the start and end points fixed at 3.5 m height, and five intermediate points with variable z -axis coordinates. The parametric surface generated by these four curves is not only able to create a wide variety of roof designs but also to shape the building curtain wall with a high degree of variability. This allows the system to search for different glass façade height variation patterns to control the interior light levels. The roof can also generate an overhang for shading purposes, with a variable depth (d) between 2 and 6 meters. Finally, four circular skylights, with parametric radius (r), can be generated and parametrically placed on the roof through the specification of its centroid through the variable surface coordinate parameters u and v .

Figure 10 shows a plan of the implemented parametric model with the variable parameters marked in red. Figure 11 presents an axonometric view of the building showing the roof generating curves and their parametric control points. Table 1 lists the variable parameters types encoded into the model, describes them, and shows their respective domain. Notice that the domain for the skylight center location variables (u and v) varies between 0 and 1 because the surface domain is reparameterized to this domain due to algorithmic generalization purposes.

Figure 10. Plan of parametric model. All variable parameters are marked in red.

Figure 11. Axonometric of parametric model showing the generating parametric curves of the roof. All variable parameters are marked in red.

Parameter	Description	Minimum	Maximum
d	Depth of roof overhang	2 m	6 m
r_n	Skylight radius	0 m	2.5 m
$Skylight_n u_n$	u coordinate (u_n) of each skylight center point ($Skylight_n$)	0	1
$Skylight_n v_n$	v coordinate (v_n) of each skylight center point ($Skylight_n$)	0	1
$Crv_n Pt_{zn}$	Move in the z -axis for each point (Pt_{zn}) of each generating roof curve (Crv_n)	-1.5 m	5 m

Table 1. Variable model parameters types and their domain bounds.

After encoding the parametric geometry, a specific Radiance material was assigned to each surface of the model. Table 2 describes the Radiance material properties, both for opaque and transparent surfaces, used in all experiments.

Surface	Reflectance (%)	Visual Transmittance (%)
Outside ground	20	N.A.

Roof cladding	50	N.A.
Curtain wall glass	N.A.	64
Skylight glass	N.A.	64
Interior floor	40	N.A.
Interior ceiling	80	N.A.

Table 2. Reflectance and Visual Transmittance of materials used in parametric model.

The ambient settings used for the backward raytracing were:

- Ambient bounces (-ab) – 4, four ambient bounces guarantees enough accuracy without a huge computational burden, making it a suitable number for an optimization cycle which will run a large number of models.
- Ambient resolution (-ar) – 300, a reasonable number for a point grid calculation with Radiance's subroutine rtrace.
- Ambient divisions (-ad) – 1000.
- Ambient sampling (-as) – 500, typically half of -ad.
- Ambient accuracy (-aa) – 0.1.

In order to compare results, the simulation settings did not change or adapt to the requisites of more complex geometries. For an optimization process, it is also important that the settings do not change to avoid any type of distortion. However, to improve accuracy, future work will extend PWL's features to pre-calculate the simulations settings based on a parametric model.

4.2.2 Metrics

Painting with Light supports all current daylight metrics available through DAYSIM [52] and DIVA [5], including climate-based metrics like Daylight Autonomy, Useful Daylight Illuminance or Continuous Daylight Autonomy. However, Daylight Factor was used as the daylight metric in this work, due to two fundamental reasons:

- Computational time. This paper presents an initial prototype for an interactive evolutionary generative system for daylight performance-based design. Due to the proof-of-concept nature of the research, a fast and easy to compute daylight metric was selected, with the goal of quickly running several optimization tests. Running more detailed metrics, such as climate-based ones, would require significantly higher computational times, not compatible with this phase of tool development.
- Daylight Factor is a well-known, non point-in-time daylight metric, thus more intuitive for designers than more recently developed metrics such as climate-based ones. As this metric is more familiar to designers that are not daylight experts, it makes the system easier to use and to predict daylighting performance, a factor that is highly relevant for the quality of the user painted objectives.

To assess the quality of the solutions found by the PGDS and to compare the different experiments conducted in this work the following statistical metrics were used:

- **Absolute difference error** is defined in equation (1):

$$\text{Absolute difference} = |y - \hat{y}| \quad (1)$$

where y are the simulation results of the solution found by the system, and $\hat{y}$ are the objective values set by the user.

- **Percentage error**: a very well known and simple metric commonly used in comparison tasks of different simulations. However, this metric often led to a compensation effect, whereby over-estimations cancel under-estimations [53, 54]. To overcome this limitation, more advanced statistical indices, that better represent the performance of a model, were also used (i.e.; *NMBE* and *CVRMSE*). Percentage of error is defined as:

$$\% \text{ error} = 100\% \times |y - \hat{y}| / \hat{y} \quad (2)$$

where y are the simulation results of the solution found by the system, and $\hat{y}$ are the target values set by the user.

- **Percentage of area $\leq 25\%$ error**: a simple calculation that returns the percentage of area that is in the first % error quartile. The threshold of the % error is intentionally higher because of the limitation inherent to this index (cancellation effect), and because more accurate metrics (i.e.; *NMBE* and *CVRMSE*) use a wider acceptance range. Thus, the purpose of considering 25 % error was to get a rough idea about the overall quality of the spatial distribution of % error by simply quantifying the area that is within the first quartile of % error.
- **Percentage of area $\leq 10\%$ error**: it uses the same method described percentage of area $\leq 25\%$ error but with the typical acceptance threshold used in the early years of building simulation, 10%. Similar to the previous metric, this index was computed to have an initial idea of the spatial % error distribution and performance. However, it should not be considered as the final metric on the deviation error assessment between the solution found by the system and the desired targets.
- **Coefficient of Variation of the Root Mean Square Error (*CVRMSE*)**: an index commonly used in energy model calibration that allows to determine how well a model fits with the target data by capturing offsetting errors between the defined target and simulated data. This metric does not suffer from the cancellation effect [55]. Equation (3) defines *CVRMSE*:

$$\text{CVRMSE} = 100\% \times \frac{\sqrt{\frac{\sum_{i=1}^N (y_i - \bar{y})^2}{n-1}}}{\bar{y}} \quad (3)$$

where $\bar{y}$ is the mean of the simulation results of the solution found by the system, and n is the number of data points (i.e.; number of sensor nodes). The range of this metric is 0 to $+\infty$. Because there is no guidelines regarding the use of this metric in daylight simulations the acceptance criterion used for *CVRMSE* was the ASHRAE guideline 14 [56], for hourly based energy simulations: 30%. The hourly threshold of guideline 14 was favored against the monthly one

because the amount of considered data points is clearly greater than 12, thus, better approximating the size of energy hourly simulations.

