Incorporating human factors in adaptive lighting systems using a dynamic human-object interface

Cesar Armando Torres[1], Advisor: Joseph Paradiso[2], Supervisor: Matthew Aldrich[2]

[1] Department of Art, Department of Computer Science, Stanford University, Stanford, CA 94305

[2] Responsive Environments Group, MIT Media Lab, Massachusetts Institute of Technology, Cambridge, MA 0213

August 16, 2011

Abstract

Although adaptive solid-state lighting systems minimize energy consumption, current systems lack human factor design necessary for optimizing user interaction. A series of studies observing adaptive lighting usability found that despite greater user control, occupants were less likely to interact with their lighting environment leading to workplace luminance levels significantly below recommended standards limiting user task performance and safety [7, 14]. This study discusses improvements in lighting system usability using a vision-based system with light invariant color-tracking. We introduce the interaction method of transferrable control onto physical artifacts which uses the natural relationships between users and their object environments to facilitate changes within a digital system. Furthermore, a new luminance mapping with the CIELAB color space is proposed to model perceived brightness while energy output is minimized using a previously developed closed loop linear optimization protocol [1].

1 Introduction

Current adaptive lighting systems (ALS) have shown considerable energy savings over the course of their inception with savings ranging from 65% to 90% energy output. With over 22 percent of the current electricity budget in the United States devoted to lighting, these systems serve as a model for future sensor network-driven energy saving solutions [11]. However, these improvements provide significant drawbacks - the most critical dealing with unhealthy user behavioral patterns. Studies have found that many occupants choose levels of working plane illuminance that fall significantly below CISBE [Chartered Institution of Building Services Engineers] Code recommendations denoted in Table 1 [3, 14] . This ultimately results in poor lighting conditions over long periods of time (over a month between interactions) detrimental to user performance as well as potential energy savings [14].

Although lighting systems impart more user freedom and control of luminance levels [3], this control must be mitigated with interaction design to promote healthy lighting conditions. Lighting controls culminate into two different types of interfaces: power controllers (switches, dimmers, presets, etc. . .) and sensor controllers (photosensors, occupancy sensors, cameras, etc. . .). The effective combination of these controllers to facilitate both energy minimization and user satisfaction is the focus of this paper.

Table 1: CIBSE recommendations for lighting levels in office settings (2002)

Description	Illuminance (lux)
Filing, copying, etc. . .	300
Computer-based work, reading	500
Technical drawing	750

1.1 Related work

Dugar & Donn [6] recently compiled an overview of current lighting controllers in adaptive systems and described their limitations and affordances. Among the most enabling characteristics of an ideal lighting system was the need to displace cognitive resources towards a perceptual-motor load. The disconnect between user knowledge of energy utilization methods and system knowledge of user preference is a growing area of lighting technology research. Mozer addresses this issue in the development of a prototype smart house that regulates gas and electricity output by learning user preference using neural networks [16]. In mediating user control and autonomous control, Guillemin & Morel [10] proposed a user control feedback loop which used a Genetic Algorithm (GA) which would learn user lighting preferences and maximize on energy conservation based on the presence of the user. Although this approach minimized energy output, it was postulated that the learned lighting preference was satisfactory on the premise that the user did not further interact with the system, which as previously stated is a result of behavioral patterns.

Interaction design with intangible medias such as light and sound has become a growing area of research within HCI. In the case study of the audio-scape environment Signal Play, Williams identified a relationship between bodily communication and interactions with space [19]. Williams argues that space and social action are closely related to the arrangement of bodies, artifacts and activities. It was noted that participants tended to focus on physical artifacts [which triggered changes in the auditory environment] as the source of sounds and regarded the digital system [the speakers that created the sound] as transparent. We take this transfer of focus to physical artifacts as a major interaction design exploration with light systems.

Outline Due to the multiple areas under consideration within this paper, we have separated content into four sections: Section 2 provides an overview of human-factors in adaptive lighting including a review of usability studies and physiological factors. Section 3 describes the use of human-object relationships for stronger object action-potential couplings based on behavior. Section 4 is a technical design overview of a vision-based adaptive lighting system with human-factored design including: modeling luminance to perceived brightness, minimizing energy output with light/dark adaptation constraints, and a gesture recognizer using a lighting invariant color tracking technique. Section 5 presents a prototype human-object interface for adaptive lighting systems. Lastly, Section 6 presents future work and Section 7 features a discussion and conclusion of the study.

2 Human factors in lighting

We review seven studies on adaptive lighting control usability spanning several sample sizes, scopes, and time periods compiled from the *Lighting Research and Technology* journal [3, 6, 7, 14, 15], *Journal of Illuminating Engineering Society* [4], medical research in *Clinical and Experimental Optometry* [2], and previous work [1, 22]. A wide array of adaptive system controls are discussed within these papers including GUI-based, sensor-based, user-controlled, and mixtures in between. We review usability issues concerning system interaction frequencies outlined in Table 2.

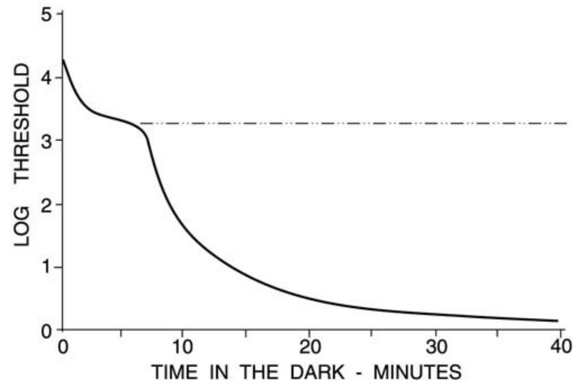

Figure 1: Dark adaptation curve plotted against time and stimulus intensity. The dashed line indicates the mesopic/photopic threshold.