- **Normalized Mean Bias Error (NMBE):** this metric describes the variation between the mean measured and predicted values and it is a good indicator of bias in the solution. It captures the mean difference between performance target and solution simulation results. *NMBE* is defined as:

$$NMBE = 100\% \times (\sum(y_i - \hat{y}_i) / (n - p) \times \bar{y}) \quad (4)$$

where p is the P-value, which for this purpose is equal to 1 [55]. The range of this metric is $-\infty$ to $+\infty$. Due to the same reasons pointed-out in CVRMSE, the acceptance criterion used for *NMBE* is based on the ASHRAE Guideline 14 [56], for hourly based energy simulations: $\pm 10\%$.

- **Percentage of improvement** [57]: this index informs how much the performance of a solution improved in relation to a baseline case, thus, it was used in the calibration of the system. Percentage of improvement is defined in equation (5):

$$\% \text{ improvement} = 100\% \times (x - \bar{x}) / x \quad (5)$$

where x are the results of the base case and $\bar{x}$ is the design performance metric under study.

Although several error deviation metrics were used, CVRMSE and NMBE were the most important indices in the calibration of Painting with Light, as they have been report has the standardized statistical indices that better represent the performance of a model [55].

4.2.3 Design of Experiments (DoE)

The DoE set for Painting with Light calibration tried to emulate how a typical designer would use the system. In all DoE experiences, the fitness function was elaborated based on the goal programming method [58-60]. This technique was used due to its computational efficiency and to always yield a dominated solution if the goals are in the feasible domain of the problem [58]. To further improve the system's search procedure, penalties and weight factors were aggregated to the goal programming method.

Five testing experiments were performed. The first four experiments consisted on progressive refinements on the calculation of the Genetic Algorithm objective function, by experimenting with methods for error computation as the main way for improving Painting with Light prediction capabilities. After calibrating the system, a final experiment was conducted to assess if Painting with Light was able to find a pre-calculated solution, for which the 'optimum' value was already known, by performing a DF calculation of a given geometry.

Each of the five experiments is briefly described below:

Experiment 1 - In the first experiment the search mechanism used was Galapagos [61]. The Galapagos Genetic Algorithm is a standard, single function GA, which only accepts single values as objective function. The fitness function used was a simple minimization of the sum of the averages of the absolute difference between simulation results and performance targets.

Experiment 2 - Penalties were introduced in the cost function in order to penalize huge deviations between the simulated results and their correspondent goals on an attempt to force the system to find solutions that better match the painted objectives.

Experiment 3 - A weighting factor was introduced in the second experiment fitness function, all the other factors remaining the same. The weighting factor aimed at better distributing the penalties, balancing them accordingly to the weight of the respective goal area in order to minimize distortion effects in the computation of the overall cost. In this way a huge deviation could be smoothed out if it occurs in a relative small goal painted area.

Experiment 4 - A slightly improved painted target was provided and the Octopus Pareto GA was used as the search mechanism. A new GA was tested since Galapagos GA tends to found more local minima as the complexity of the problem increases [61]. Thus, the more robust Strength Pareto Evolutionary Algorithm 2 (SPEA 2) [62] available through Octopus Grasshopper add-on was used. The fitness function was the same as in experiment 3 but the penalty function was redefined to penalize less sensor nodes that reported a performance closer to their corresponding target, and to highly penalize the ones that had higher deviations. In order to use a Pareto GA to optimize a problem with a cost function that returns a single value, genetic diversity was assigned as an inoculum second optimization vector.

Experiment 5 - finally, after the progressive calibration of Painting with Light, the results of a pre-defined design were used as painted goals, in order to assess if the system was able to find a solution similar both in performance and in design features. This final experiment also allowed evaluating the impact of a highly accurate painted goal target space in the system, considering that the interface developed to paint the daylight goals roughly approximates the phenomena of light decay. The search mechanism and the fitness function applied on this experiment were the same used in experiment 4.

5. Results

Results for each experiment are reported and shortly reviewed in the following subsections.

5.1 Experiment 1

Figure 12 shows the geometry proposed by Painting with Light in experiment 1. Figure 13 shows the painted performance targets, the proposed solution Daylight Factor (DF) performance, and the spatial mapping of their difference. The generated overhang (d) is 5 m deep.

Figure 12. Views of experiment 1 solution. Left: southeast perspective. Right: northwest perspective.

Figure 13. Experiment 1 visual results overview. Left: painted objectives. Center: Daylight Factor simulation results of the solution proposed by the system. Right: difference (Δ) heatmap between painted objectives and simulation results.

The colors used in the painted targets were: Red: 10% DF; Orange: 7% DF; Yellow: 5% DF; Light yellow: 4% DF. Table 3 shows the values of the roof design variables while table 4 reports the skylights variables.

	Ptz1 [m relative to 3.5 m]	Ptz2 [m relative to 3.5 m]	Ptz3 [m relative to 3.5 m]	Ptz4 [m relative to 3.5 m]	Ptz5 [m relative to 3.5 m]
Crv1 (south)	-1.1	-1.5	1.8	-0.8	0.1
Crv2 (east)	2.7	2.9	3.9	1.4	3.8
Crv3 (north)	0.8	-1.1	2.1	1.1	-1.3
Crv4 (west)	2	-0.2	2.1	1.4	3.3

Table 3. Experiment 1 roof design variables.

	u [0-1]	v [0-1]	radius [m]
Skylight1	0.54	0.48	2.1
Skylight2	0.62	0.54	0.4
Skylight3	0.57	0.44	1.24
Skylight4	0.56	0.48	2.5

Table 4. Experiment 1 skylights' variables.

The system proposed a solution with a deep overhang of 5 m, to respond to the fact that the user did not predict in the painted targets that higher Daylight Factor levels usually occur closer to the facades. As for the four skylights, only one of them was wrongly positioned, but its radius is so small that had little effect on the design performance. The other three are quite accurately positioned and have radiuses that vary from 1.24 m to 2.10 m. It is possible to observe that the system tried to aggregate two skylights in an attempt to match the elliptical red area painted in the southeast part of the building. The system also slightly elevated the northwest corner of the roof to increase the glazing area, to achieve higher light levels close to the red spot painted in the building's northwest corner. However, as we can see in the difference map, the system missed most of the DF targets by underestimation. It totally failed to meet the targets set for the northwest corner, which requested the positioning of a skylight directly above them.

Table 5 shows the average of absolute difference and % error for each target color/DF levels, and for the overall space. It shows that the biggest deviation between painted objectives and proposed solution performance happened in the 10% and 7% DF target areas. The percentage of area $\leq 25\%$ and $\leq 10\%$ error are both very small, 21.87% and 6.59% respectively. *NMBE* of this experiment is -53.06% and the *CVRMSE* is 77.46%, both very far for the acceptance criteria assumed for those metrics.

Goal area	Absolute difference (average) [DF]	% error (average)
10% DF (Red)	3.37	56.2
7% DF (Orange)	2.08	48.1
5% DF (Yellow)	1.1	41.6
4% DF (Light yellow)	1.97	29
Total	2.2	44

Table 5. Total and partial averages for absolute difference and percentage of error between desired goals and system output in Experiment 1.

The poor results of this experiment are mainly due to two reasons: the unfeasibility of the painted target solution due to the user lack of understanding how DF works spatially; and the distortion introduced by the necessary averaging on the fitness calculation for Galapagos GA.

5.2 Experiment 2

The design solution found by the system in this experiment is illustrated in Figure 14. The performance of the solution and the painted objectives can be seen in Figure 15. Painting with Light modeled one overhang of 2.36 m. The painted objectives are more feasible in this experiments, namely on the perimeter, where higher values were painted. Table 6 and 7 report the design variable values for this experiment. Table 8 quantifies the average of absolute difference and % error between painted objectives and the design solution proposed by Painting with Light.

Figure 14. Views of the solution of experiment 2. Left: southeast perspective. Right: northwest perspective.

Figure 15. Experiment 2 visual results overview. Left: painted objectives. Center: Daylight Factor simulation results of the solution proposed by the system. Right: difference (Δ) heatmap between painted objectives and simulation results.

	Ptz1 [m relative to 3.5 m]	Ptz2 [m relative to 3.5 m]	Ptz3 [m relative to 3.5 m]	Ptz4 [m relative to 3.5 m]	Ptz5 [m relative to 3.5 m]
Crv1 (south)	0.3	0.5	-1.2	2.2	1.1
Crv2 (east)	0.1	1.4	1.8	0.4	-1.1
Crv3 (north)	1.2	2	1.8	-1.5	1.1
Crv4 (west)	0.7	2.2	-0.3	2	2.2

Table 6. Experiment 2 roof design variables.

	u [0-1]	v [0-1]	radius [m]
Skylight1	0.54	0.53	0.75
Skylight2	0.38	0.62	1.42
Skylight3	0.46	0.47	1.52
Skylight4	0.55	0.49	1.95

Table 7. Experiment 2 skylight's variables.

Goal area	Absolute difference (average) [DF]	% error (average)
10% DF (Red)	4.15	41.5
7% DF (Orange)	2.9	41.6
5% DF (Yellow)	2	40
4% DF (Light yellow)	1.43	38.9
Total	2.65	40

Table 8. Total and partial averages for absolute difference and percentage of error between desired goals and system output in Experiment 2.

The system proposed a solution with an almost flat roof that is slightly tilted towards East. This generates a glass facade with approximately the same height in all directions. The overhang produced is less deep than the one created in the first experiment, measuring 2.36 m. This is consistent with the higher DF targets set for the perimeter in this experiment. For the skylights, this experiment produced a solution that is more compatible with the defined goals. It created four skylights, a group of two to match the red target area (10% DF) on the Southeast area of the building, that once again attempts to match the elliptical shape

of the red target; one for the Southwest red target values, and one for the Northwest red target area. However, through the analysis of the difference map we can observe that, although it did better than in the first experiment, the system was still unable to accurately locate the Northwest skylight, generate the correct skylight radius, and to produce a balanced solution between the curtain glass height and the overhang. This led to some overlit areas (the more narrow convex areas of the building) and to a large blue dim area which does not correspond to the DF pattern target painted by the user.

In general, the average % error decreased in all the target areas for experiment 2. However, the average absolute difference increased due to some outlier data points. Although both the percentage of area $\leq 25\%$ and $\leq 10\%$ error increased, to 31.4% and 15.5% respectively, they still indicate that a large portion of the sensor node grid misses the target criteria. This is reinforced by the *NMBE* of -38.3%, and the *CVRMSE* of 69.04%, both still far from their acceptance range criteria.

5.3 Experiment 3

Experiment 3 solution and results are reported in Figures 16 and 17, Table 9, 10, and 11. The generated overhang (d) for this experiment is 2 m deep.

Figure 16. Views of the solution of experiment 3. Left: southeast perspective. Right: northwest perspective.

Figure 17. Experiment 3 visual results overview. Left: painted objectives. Center: Daylight Factor simulation results of the solution proposed by the system. Right: difference (Δ) heatmap between painted objectives and simulation results.

	Ptz1 [m relative to 3.5 m]	Ptz2 [m relative to 3.5 m]	Ptz3 [m relative to 3.5 m]	Ptz4 [m relative to 3.5 m]	Ptz5 [m relative to 3.5 m]
Crv1 (south)	0.2	0.5	-1.2	2.1	1
Crv2 (east)	0	1.2	1.8	0.4	-1.1
Crv3 (north)	1.1	2	1.7	-1.5	1.1
Crv4 (west)	1	2.5	1.5	2	2.5

Table 9. Experiment 3 roof design variables.

	u [0-1]	v [0-1]	radius [m]
Skylight1	0.54	0.52	0.7
Skylight2	0.38	0.62	1.85
Skylight3	0.46	0.47	1.95
Skylight4	0.55	0.49	1.9

Table 10. Experiment 3 skylight's variables.

Goal area	Absolute difference (average) [DF]	% error (average)
10% DF (Red)	3.6	35.8
7% DF (Orange)	2.5	35.6
5% DF (Yellow)	1.68	33.5
4% DF (Light yellow)	1.4	38.1
Total	2.3	36

Table 11. Total and partial averages for absolute difference and percentage of error between desired goals and system output in Experiment 3.

The design solution found by the system is very similar to experiment 2. However, the glass curtain wall is slightly higher, what improved the performance of the solution in the perimeter. The radiuses of the skylights are also slightly larger, further approximating DF criteria for inner areas. However, the proposed skylight locations are still inadequate, namely for the Northwest targets. This experiment reported a higher percentage of area in the first quartile of 25 % error and in 10 % error: 44.77%, and 21.84% respectively. Although *NMBE* improved to -17.2% and *CVRMSE* to 55.23%, they still not meet their acceptance range.