Figure 2: Simultaneous brightness effect. Stimulus looks significantly brighter when the proportion of luminance surrounding inner rectangle to the inner rectangle is large. All inner rectangles are photometrically equivalent.

Light and dark adaptation When changes in lighting conditions occur, our eyes experience a physiological change known as light adaptation under changes to well-lit lighting conditions and dark adaptation with changes to low lighting conditions. These changes are attributed to sensitivity thresholds of rods and cones in the eye which activate at different levels of luminance [9, 21]. These physiological changes takes as much as ten minutes for light adaptation and can range from twenty to thirty minutes for dark adaptation as noted in Figure 1. This can lead to varying perception of light intensity and discomfort; we factor these adjustment periods by constraining the energy optimization protocol to follow natural patterns of light and dark adaption in Section 4.

Brightness constancy User input of ALS lighting levels has been the main form of system manipulation. However, while the energy optimization follows a linear adjustment of energy output, human per-

Table 2: Major observations in adaptive lighting system usability

Perceived interaction is high; actual interaction is low	Occupants, when surveyed about their perceived use of their ALS, report a considerably greater frequency of use compared to observed long term behavior [7, 14]. Interaction occurs typically less than once a month.
Optimal lighting levels are higher	Objectively defined optimal task lighting is greater than the subjectively-defined task lighting. [2]
Switching only occurs with arrival and departure	Switching behaviors are defined by the anticipation of intended use, where a set office lunch period results in greater switching behavior over time as opposed to leaving the workspace to run an errand with the intention of returning to the workspace in a short interval of time [14].
Interface complexity and reverting to the old ways	Presets are rarely used. If the user doesn't understand the system, they revert to typical patterns of use of conventional systems [7].
Flexibility is constrained to current interface language	Since light is intangible, it suffers from no natural mapping of its interface and must thus rely on past interfaces i.e. switches, dimmers, and presets. [6].
Control is overrated?	The ability to control lights does not affect subject mood, alertness, or performance [4]. However, users carry more positive attitudes adaptive lighting control and satisfaction with illuminance [15].

ception of brightness experiences several phenomena that can hinder optimal lighting optimizations. In fact, luminance, or the amount of reflected light entering the eye, *does not* correspond to the perceived brightness of objects.
Instead, brightness relies on the lateral interactions of retinal ganglion cells [8], implying that brightness depends on a proportion of luminance from the non-foveal region to the foveal region in question. Figure 2 illustrates this effect. Any target predominantly surrounded by an area of higher luminance should look darker than the same target predominantly surrounded by an area of lower luminance [13]. As such, the brightness of a user defined region is dependent on the brightness of the overall work area. We take this into account in our system design (Section 4.1) and employ a L*a*b*-based brightness model to adjust system lighting levels based on the proportion of region-to-surround brightness.

3 Interaction design

Several studies examined control system interfaces and concluded the following: a user-friendly interface is needed [6, 7]. However this problem is constrained by the inherent intangibility of light. In other words, there does not exist a natural physical mapping that can convey to the user the intended role of the lighting system's controls. Thus, the formulation of a natural mapping between a light system and the real world is defined by the "switch", the "dimmer"[rotation, linear], and less commonly the "preset" button. A similar problem is found in consumer electronics, where although buttons afford pushability, they only partly convey their intended purpose. These types of interfaces rely on semantic clues, such as making buttons on an electronic arrow-shaped to convey increase/decrease in some stimulus value. This approach however is considerably more cognitively demanding and hence *unfriendly.* In the case of lighting system control, it is especially difficult to satisfy usability heuristics such as the recognition of system operators — a switch may be recognizable and afford a binary operation, but the object controlled by said switch is ambiguous. With trial and error, this uncertainty can be resolved; however this can cause unwanted behavior and cognitive loads (searching for changes in system, turning on the wrong light, etc...) which increases exponentially as the amount of controls increase within a given control. "Presets", or predefined system configurations, are even more inflexible as they greatly increase the risk factor associated with changing a larger group of objects.

These incongruencies with inherently intangible contexts has been a significant research problem in the field of interaction design. Djajadiningrat argues that usability lies not in communicating the necessary action, but instead shifting attention to strengthening the coupling between the action and the feedback to maximize action potential through physical objects i.e. tangible interfaces [5]. This approach takes focus away from communicating the necessary action towards communicating the purpose of the action (feedforward) and associating more strongly the action and the feedback (inherent feedback). We take these metrics into account in selecting an appropriate interface.

3.1 Gestural interface

The gesture is one of the most natural interfaces and coincides with a low cognitive load and a greater motor load characteristic of an optimal lighting interface. Gestural interfaces augment the coupling of perceptual and motor skills with interactions with technology [6, 20]. However, in the case of conferencing technologies, gestures lose much of their effectiveness across video connections [19]. Gestures occur between bodies of communicating partners which unfortunately does not occur naturally with digital systems. Although the use of gestures would create a unique system, it would not elicit natural participation needed for increasing frequency of interaction with adaptive lighting systems. Furthermore, determining whether a user was eliciting a system change or merely exhibiting natural gestural movement could result in unwanted system responses.

However, coupling gestures with an external interface allows for a natural mixture of gesture and use of the color tracking methodology popular in gesture recognition. Our prototype of human-object interaction uses two indicators of participation: touching an object (with a natural pointing gesture), or grabbing the object (occlusion of a color cue). We also incorporated usability heuristic such as exiting system manipulation as a closed fist. Implementation of a vision-based gesture recognizer is discussed in Section 4.3.

3.2 Human object interface

Natural mapping is difficult to associate with lighting without referring to the semantic relation that has been placed with current lighting interfaces. The basic idea behind this model is to create meaning through the interaction with an object. In other words, improve the action-potential of an object and make it more "grabable". There are several techniques such as using visual and tactile properties such

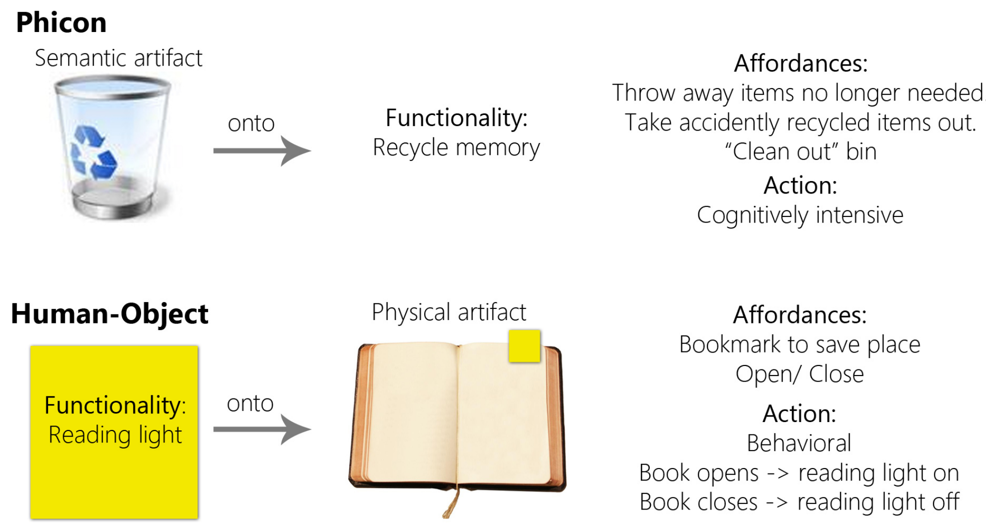

Figure 3: Phicon vs. Human-Object

as weight, texture, sound, material, etc.... Physical icons, or phicons, are embodiments of a semantic meaning onto a digital object (such as the recycling bin in desktops), and is a commonly used method of drawing upon metaphor and conveying potential actions. We take a new approach that *embodies a functional meaning onto a physical object.* The natural relationship with the user and the object serves as a new interface that minimizes on cognitive load and allows for stronger action potentials to be achieved based on the user's natural behavior with an object (see Figure 3).

For instance, suppose that the function of a lighting preset is 500 lux or "reading light". If the user transfers control of the preset functionality onto a book, a highlighter, or to their reading glasses, this would greatly increase user participation through their inherent behavior patterns (reaching for their glasses to read) and allow for optimal lighting conditions. Or as another example, suppose that a cell phone is commonly taken with a person in departure from their office workspace. This could very well be a sensory rich relationship that can control the on/off behavior of lights and correlate strongly to natural habitual behavior between a user and the system. Furthermore, a natural coupling is achieved with cell phone presence and light. Should a user see his or her light on, it is an indication that the user has forgotten his or her cell phone.
Implementation of transferrable control using optically-trackable cues is described in Section 4.3 and Section 5. Inherent feedback is naturally conveyed through light stimulus. Should users make a change to the system, their actions are reflected in the light status of the system. We explore other forms of feedback such as multimodal stimuli in Section 5.1.

4 Human factor design in ALS

Light sensors are usually interspersed in a traditional adaptive lightings system and collectively contribute data on ambient and controlled illuminance levels. Some systems have sensors near sources of light and provide variable light measurements depending on the location of the workplace of interest. In previous work, we have used sensors directly on the workplace surface such as pucks [1]; however, this method is requires that the sensor not be occluded and has calibration protocols that can hinder usability.
. The use of a camera in adaptive system control has many advantages including removing the need for obstructive photosensors on the workplace or region of interest. Addingly, they are cost-effective technology with potential for many applications including security, computer vision applications, and line-of-sight measurements[22]. We used a CMOS (Complementary metal oxide semiconductor) webcam sensor to measure luminance incident to the plane of interest based on a previously designed ALS[17, 22]. The same testbed lighting system as previous studies was used consisting of four IW2 BLAST 12 LED fixtures with 24VDC, 150W power source, and a UDP output-configurable communication [22].

4.1 L* as a linear model of brightness perception

Based on the physiological factor of brightness constancy, we found that photometric quantities did not correlate to brightness perception. In order to better correlate to user perception of brightness, we used the CIE L*a*b* as a perceptive model and calibrate the system using systematic transforms.
The CIE L*a*b* is the principal color space used throughout the framework design in order to best model perceptual changes in luminance [12]. L*a*b* is a color-opponent space derived from a modified Steven's law representation to characterize perceptual differences in color and brightness described by:

$$S = kI^b \tag{1}$$

where the perceptual strength of a sensory stimulus S is related to the scaling constant k, the physical intensity I, and the exponent b which is specific to the just noticeable differences (jnd) of a particular sensory experience. Brightness perception has a $b > 1$ indicating that small magnitudes changes create larger perceptual differences as opposed to the same change in larger magnitudes.