5.4 Experiment 4

Figure 18 and 19 shows the solution found in the fourth experiment. Table 12 and 13 shows the design variables. The overhang depth (*d*) is 2.9 m. Table 14 reports the absolute deviation and the percentage of error between the painted targets and the design performance.

Figure 18. Perspective views of the solution of experiment 4. Left: southeast perspective. Right: northwest perspective.

Figure 19. Experiment 4 results spatially mapped. Left: painted objectives. Center: Daylight Factor simulation results of the solution proposed by the system. Right: difference (Δ) heatmap between painted objectives and simulation results.

	Pt 1 [m relative to 3.5 m]	Pt 2 [m relative to 3.5 m]	Pt 3 [m relative to 3.5 m]	Pt 4 [m relative to 3.5 m]	Pt 5 [m relative to 3.5 m]
Crv1 (south)	-1.5	-0.6	4.4	2.5	-0.9
Crv2 (east)	-1.2	0.2	4.4	0.1	3.4
Crv3 (north)	-1.5	2.7	1.7	2.1	4.4
Crv4 (west)	4.3	1.8	2.9	-0.2	4

Table 12. Experiment 4 roof design variables.

	u [0-1]	v [0-1]	radius [m]
Skylight1	0.42	0.58	1.75
Skylight2	0.57	0.48	2.4
Skylight3	0.59	0.34	0
Skylight4	0.45	0.47	2.15

Table 13. Experiment 4 skylight's variables.

Goal area	Absolute difference (average) [DF]	% error (average)
10% DF (Red)	1.62	16.15
7% DF (Orange)	0.76	10.8
5% DF (Yellow)	0.92	18.4
4% DF (Light yellow)	0.88	22
Total	1.08	16.3

Table 14. Total and partial averages for absolute difference and percentage of error between desired goals and system output in Experiment 4.

Building from the last experiences a more feasible painted solution space was provided in this experience. All the perimeter nodes were set to higher values, and the inner red focus assumed a more circular shape due to the fact that the parametric model could only generate circular skylights.

The solution found by the system generated a roof that achieves a good balance between facade height and overhang depth to better match the painted objectives in the building perimeter. Thus, the system lowered the facade in the narrow southeast and southwest zones and raised it in the northwest quadrant, with a deeper floor plan, generating a wavy rooftop with a central elevated bump area. Regarding the skylights, Painting with Light search mechanism only generated three skylights this time, probably due to the approximation of the interior red spots to a circular shape. The skylights match the location of the inner red painted areas and their radius are consistent with the defined target performance pattern. We can observe that the largest skylights are also placed in the highest area of the rooftop while the northwest skylight is smaller and placed in a location where the roof is relatively low, thus not needing to open as much to provide the amount of light needed to meet the desired targets for that area. Overall, the difference map indicates that the solution met the great majority of the desired results.

The % error average was kept below 20% ranging from 10.4% (for the 7% target areas) and 18,4% (for the 5% target zone). The percentage of area ≤ 25 % error increased dramatically to 77,24%. The percentage of area ≤ 10 % error is now 52.21%. The reported *NMBE* of -1.54% is significantly below the acceptance threshold of +/- 10%. Similarly, the *CVRMSE* registered a value of 20.17%, almost 10% below the acceptance criterion used for this index, 30%.

Results for this experiment produced a pattern that matches well the painted DF target distribution, and statistical error measurements for *NMBE* and *CVRMSE* were considerable acceptable. However, it is possible to further calibrate the system by manipulating the definition of the fitness function, by redefining other fitness and penalty functions, and or by the adjustment of an improved DF target.

5.5 Experiment 5

Figure 20 shows on the top two images, the pre-defined building design that Painting with Light had to find, as the result from that Daylight Factor simulation was used as the painted target to the system. The two images in the bottom of Fig. 20 show the design found by the Painting with Light, which is very close to the original design. Figure 21 shows the performance of the solution found by the system and the desired goals. Table 15 and 16 report the design variables of the pre-defined building, while Table 17 and 18 shows the values found by PWL. Table 19 reports the absolute deviation and the percentage of error between the pre-defined test geometry and the Painting with Light solution results. The overhang of both solutions is 0.5 m.

	Ptz1 [m relative to 3.5 m]	Ptz2 [m relative to 3.5 m]	Ptz3 [m relative to 3.5 m]	Ptz4 [m relative to 3.5 m]	Ptz5 [m relative to 3.5 m]
Crv1 (south)	2.4	0.2	0.2	-1.5	1.7
Crv2 (east)	-1.5	-1.5	-1.5	-1.4	-1.4
Crv3 (north)	0.8	-1.4	1.7	2.8	1.8
Crv4 (west)	-1.5	-1.5	-0.2	1.5	0.4

Table 15. Pre-defined building roof design variables.

	u [0-1]	v [0-1]	radius [m]
Skylight1	0.62	0.41	2.0
Skylight2	0.4	0.59	1.5
Skylight3	0.6	0.55	1
Skylight4	0.5	0.51	2.5

Table 16. Pre-defined skylight's variables.

	Ptz1 [m relative to 3.5 m]	Ptz2 [m relative to 3.5 m]	Ptz3 [m relative to 3.5 m]	Ptz4 [m relative to 3.5 m]	Ptz5 [m relative to 3.5 m]
Crv1 (south)	1.4	0.3	1.6	-1.5	0.1
Crv2 (east)	-1.5	-1.5	-1.5	-1.4	-1.3
Crv3 (north)	2	-1.4	0.3	2.8	4.1
Crv4 (west)	-1.5	-1.5	-0.2	1.5	0.4

Table 17. Experiment 5 building roof design variables.

	u [0-1]	v [0-1]	radius [m]
Skylight1	0.5	0.52	2.5
Skylight2	0.72	0.34	0
Skylight3	0.62	0.41	2.06
Skylight4	0.73	0.6	0

Table 18. Experiment 5 skylight's variables.

The overall shape of the roof that Painting with Light generated is very similar to the target building, what is particularly remarkable given the high degree of variation that the shape of the building can assume. As for the skylights, although the system only generated two skylights rather than the four skylights of the pre-defined solution it was capable to generate a solution which performance is very close to the painted one. In terms of % error in average the solution found by the system never goes beyond 10%.