The L*a*b* space can be separated into three channels: where L* represents brightness (which carries much of its information from the green channel in the commonly used RGB color space); a* and b* model chromatic opponency observed in color perception and likewise follow a Steven's law representation that maps to uniform changes in perceived color. This color space allows for a Euclidean distance calculation to gather linear changes in perceived color (ΔE):

$$\Delta E_{ab}* = [(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2]^{\frac{1}{2}} \tag{2}$$

The perceived change in brightness is similarly computed as the change in L*:

$$\Delta E_{l}* = \Delta L^* \tag{3}$$

Mapping illuminance to L* The functional luminance of the L* channel foliates color space into constant planes of luminance, such as the alychne visually portrayed by Figure 4. Although luminance is extracted using the CIELAB color space, pure colors, as Hering defined, have no brightness [12]. This implies that an L*a*b* intensity can be decomposed into the luminance of the color and the pure color on a grayscale. Within this model, we lose the ability to distinguish between luminance data and color data. In order to parse luminance, we developed a novel approach to separate these two components. We assert that spectral white contains only luminance data and we use this fact as an extraction method for luminance data from pure color representation in L*.

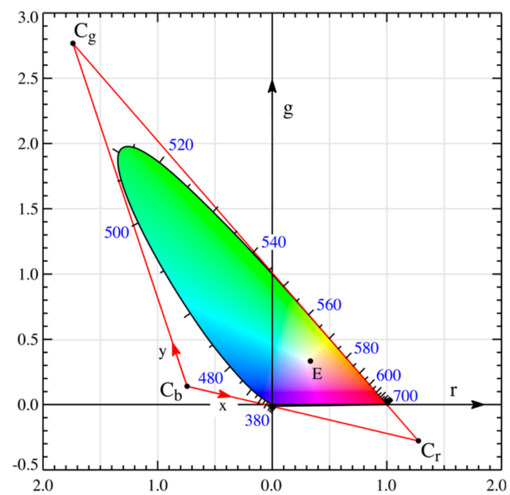

Figure 4: Alychne defined by the line connecting C_b and C_r

To better conceptualize this problem, suppose that we have a region R that consists of a white plane. We take the average L* value and find it to be close to 100. If a red box (L* = 53 under the same illuminant) enters the region, it would result in a lower average L* value. Color constancy works on both ends of this equation. Although the light has a different spectral distribution, we still interpret it to have the same brightness as the whitepoint [8]. Thus, we can say that although the box has L* = 53, because of the whitepoint under the same illuminant = 100, we can say that every red value should have 47 added to it, allowing for again the global average luminance to be equal to 100 as represented in Figure 4.1.

We define R to be a set of all pixel within a rectangular region from a camera capture to encompass a smooth white uniformly lit surface and convert to CIE L*a*b* color space. We extract the L* channel and take the mean value over the region R.

$$L^* = f_{RGB \to L*}(x) \tag{4}$$

$$\bar{R}_{L^*} = \frac{1}{wh} \sum_{(i,j)=(x_0,y_0)}^{i=w,j=h} R_{L^*}(i,j) \tag{5}$$

Let light i be at maximum lumen output such that:

$$R_{L^*,max} = L_{i,max} \tag{6}$$

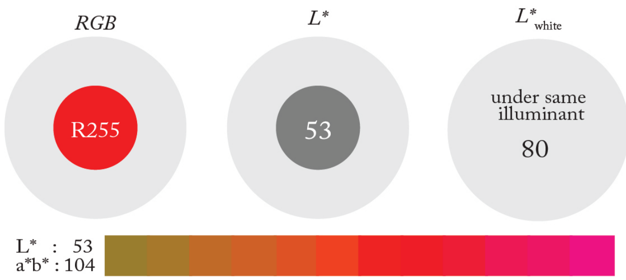

Figure 5: Pure color representation under L*a*b* space. The pure color red is represented in the L* color space as 53, as are all the colors in the color bar. Based on the luminant, we can transform the pure color red and all its equivalents to correlate with a white point intensity value under the same illuminant.

In the same region R, we take a color gamut and map values from the a* and b* channels using their Euclidean distance:

$$\text{purecolor}_{ab} = R_{ab}* = [(\Delta a^*)^2 + (\Delta b^*)^2]^{\frac{1}{2}} \tag{7}$$

and map these values to the differences between the L* value and a whitepoint.

$$\text{purecolor}_{ab} \mapsto R_{L^*,white} - R_{L^*} \tag{8}$$

We repeat this process for every illumination level for light i so that for any L* value there exists a mapping from the purecolor to the luminance of the whitepoint.

$$\text{purecolor}_{ab,L*} \mapsto R_{L^*_{ab},white} - R_{L^*_{ab}} \tag{9}$$

At this point, for any given region R, we apply this mapping such that every pixel P in R' is the sum of the original pixel L* value and the identity of function $\text{purecolor}_{ab,L*}$ for pixel P. The process can be visualized in Figure 6.

Calibration of luminaires to L* intensity values Solid-state lighting has the unique property of regulating lumen output. This regulation is luminaire specific, although most follow a form of Equation 1. In order to reduce error and accurately model visual perception, we define a protocol to map luminaire output to L* intensity values from CMOS webcam sensor. The process involves measuring L* values for all light intensities of a single luminaire and applying a linear transform based on the maximum output to obtain the mapping for any similar luminaire. Setting L_i to its maximum lumen output, we plot $\bar{R}_{L^*}$ for every lumen output of light i. We define the mapping from lumen output to L^* intensity at region R_i as:

$$f_{L_i,R_i} : X \to Y \tag{10}$$

where X is the lumen output and Y is the L* value. For every light n of similar structure to L_i, the lumen to L^* mapping function can be found by scaling the f_{L_i} such that the scaling factor is the max output over the maximum output of the original calibration function.

$$f_{L_n,R_n}(x) = \frac{X_{i,\max}}{\arg\max_x f_{L_i,R_i}(x)} f_{L_i,R_i}(x) \tag{11}$$

this process is invertible so that for any L* value, we can obtain a percent output (e.g. an observed L* value of 50 needs 78% output from luminaire A.)
Assuming consistency with luminaire output, this method has several limitations - the most severe dealing with highly reflective surfaces. For our implementation, we assume a consistent matte surface, however deviation from this model could be used as feature vectors for a surface recognition classifier which we leave for future work.

4.2 Dark and light adaptation constraints with energy optimization

In order to solve the optimization problem for minimizing energy output, we use the simplex linear programming algorithm defined in previous work [1, 17, 22]. All input values of the optimization routine are in the L* domain. The total energy consumption and the cost function of the system is defined as:

$$E_{system} = \sum_{i}^{n} E_{i,max} \cdot c_i \tag{12}$$

where E_i is the energy consumed by light i at maximum output and c_i is the weighted contribution of each light to the E_{system}.
Thus, the objective function is the energy consumption of the light system.

$$\min \zeta = \sum_{i}^{n} E_{i,max} \cdot c_i, c_i \geq 0 \tag{13}$$

We also add a special constraint to account for light and dark adaptation. For rapid changes in light, or differences greater than ϵ, we add the constraints

$$c_i \leq 1$$

and

$$U_L^* = \begin{cases} f_{\text{light adaptation}} & \Delta\epsilon > 0 \\ f_{\text{dark adaptation}} & \Delta\epsilon < 0 \\ \text{user input} & \text{otherwise} \end{cases}$$

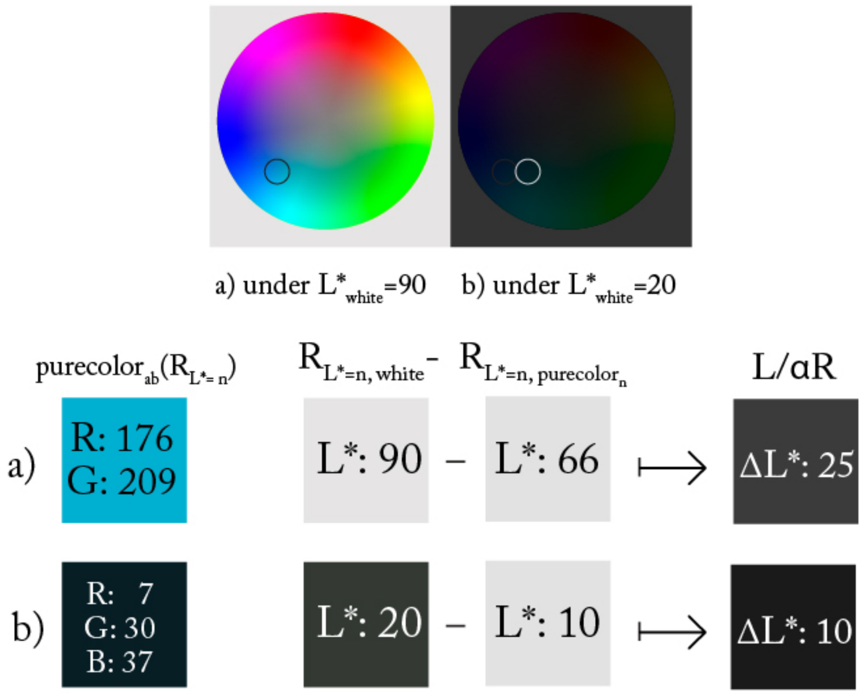

Figure 6: Extraction of pure color from the L* color space and the mapping of those values through a reference white point. Two point values being referenced are from the same location on the color gamut. Each color gamut is uniformly lit with L* corresponding to 90 and 20 respectively.

for

$$U_L^* = (\sum_i^n L_{i,max} \cdot c_i) + L_{ambient}$$

where U_L^* is the desired luminance level inputted by the user. The routine for updating the a closed loop is diagrammed with Figure 7. This implementation allows us to account for error in the lumen to L* mapping by adjusting the needed light L by the difference between user-defined level and L_L^*, or ΔL.

The dark adaptation function $f_{\text{dark adaptation}}$ is the mapping of maximum intensity values from the change ϵ to time elapsed. Dark adaptation from max illuminance to no illuminance should span a maximum of thirty minutes. Similarly, $f_{\text{light adaptation}}$ should range at max one minute. In order to account for brightness constancy, we further define the user input as

$$\text{new user value} = \text{surround luminance} \cdot \text{user ratio}$$

$$U_{L^*} = \bar{R}'_{L^*} \cdot U_\rho \tag{14}$$

where ρ is the preferred brightness constancy ratio.

Special cases Although not represented in the Figure 7, logical cases exist to ensure that the optimization problem sent to the simplex block is not over constrained. For instance, should the $L_{system,max} = \sum_i^n L_{i,max}$ be less than the desired user level U_L^*, then the system would sent the vector $\mathbf{c} = \begin{bmatrix} 1 & \dots & 1 \end{bmatrix}$ where $\mathbf{c} \in \mathbb{R}^n$ for maximum lumen output.

4.3 Gesture Recognition

Gesture recognition can be decomposed to two separate areas: feature extraction and classification. We take into special consideration recognition under extreme lighting conditions.