The difference map of Figure 21 shows minimal deviation, except for a NE area which corresponds to the NE skylight of the pre-defined solution that the system did not generated. The percentage of area ≤ 25 % error increased to 94.4%. The percentage of area ≤ 10 % error increased to 67%. *NMBE* is -0.34% and *CRMSE* is 12.42%, yielding the best results of all the experiments that were conducted.

SE - goal geometry

NW - goal geometry

SE - result geometry

NW - result geometry

Figure 20. Perspective views of experiment 5 solutions. Top: Pre-defined solution which DF results were used as painted targets for the system. Bottom: Solution found by Painting with Light.

Figure 21. Experiment 5 results spatially mapped. Left: painted objectives, deriving from DF simulation of pre-determined geometry. Center: Daylight Factor simulation results of the solution proposed by Painting with Light. Right: difference (Δ) heatmap between painted objectives and simulation results.

Goal area	Absolute difference (average) [DF]	% error (average)
15-20%	1.5	8.3
10-15%	0.94	8.1
5-10%	0.8	9.1
0-5%	0.5	9.5
Total	0.8	8.8

Table 19. Total and partial averages for absolute difference and percentage of error between desired goals (pre-defined solution DF results) and system output in Experiment 5.

The area weights used in the penalty function could be one of the reasons for why Painting with Light did not found the two smaller skylights of the pre-defined solution. Because the area affected by them was quite small, they had little impact on the overall performance, thus making the system focus more on finding a better match to the larger areas of the painted goals. Figure 22 also shows that Painting with Light proposed minor differences on the facade, namely in NW quadrant, in order to compensate for the absence of the NW skylight that contributed to higher DF values on that corner of the building.

This final experiment showed that Painting with Light is currently able to find solutions reasonably close to the optimum, giving a well painted, feasible solution space.

6. Discussion

By comparing the results it is possible to have a better understanding of the impact of the different changes proposed in the PGDS calibration DoE.

Figure 22 shows the progress of the total percentage of area $\leq 25\%$ error, and Figure 23 reports the total percentage of area $\leq 10\%$ error. Figure 24 reports the evolution for the other metrics used in this study. The bar chart of Figure 25 shows the percentage of improvement of all the experiments in relation to experiment 1. However, because the first experiment was an initial trial that reflected the little knowledge of the user about how DF would behave in the implemented parametric model, a second percentage of improvement study was conducted between the experiments that had a similar DF target pattern; experiment 2, 3, and 4. Figure 26 shows the percentage of improvement pattern of experiment 3 and 4 when compared with experiment 2.

Figure 22. Total percentage of area $\leq 25\%$ error between painted targets and simulation results on each experiment. Below each experiment bar there is a plan that locates the areas where percentage error is above 25%. On the right, a heatmap shows the nodes between 25% to 35% of error, to visualize the deviation pattern for the area above 25%.

Figure 23. Total percentage of area $\leq 10\%$ error between painted targets and simulation results on each experiment. Below each experiment bar there is a plan that locates the areas in which % error is above 10%. On the right, a heatmap shows the nodes from 10% to 20% of error, to assess the deviation pattern for the area above 10%.

Figure 24. Difference between painted goal averages and DF results of the solutions found by Painting with Light in the five experiments.

Figure 25. Percentage of improvement in the different analyzed types of error of Experiment 2, 3 and 4 when compared with Experiment 1.

Figure 26. Percentage of improvement in the different analyzed types of error of Experiment 3 and 4 when compared with Experiment 2.

When comparing experiments one and two, only a slight improvement exists. The percentage of area ≤ 25 % error increased 9.53%, the percentage of area ≤ 10 % error 8.93%, the global % error average decreased 4%, the *NMBE* 14.76%, and the *CVRMSE* only 8.42%. This indicates that although providing a painted performance target with patterns closer to the real DF decay effect, the system still had difficulties in finding a good solution. This unsatisfying improvement is related with the averaging problems of the fitness function. Although a penalty function was used in experiment 2 the distribution of the penalties did not consider the spatial weight of each target areas. Thus, the resulting fitness function could be biased by some outlier data points that had little relevance in the overall performance.

By introducing weighting factors that express spatial bias in Experiment 3 a more noticeable improvement was reported. However, those improvements were still far from the acceptance threshold set for *NMBE* and *CVRMSE*. Compared with the first experiment, Experiment 3 improved 18.18% in % error, 67.58% in *NMBE*, and 32.22% in *CVRMSE*. The more expressive improvement in *NMBE* is normal because this index can suffer from the cancellation effect where a positive bias compensates for a negative one. This seems to be happening just by observing the high contrast between blue and red areas in experiment 3 solution DF results. *CVRMSE*, by the contrary, does not suffer from the cancellation effect reporting, thus, a more timid improvement. It is worthy to mention that regarding the average of absolute difference both experiment 2 and 3 appears to have a worse performance than experiment 1. However, absolute difference in this particular case is being biased by outlier data points that distort its average. In summary, when compared with experiment 1, experiment 3 seems to indicate that the approach could be further developed in the calibration the system.

However, when we compare experiment 3 to experiment 2 we can observe that the improvements are far from being relevant. For example, *CVRMSE* only improved 24%, and mean % error 10%. This indicates that the optimization could be suffering from inherent limitations of the search algorithm, thus, trapping the system in a local minimum. A comparison between the designs of the two solutions also supports this observation. In matter of fact the design solution of experiment 2 and 3 were very similar.

Experiment 4 reports a considerable increase of the system performance in finding the goals. Steeping the penalty function curve but especially using a more robust GA (SPEA2) produced a huge impact on the quality of the system's outcome. The percentage of area ≤ 25 % error was 77.24%, increasing 55.37%, 45.84%, and 33.19%, in relation to experiment 1, 2 and 3. Regarding the 10% threshold for % error, the percentage of area ≤ 10 % surpass 50%, registering 52.21% meaning that increased 45.62%, 36.7%, and 30.37% in comparison with experiment 1, 2 and 3 respectively. All the statistical indexes report significant improvements. When compared with experiment 1, the percentage of improvement raised to 97.10% in *NMBE*, and 73.96% in *CVRMSE*. When compared with improved experiment 2, this experience also shows remarkable progress, registering percentages of improvement that range from 53% (absolute difference) to 96% (*NMBE*). Both *NMBE* and *CVRMSE*, the most important indices in the calibration of the system, are comfortably in their acceptance range. Visually, the daylight performance pattern of the generated design is similar to the one painted by the user. This demonstrates that even resorting to averaging based processes, it is possible to achieve a solution that closely approximates a patterned performance target, with a high spatial resolution, using Painting with Light.