Color extraction Also aligned by brightness, a variation of the L*a*b* color space is known as the Luv space which uses the same L channel; however its color components constitute the relationship between chroma and hue. They prove especially important in design of color-tracking application with dynamic uniformly illuminated lighting . Back projection is a color segmentation method used to determine the fit of a pixel to a distribution of pixels in a histogram model. This differs from a conventional point-sample color segmentation which compares a given pixel to a selected pixel with a tolerance in order to identify pixels of that color. Since flesh cannot be accurately represented within a set interval of any one color, a histogram approach allows for greater inclusion of lighting variations in a target color. By converting

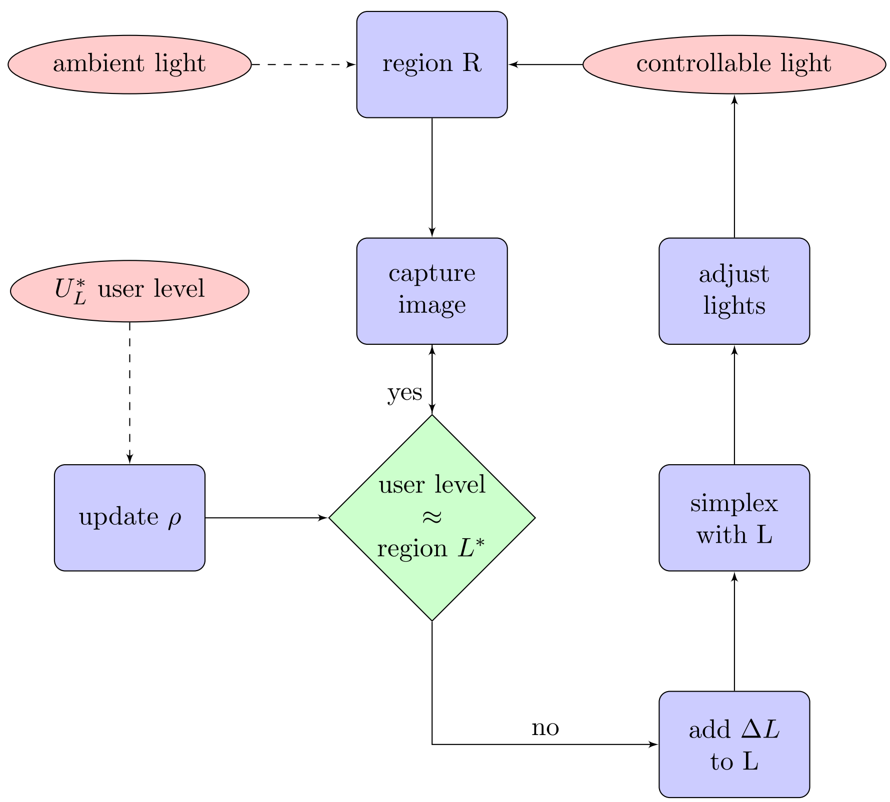

Figure 7: Closed loop optimization of an adaptive lighting system. L is defined to be the sum of all L^* output of all lights in the system at time T. Dashed lines represent outside input.

into the Luv color space and using u and v channels, lighting robust color segmentation can be obtained through a Bayesian extraction described by:

$$P(T \mid C) = \frac{P(T)}{P(C)} P(C \mid T) \tag{15}$$

where T is the probability that a pixel is the target color and C is the color of the pixel. This is the most commonly used method for hand segmentation. Several extensions have been made including the inclusion of an inverse probability such that a histogram of non-target colors provides further segmentation [18].

Hand feature extraction After color segmentation, a set of skin colored regions remain. In order to extract the hand of interest from this set, we define a series of characteristics that the regions must satisfy. We begin by applying a connect blobs algorithm, and remove skin candidates with areas greater than an experimentally derived threshold. We then proceed to extract the contour of the region using the convex hull method where the region R_i is decomposed into a sequence of clockwise points $P_1, P_2, \ldots, P_n$. We take into account the natural physiology of an outstretched hand and follow, with some variation, the outstretched hand detection method [18]. We use a simplified outstretched hand detector that calculated the best fit ellipse which corresponds to fingertip locations. We calculate fingertips as the point P of the convex hull sequence and use region moments to calculate the center of the region. We then select the region with the minimum square error of the fingertip distance to the radius of the best fit ellipse for R_i such that:

$$\arg\min_i \mathrm{SE}(\hat{\phi}_i) = E[(\hat{\phi}_i)^2] \tag{16}$$

where the error is the distance between a point on the convex hull and center of the best fit ellipse:

$$\hat{\phi}_i = R_{\text{ellipsecenter}} - P_j$$

Consequently, we also define the largest error within a region:

$$\arg\max_{P_j} \hat{\phi}_{min} = R_{\text{ellipsecenter}} - P_j \tag{17}$$

to be the pointing finger.

Library preparation We prepared a library of gestures by first defining a common metric for extraction. Gestures incorporated are open hand, pointing finger hand, and closed fist hand. We begin by scaling the image to 100 x 100 pixels with bicubic interpolation using a 3x3 kernel. We apply a Gaussian filter and calculate the Jacobian to acquire $\frac{\partial}{\partial x}$ and $\frac{\partial}{\partial y}$ which we then separate into a two dimensional histogram $H_{\text{Lib},i}$ for each image from i to n.

Classifier After extracting the hand vector, we are ready to build the classifier for gesture recognition. Let the recognizer be defined by S. In order to extract and process the feature vector, we use a orientation-based classification methodology for hand recognition [20]. This method takes the directional derivative of the hand vector and distributes them into bins on two dimensions. For our implementation, we extended the scope of this method by stretching the hand on its horizontal axis (perpendicular to the fingers) since the largest amount of orientation data is contained on that dimension. We generate a two-dimensional histogram H and compared it to the previously generated library, Lib.

$$\arg\min_i H - H_{\text{Lib},i}, i \leq n \tag{18}$$

We add a weighted vector constructed from Equation 16 to help distinguish between similar gestures.

5 Prototype

5.1 Overview of vision-based ALS system

Table 3 models the underlying class structure in the implementation of the ALS system.

Light and dark adaptation constraint in energy optimization The rate of dark adaptation is dependent on the time of exposure under photopic conditions known as bleaching and the overall lighting level. A closer look at Figure 1 reveals that rods begin changing in sensitivity around the 10 minute mark (under full bleach) and increase over the course of 20-30 minutes. Past the photopic/ scoptic threshold, cones take over in about a minute and continue to improve in visual acuity and color over the next 10 minutes. We integrated these constraints in the development of the ALS system in Section 4.2. If the system experiences a change in photopic/scoptic condition, it will dim to the desired intensity over the maximum course of a minute (photopic to scoptic) or thirty minutes (scoptic to photopic) depending on the change in intensity.

Multimodal cues To strengthen the association between lighting levels and sound, we utilize a MIDI sound generator to associate motions with the dimmer functionality. For instance, for setting the lighting intensity, we added an incrementing tone to correlate with the intensity of the light. In our usability review, we found that a significant amount of users underestimate optimal luminance levels. We nonlinearly mapped the tone values such that higher tones (associated with higher luminance levels) would occur later so as to influence more powerful light settings.

Table 3: Adaptive Lighting System prototype

Hand	Loads a gesture recognizer from a preassembled gesture library and classifies hand. Uses a skin lookup histogram to track hand position, size, bounding rectangle, and pointing finger.
CTracker	Creates a color lookup table, back projects, and then creates CObjects of all regions of a specified area range.
CObject	A colored object that keeps track of whether it is being pressed, presence, hue and location. On each update, its stored CFunction.
CFunction	Acts as an ActionListener and incorporates the functionality mechanisms specified in Figure 4.
LightingSystem	Updates the system lights and contains the simplex object to optimize lights. Stores total energy used.
LightSocket	Controls individual light structures. Keeps track of light map L* and current lumen output of a fixture at time t.
Simplex	Implementation of simplex linear programming algorithm with added light adaptation and dark adaptation constraints

Proportional luminance measurements By mapping luminance to L*, we were able to manipulate light intensities linearly based on the perception of brightness. Furthermore, we took into account simultaneous contrast by optimizing on the ratio of luminance at a region to luminance of the surrounding region (such as a illuminant appearing brighter in a dark office than a bright office).

5.2 Human-Object Interface

Post-Its as optical cues In order to construct a human-object interface, we found that Post-Its are highly optically-trackable and have natural affordances associated with personalized memo creation and demarcation of space. Their low cost also presented an added benefit for high design iteration. We incorporated functionality mechanisms described by Figure 4.

Natural mapping Post-it actors could be defined to control a light structure and naturally map onto the arrangement of the system and control individual lights. Furthermore, because Post-its, or the objects attached to them, are physically configurable, a natural mapping based on geophysical location is achieved.

Presence detector Based on the presence of a color cue, the system triggers preset values. We attach color cues to: an eraser for technical drawing presets; the spacebar and mouse for computer lighting; books, bookmarks, and folders for reading presets; cell phones and watches for on/off levels. Should occlusion occur or multiple triggers cause rapid changes in lighting levels, a tolerance is added to the amount of time a stimulus is present. If several presets trigger changes in the system, then the preset with the highest value is selected.

Region selector In order to optimize energy output, we defined a default region for the system to take light measurements. However, we also added the ability to select a region with switching on a region select cue and with a pointing gesture selecting the region of interest on the workplane. A closed hand would end the communication. Usability testing is needed to determine how best to convey system status which we leave for future work.

6 Future Works

We plan to enact a series of usability tests to rigorously test transferrable control in interaction design. Furthermore, we see the framework for human-object interaction as also applicable to rapid design iteration of physical interfaces.

We are also interested in the use of the YCrCb space as a stronger indicator of chromaticity and luminance extraction. Another area of exploration could be conversion to CMYK space to produce printable control interfaces. In our work, we encountered a potential means of surface classification by noting incongruencies from our brightness extraction models. We also plan to implement further control structures such as three dimensional space textitpoint-and-control gesture recognition.

7 Discussion & Conclusions

Through this proof-of-concept study, we found a potential extraction method to replace obstructive photosensor measurements and create closer bonds with human behaviors. Through the use of transferrable control, a link between the user and the items he or she uses forms a new form of interaction design. We were limited by the color rendering index available in our calibration studies, but nonetheless a rigid adaptive system was able to be constructed from CMOS camera sensor data. Although skin segmentation did not hold robust under lighting, unique colors such as those within the orange hues were identifiable at as little as 20 L*.

The system described in this paper used saturated hues as a primary method of detection. Although we were able to successfully identify color cues in complex backgrounds, we realize that there exist several factors that can effect color constancy. Sunlight, at dusk, produces a larger distribution of higher frequencies in its spectrum, leading to a linear change in all environments under this illuminant. More research needs to be conducting in a codebook method to follow small changes in color and adjust an image so that hues stay constant.

Lighting control is currently a habitual flip-of-a-switch. Even in the case of IR remotes that bring switches to the sphere of a user, it is not enough to elicit changes in the system. Gesture recognition elicits participation notably from all age groups. However, natural communication between a human and machine through gestures would result in a similar flip-of-switch behavior, although more conveniently located.