After the calibration done through experiments 1 to 4, experiment 5 tested if Painting with Light is able to find a pre-defined solution, to assess if a realistic and accurately painted solution would improve the prediction ability of the system. This experiment showed that when a more detailed solution space (i.e. a target spatial distribution that correctly conveys the granularity of light distribution and light decay patterns) is provided to the system, Painting with Light is able to find solutions that are very close to the desired targets. Experiment 5 yield the best results in all statistically indexes. In comparison with experiment 1, the improvements range from 63.64% (absolute difference) to 99.36% (*NMBE*), and when compared with experiment 2, improvement percentages are 69.81% (absolute difference) and 100.89% (*NMBE*). Even %error achieved the typical acceptable thresholds, below 10%. The more advanced statistical methods lie very comfortably in the acceptance range (*NMBE* registers -0.34% while *CRMSE* 12.24%). In terms of the architecture design solution, the system was able to find the overall shape of the roof and facades, with minor modifications on the northwest corner, and on the two main skylights. The two small skylights were not found by the system, but this is due to their small area when compared with the overall area of the sensor grid plane, what makes their presence less relevant in the overall performance of the building. Because the fitness function prioritizes larger areas of painted goals, the low effect of those small skylights was probably neglected by Painting with Light.

7. Conclusions

The research presented in this paper proposes a PGDS prototype, Painting with Light, for daylighting-based design that supports a visual description of the desired daylight patterns on an analysis workplane. The implemented target editor directly addresses current limitations on target definition of current PGDS for building performance metrics with a high spatial granularity by: (1) using brushing and linking techniques in Rhino through the drawing of polylines or closed NURBS curves; (2) controlling light decay pattern with sliders; (3) supporting diagnostic visualization techniques by automatically generating a difference heatmap between desired performance targets and simulation results; and (4) computing and displaying different statistics indexes that inform about the quality of the system's output solution, in term of its deviation from the painted objectives.

The experiments showed that by setting/painting target patterns within the feasible region and progressively refine the fitness function it is possible to calibrate the PGDS to produce high quality solutions. Experiment 4 results demonstrate this by presenting a solution with a *NMBE* of -1,54% and a *CVRMSE* of 20.17%, both in the range of acceptance typically used in hourly-based energy simulations: +/- 10% for *NMBE*, and 30% for *CVRMSE*.

The fourth experience also demonstrated that the robustness and quality of the search algorithm is extremely relevant in the production of good results. It was observed that Galapago's standard GA had a tendency to get stuck in local minima as the complexity of the problem increased, thus a more powerful evolutionary optimization algorithm, the SPEA2, was used. Finally the last experiment showed that when a good search mechanism is combined with a realistically painted goal pattern, Painting with Light is able to produce high quality solutions with a performance very close to the desired light values and spatial distribution.

Despite the quality of the solutions that can be generated by the proposed PGDS, different types of limitations can still be found, namely:

- User related limitations:
 - The user can poorly design the target space due to the lack of understanding how different daylight metrics work. Because goal programming was the optimization method used, the imposition of an unfeasible goal makes the search method very inefficient. [58]
 - Overconstrain the parametric model, making it impossible to meet desired performance targets.
- Workflow limitations:
 - In order to produce quality results, a calibration of the fitness function is still needed. This is also related with the main weakness of the goal programming method: the designer needs to devise the appropriate weights in order to better steer the optimization search; a task that, depending on the problem, may require some iterations or expert knowledge.
 - All the experiments were done just for one daylight metric, Daylight Factor. More trials are required to test the system with more complex daylight metrics.
- Interface limitations:
 - The method that controls light decay on the painting process was simplified to a linear strength factor implemented as a slider. This approach requires a deep intuition about light decay patterns from the user, thus, making the system more prone to human error.
 - The fitness function is based on weight factors that are related with the area of the different painted goals. This may lead the system to overlook small effects or light patterns that the designer has included into the goal space.
 - The statistical methods used in the assessment on the quality of the solutions could benefit from adequate visualizations that would to improve diagnostic tasks throughout the design workflow.

To address these limitations future work on the development of this tool is proposed. To solve the user limitations an expert target editor that would steer the painting of performance patterns to more feasible regions, or warn about the painting of unfeasible target areas, needs to be implemented. In the development of this expert editor, special functions that compute patterns of light propagation and decay for different daylight metrics need to be developed to effectively answer to the light decay control limitation. To address the workflow limitations, further experiments need to be conducted using Pareto-based approaches such as multi-objective GA. By making each target area an objective to the multi-objective Pareto GA, the system could improve its ability to predict small pattern variations, because it will no longer be biased by the weights related with target areas. Further experiences should also be performed in order to extend the daylight metrics used to annual climate-based metrics such as Daylight Autonomy, Useful Daylight Illuminance, and Spatial Daylight Autonomy. Finally, in order to improve evaluation tasks, automatic data visualizations of the statistics indexes should be implemented and further developed to facilitate comparison and design performance assessment tasks.

Acknowledgements

Work presented in this paper was developed as part of research project PCDT/AUR-AQI/103434/2008, Intelligent Shading Facades: Design, Prototyping and Testing of Innovative Shading Systems for Energy Efficient Buildings, funded by Fundação para a Ciência e a Tecnologia (FCT), Portugal.

References

- [1] Asadi, S., and Woo Geem, Z. "Sustainable Building Design: A Review on Recent Metaheuristic Methods." *Recent Advances in Swarm Intelligence and Evolutionary Computation*. Springer, 2015. 203–223.
- [2] Machairas, V., Tsangrassoulis, A., and Axarli, K. "Algorithms for Optimization of Building Design: A Review." *Renewable and Sustainable Energy Reviews* 31 (2014): 101–112.
- [3] Celani, G., and Vaz, C. "CAD Scripting and Visual Programming Languages for Implementing Computational Design Concepts: A Comparison from a Pedagogical Point of View." *International Journal of Architectural Computing* 10.1 (2012): 121–138.
- [4] Leitão, A., Santos, L., and Lopes, J. "Programming Languages for Generative Design: A Comparative Study." *International Journal of Architectural Computing* 10.1 (2012): 139–162.
- [5] Jakubiec, Alstan, J., and Reinhart, Christoph F. "DIVA 2.0: Integrating daylight and thermal simulations using Rhinoceros 3D, Daysim and EnergyPlus." *Proceedings of Building Simulation*. Vol. 20. 11. 2011: 2202-2209
- [6] Roudsari, Mostapha S., Pak, M., and Smith, A. "Ladybug: A Parametric Environmental Plugin for Grasshopper to Help Designers Create an Environmentally-Conscious Design." *Proceedings of the 13th International IBPSA*, Lyon, France, 2013: 3128-3135.
- [7] Crawley, Drury B, Lawrie, Linda K., Buhl, Walter F., Huang, Y. J., Pedersen, Curtis, O., Strand, Richard, K., Liesen, Richard J., Fisher, Daniel E., and Witte, Michael, J. "EnergyPlus: Creating a New-Generation Building Energy Simulation Program." *Energy and buildings* 33.4 (2001): 319–331.
- [8] Ward, Gregory J. "The RADIANCE Lighting Simulation and Rendering System." *Proceedings of the 21st Annual Conference on Computer Graphics and Interactive Techniques*. ACM, 1994: 459–472.
- [9] Kollat, J. B., and Reed, P. CA framework for visually interactive decision-making and design using evolutionary multi-objective optimization (VIDEO)." *Environmental Modeling & Software* 22.12 (2007): 1691-1704.
- [10] Winer, E. H., and Bloebaum, C. L. "Development of visual design steering as an aid in large-scale multidisciplinary design optimization. Part I: method development." *Structural and Multidisciplinary Optimization* 23.6 (2002): 412-424.
- [11] Winer, E. H., and Bloebaum, C. L. "Development of visual design steering as an aid in large-scale multidisciplinary design optimization. Part II: method validation." *Structural and Multidisciplinary Optimization* 23.6 (2002): 425-435.
- [12] Evins, R. "A review of computational optimization methods applied to sustainable building design." *Renewable and Sustainable Energy Reviews* 22 (2013): 230-245.
- [13] Gagne, J., and Andersen, M. "A generative facade design method based on daylighting performance goals." *Journal of Building Performance Simulation* 5.3 (2012): 141-154.

- [14] McNeel, R. et al. "Rhinoceros 3D." *Retrieved Jan 15* (2009).
- [15] Attia, S., Hamdy, M., O'Brien, W., and Carlucci, S. "Assessing Gaps and Needs for Integrating Building Performance Optimization Tools in Net Zero Energy Buildings Design." *Energy and Buildings* 60 (2013): 110–124.
- [16] Nguyen, Anh-Tuan, Reiter, S., and Rigo, P. "A Review on Simulation-Based Optimization Methods Applied to Building Performance Analysis." *Applied Energy* 113 (2014): 1043–1058.
- [17] Huang, Yu, and Niu, Jian-lei. "Optimal Building Envelope Design Based on Simulated Performance: History, Current Status and New Potentials." *Energy and Buildings* 117 (2016): 387–398.
- [18] Turrin, M., von Buelow, P., and Stouffs, R. "Design Explorations of Performance Driven Geometry in Architectural Design Using Parametric Modeling and Genetic Algorithms." *Advanced Engineering Informatics* 25.4 (2011): 656–675.
- [19] Granadeiro, V., Duarte, José P., Correia, João R., and Leal, Vitor M.S. "Building Envelope Shape Design in Early Stages of the Design Process: Integrating Architectural Design Systems and Energy Simulation." *Automation in Construction* 32 (2013): 196–209.
- [20] Caldas, L., and Santos, L., "Generation of Energy-Efficient Patio Houses with GENE_ARCH: combining an Evolutionary Generative Design System with a Shape Grammar" *30th eCAADe Conference Proceedings: Digital Physicality/Physical Digitality*, Prague, Check Republic, volume 1, 2012: 459-470.
- [21] Basbagill, John P., L. Flager, F., Lepech, M., and Fischer, M. "Application of life-cycle assessment to early stage building design for reduced embodied environmental impacts." *Building and Environment* 60 (2013): 81-92.
- [22] Basbagill, John P., L. Flager, F., and Lepech, M. "A Multi-Objective Feedback Approach for Evaluating Sequential Conceptual Building Design Decisions." *Automation in Construction* 45 (2014): 136–150.
- [23] Flager, F., Welle, B., Bansal, P., Soremekun, G., and Haymaker, J. "Multidisciplinary process integration and design optimization of a classroom building." *Journal of Information Technology in Construction* 14 (2009): 595-612.
- [24] Brown, N., and T. Mueller, C. "Design for Structural and Energy Performance of Long Span Buildings Using Geometric Multi-Objective Optimization." *Energy and Buildings* 127 (2016): 746-761.
- [25] Lin, Shih-Hsin E., and Jason Gerber, D. "Evolutionary Energy Performance Feedback for Design: Multidisciplinary Design Optimization and Performance Boundaries for Design Decision Support." *Energy and Buildings* 84 (2014): 426–441.
- [26] Lin, Shih-Hsin E., and Jason Gerber, D. "Designing-in Performance: A Framework for Evolutionary Energy Performance Feedback in Early Stage Design." *Automation in Construction* 38 (2014): 59–73