There is also the question as to whether a constant non-changing light stimulus can trigger changes in physiologic behavior. Several studies have investigated the use of lighting to adjust circadian rhythms for nighttime workers **??**. If stable lighting is achieved, the lack of fluctuation could produce unhealthy sleeping behavior.

We established a light-invariant vision system with a framework for future design and prototyping. A common problem of smart lighting is an overly complex

Table 4: Functionality mechanisms using Post-Its as optical cues.

Switch	Button-interface. Only active with a pointing gesture. Includes: On, Off, Presets: Reading, Writing, Computer Work, Sleeping, Standby, Sound On, Mute
Dimmer	Sound cue. Pass hand over post it and drag down to lower intensity or up to raise intensity of stimulus. Close hand to set.
Individual control	Sub-function. Can be coupled with other functions to create multiple individual controls. Associates a light or set of lights to an object.
Detector	Sub-function. Can be coupled with other functions to create multiple individual controls. Tracks object presence in the environment.

and hard-to-remember control interface. An human-object interface would allow for a subjective linking of action as well as habit. We noted the limitation that a visiting user would not be familiar with the interface construction of another user. We resolve a traditional switch may be the best answer for retaining flexibility. Overall, we found that sensor networks can not only pull data from the user, but also learn their behaviors through object and space interaction.

8 Acknowledgements

A special thanks to Phillips Color Kinetics for their donation of IW-2 Light Blast modules which made this research possible.

This work was supported by the MIT Summer Research Program under the Office of the Dean for Graduate Education and by the MIT Media Lab.

References

[1] Aldrich, M., Zhao, N., and Paradiso, J. Energy efficient control of polychromatic solid state lighting using a sensor network. vol. 7784, SPIE (2010), 778408.

[2] Bowers, A. R., Meek, C., and Stewart, N. Illumination and reading performance in age-related macular degeneration. *Clinical and Experimental Optometry 84*, 3 (2001), 139–147.

[3] Boyce, P. R. *Human Factors in Lighting*, 2 ed. CRC Press, 2003.

[4] Boyce, P. R., Eklund, H., and Simpson, S. N. Individual lighting control: Task performance, mood, and illuminance. *Journal of Illuminating Engineering Society* (2000), 131–142.

[5] Djajadiningrat, T., Overbeeke, K., and Wensveen, S. But how, donald, tell us how?: on the creation of meaning in interaction design through feedforward and inherent feedback. In *Proceedings of the 4th conference on Designing interactive systems: processes, practices, methods, and techniques*, DIS '02, ACM (New York, NY, USA, 2002), 285–291.

[6] Dugar, A. M., and Donn, M. R. Tangible intervention: Improving the effectiveness of lighting control systems. *Lighting Research and Technology* (2011).

[7] Escuyer, S., and Fontoynont, M. Lighting controls: a field study of office workers' reactions. *Lighting Research and Technology 33*, 2 (2001), 77–94.

[8] Gilchrist, A., Kossyfidis, C., Bonato, F., Agostini, T., Cataliotti, J., Li, X., Spehar, B., Annan, V., and Economou, E. An anchoring theory of lightness perception. *Psychological Review 106*, 4 (Oct 1999), 795–834.

[9] Goldstein, E. *Sensation and perception.* Wadsworth Cengage Learning, 2009.

[10] Guillemin, A., and Morel, N. An innovative lighting controller integrated in a self-adaptive building control system. *Energy and Buildings 33*, 5 (2001), 477 – 487.

[11] Inc., N. C. Solid-state lighting research and development portfolio:multi-year program plan fy09 fy15 tech. rep. *Lighting Research and Development Building Technologies Program* (2009).

[12] Koenderink, J. J. *Color Science*. The MIT Press, 2010.

[13] Lotto, R. B., and Purves, D. The effects of color on brightness. *Nature Neuroscience 2* (1999), 1010–1014.

[14] Moore, T., Carter, D., and Slater, A. Long-term patterns of use of occupant controlled office lighting. *Lighting Research and Technology 35*, 1 (2003), 43–57.

[15] Moore, T., Carter, D., and Slater, A. A study of opinion in offices with and without user controlled lighting. *Lighting Research and Technology 36*, 2 (2004), 131–144+.

[16] Mozer, M. C., Dodier, R. H., Anderson, M., Vidmar, L., Iii, R. F. C., and Miller, D. The neural network house: An overview, 1995.

[17] Newsham, G., and Arsenault, C. A camera as a sensor for lighting and shading control. *Lighting Research and Technology 41*, 2 (2009), 143–163.

[18] Pan, Z., Li, Y., Zhang, M., Sun, C., Guo, K., Tang, X., and Zhou, S. A real-time multi-cue hand tracking algorithm based on computer vision. In *Virtual Reality Conference (VR), 2010 IEEE* (march 2010), 219 –222.

[19] Williams, A., Kabisch, E., and Dourish, P. From interaction to participation: Configuring space through embodied interaction. In *UbiComp 2005: Ubiquitous Computing*. Springer Berlin / Heidelberg, 2005.

[20] Wu, Y., and Huang, T. S. Vision-based gesture recognition: A review. In *Proceedings of the International Gesture Workshop on Gesture-Based Communication in Human-Computer Interaction*, GW '99, Springer-Verlag (London, UK, 1999), 103–115.

[21] Wyszecki, G., and Stiles, W. *Color science: concepts and methods, quantitative data, and formulae*. Wiley classics library. John Wiley & Sons, 2000.

[22] Zhao, N. Smart solid-state lighting control. Master's thesis, Helmhotlz Institute for Biomedical Technology, August 2010.