- [27] Welle, B., Haymaker, J., Fischer, M., and Bazjanac, V. "CAD-Centric Attribution Methodology for Multidisciplinary Optimization Environments: Enabling Parametric Attribution for Efficient Design Space Formulation and Evaluation." *Journal of Computing in Civil Engineering* 28.2 (2014): 284-296
- [28] Welle, B., Haymaker, J., and Rogers, Z. "ThermalOpt: A methodology for automated BIM-based multidisciplinary thermal simulation for use in optimization environments." *Building Simulation* Vol. 4. No. 4. Tsinghua Press, 2011
- [29] Diakaki, C., Grigoroudis, E., and Kolokotsa, D. "Performance study of a multi-objective mathematical programming modelling approach for energy decision-making in buildings." *Energy* 59 (2013): 534-542.
- [30] Ferrara, M., Fabrizio, E., Virgone, J., and Filippi, M. "A Simulation-Based Optimization Method for Cost-Optimal Analysis of Nearly Zero Energy Buildings." *Energy and Buildings* 84 (2014): 442–457.
- [31] Xu, Jun, Kim, Jin-Ho, and Koo, J. "A systematic approach for energy efficient building design factors optimization." *Energy and Buildings* 89 (2015): 87-96.
- [32] Fesanghary, M., Asadi, S., and Geem, Zong W. "Design of low-emission and energy-efficient residential buildings using a multi-objective optimization algorithm." *Building and environment* 49 (2012): 245-250.
- [33] Asadi, E., da Silva, Manuel G., and Antunes, Carlos H. "Multi-objective optimization for building retrofit: A model using genetic algorithm and artificial neural network and an application." *Energy and Buildings* 81 (2014): 444-456.
- [34] Murray, Sean N., Walsh, Brendan P, and Kelliher, D. "Multi-variable optimization of thermal energy efficiency retrofitting of buildings using static modelling and genetic algorithms—A case study." *Building and Environment* 75 (2014): 98-107.
- [35] Wright, Jonathan A., Brownlee, A., Mourshed, Monjur M., and Wang, M. "Multi-objective optimization of cellular fenestration by an evolutionary algorithm." *Journal of Building Performance Simulation* 7.1 (2014): 33-51.
- [36] Kasinalis, C., Loonen, R.C.G.M., Cóstola, D., and Hensen, J.L.M. "Framework for assessing the performance potential of seasonally adaptable facades using multi-objective optimization." *Energy and Buildings* 79 (2014): 106-113.
- [37] Schneider, S., and Donath D. "Topo-Metric Variations for Design Optimization: Introducing a Generative Model for Simultaneously Varying Metric and Topological Properties of Facade Geometry." eCAADe 2013: Computation and Performance—*Proceedings of the 31st International Conference on Education and Research in Computer Aided Architectural Design in Europe*, Delft, The Netherlands, September 18-20, 2013: 487-495.
- [38] Wang, J., and Caldas, L., "Selection of Energy Efficient Windows for Hot Climates Using Genetic Algorithms Optimization", *Proceedings of PLEA 2016, International Conference on Passive and Low Energy Architecture 2016: Cities, Buildings, People: Towards Regenerative Environments*, Los Angeles, July 11-13, 2016.

- [39] Andersen, M. "Unweaving the Human Response in Daylighting Design." *Building and Environment* 91 (2015): 101–117.
- [40] Andersen, M., Gagne, Jaime ML, and Kleindienst, Siân. "Interactive Expert Support for Early Stage Full-Year Daylighting Design: A User's Perspective on Lightsolve." *Automation in Construction* 35 (2013): 338–352.
- [41] Rakha, T., and Nassar, K. "Daylight as an Evolutionary Architectural Form Finder." *The Proceedings of the International Conference For Computing In Civil And Building Engineering–ICCCB*, 2010.
- [42] Futrell, Benjamin J., C. Ozelkan, E., and Brentrup, D. "Optimizing Complex Building Design for Annual Daylighting Performance and Evaluation of Optimization Algorithms." *Energy and Buildings* 92 (2015): 234–245.
- [43] Futrell, Benjamin J., C. Ozelkan, E., and Brentrup, D. "Bi-objective optimization of building enclosure design for thermal and lighting performance." *Building and Environment* 92 (2015): 591-602.
- [44] Becker, R. A., and Cleveland, W. S. "Brushing scatterplots." *Technometrics* 29.2 (1987): 127-142.
- [45] Monks, M., Oh, B. M., and Dorsey, J. "Audiooptimization: Goal-based acoustic design." *Computer Graphics and Applications, IEEE* 20.3 (2000): 76-90.
- [46] Andersen, M., Kleindienst, S., Yi, L., Lee, J., Bodart, M., and Cutler, B. "An intuitive daylighting performance analysis and optimization approach." *Building Research & Information* 36.6 (2008): 593-607.
- [47] Shea, K., Aish, R., and Gourtovaia, M. "Towards integrated performance-driven generative design tools." *Automation in Construction* 14.2 (2005): 253-264.
- [48] Wetter, M. "GenOpt - A generic optimization program." *Seventh International IBPSA Conference, Rio de Janeiro*. 2001.
- [49] Caldas, L. "Generation of Energy-Efficient Architecture Solutions Applying GENE_ARCH: An Evolution-Based Generative Design System", *Advanced Engineering Informatics*, 22.1 (2008): 59-70.
- [50] Caldas, L.G., and Norford L. K. "A design optimization tool based on a genetic algorithm." *Automation in construction* 11.2 (2002): 173-184.
- [51] Agarwala, A., Dontcheva, M., Agrawala, M., Drucker, S., Colburn, A., Curless, B., Salesin, D., and Cohen, M. "Interactive digital photomontage." *ACM Transactions on Graphics (TOG)* 23.3 (2004): 294-302.
- [52] Reinhart, C. F., and Walkenhorst, O. "Validation of dynamic RADIANCE-based daylight simulations for a test office with external blinds." *Energy and Buildings* 33.7 (2001): 683-697.
- [53] Hubler, D., Tupper, K., and Greensfelder, E. "AB-10-C028 Pulling the Levers on Existing Buildings: A Simple Method for Calibrating Hourly Energy Models." *ASHRAE Transactions* 116.2 (2010): 261.

- [54] Diamond, S. C., and Hunn, B. D. *Comparison of DOE-2 computer program simulations to metered data for seven commercial buildings*. No. LA-UR-80-2791; CONF-810101-2. Los Alamos Scientific Lab., NM (USA), 1981.
- [55] Coakley, D., Raftery, P., and Keane, M. "A review of methods to match building energy simulation models to measured data." *Renewable and Sustainable Energy Reviews* 37 (2014): 123-141.
- [56] ASHRAE Standards Committee. "ASHRAE guideline 14-2002 Measurement of energy and demand savings." (2002).
- [57] Standard, ASHRAE "90.1, 2007." *Energy Standard for Buildings Except Low-Rise Residential Buildings-ANSI/ASHRAE Standard* (2007): 90-1.
- [58] Coello, C. A. C. "A comprehensive survey of evolutionary-based multiobjective optimization techniques." *Knowledge and Information systems* 1.3 (1999): 269-308.
- [59] Ijiri, Y. *Management goals and accounting for control*. Vol. 3. North Holland Pub. Co., 1965.
- [60] Charnes, A., and Cooper, W. W. "Management models and industrial applications of linear programming." *Management Science* 4.1 (1957): 38-91.
- [61] Rutten, D. "Evolutionary Principles Applied to Problem Solving." I Eat Bugs For Breakfast. 4 Mar. 2011. Web. 22 June 2015. <<https://ieatbugsforbreakfast.wordpress.com/2011/03/04/epatps01/>>.
- [62] Zitzler, E., Laumanns, M., and Thiele, L. "SPEA2: improving the strength Pareto evolutionary algorithm." Zurich, Switzerland, Swiss Federal Institute Technology (2001).