Computational Cognitive Color Perception

Özer Ciftcioglu, *Senior Member, IEEE*
Department of Architecture
Delft University of Technology, The Netherlands
Maltepe University, Maltepe - Istanbul, Turkey
o.ciftcioglu@tudelft.nl ozerciftcioglu@maltepe.edu.tr

Michael S. Bittermann
Department of Architecture
Maltepe University
Maltepe - Istanbul, Turkey
michaelsbittermann@maltepe.edu.tr

***Abstract*—Comprehension of aesthetical color characteristics based on a computational model of visual perception and color cognition are presented. The computational comprehension is manifested by the machine's capability of instantly assigning appropriate colors to the objects perceived. They form a scene with aesthetically pleasing characteristics. The present approach to computational cognition is principally the same as contrived earlier [1]. This work distinguishes itself from the earlier work through the involvement of color differences. The color difference computations are carried out based on a standard human color observer model. The color difference information is combined with geometric perception information using the method of fuzzy neural tree based on likelihood. The study exemplifies the aptitude of the computational cognition for modeling cognition phenomenon. Cognitive color perception in computational form has generic relevance to applications involving human-like aesthetical appreciation, as is the case in building architecture, for instance and other design tasks.**

***Keywords*—visual perception; color difference; cognitive computing; genetic algorithm; fuzzy neural tree; auto-association**

I. Introduction

When a human being has experience with solving problems in a certain field, for instance his professional area, then his cognition is developed in this area. Cognition is to form a situated right strategy that requires some form of abstraction and optimization. Comprehension is the detailed abstraction and instantiation of cognition. The more thorough the comprehension of the task, the more one will be able to provide best solution with minimal reasoning effort involved. That is, the response arises spontaneously in one's mind without explicit remembrance of the concepts one had to familiarize oneself with, when one was not yet an experienced professional. This description, albeit one could hardly disagree with it based on common experience, does entail two significant problems. The first problem is to specify what is meant by the terms cognition and comprehension with minimal ambiguity. The second problem is to explain, how cognition and comprehension are accomplished, so that they yield a spontaneous, best reaction even in complex task contexts. Both issues are addressed in this study, presenting a computational cognition and comprehension approach, and employing it for a task that traditionally defies analysis, since, by definition, this task minimally involves abstractions in the form of linguistic concepts. The task is to comprehend the color aesthetics of a scene through visual perception. The comprehension entails the relations among geometry of environment, viewpoint, color of objects, perception of a scene, and aesthetical quality, so that human, when faced with an arbitrary color composition for the scene, is able to spontaneously propose modifications to the colors of the objects converting a non-aesthetical scene to an aesthetical one. This is one of the remarkable capabilities of human designers, and its computational reproduction is the aim of this work.

It is to emphasize that it is difficult to model cognition and comprehension, in particular in the domain of aesthetics. A judgment stemming from aesthetical comprehension appears to directly emanate from the perception act itself, i.e. without trace of intermediary reasoning and without reference to purpose [2-6]. Although the traditional consensus about this generic character of aesthetics in the literature may be considered as an insight into the topic, nevertheless it implies a severe challenge for modeling of the cognitive phenomenon. One possible direction one can think of to deal with the issue, is to develop non-parametric models of the relationship between a set of features of objects and an aesthetics label associated with the objects, e.g. see [7]. A flavor of associated persistent issues in such data-driven modeling approach can be obtained from the literature review in [8]. A second direction to deal with the aesthetical cognition modeling problem is to develop theoretical measures of aesthetics and to restrict the purpose of experiment to the validation of the measures [9-12]. *Complexity* has been considered as a property responsible for aesthetics, where generally high complexity of an object is deemed to yield high aesthetical appreciation of the object. An early work taking this view is due to Birkhoff [7], and several other works have taken the same view since [13, 14]. The dependence of aesthetics and complexity, however, appears questionable noting the abundant existence of aesthetical objects possessing high, as well as medium, as well as low complexity.

The present approach to computational cognition is principally the same as given earlier [1]. This work distinguishes itself from the earlier work through the involvement of color. The details of the research will be given in sections II and III. Color difference computations are carried out based on a standard human color observer model [15], and the color difference information is combined with geometric perception information using the method fuzzy neural tree based on likelihood [16]. Invoking the perception computations, the cognition is established by evolutionary computation and brought into refined form through auto-associative radial basis function network (RBF). The validity of the resulting cognitive color perception is verified by computer experiments.

The organization of the paper is as follows. In section II the visual perception, as well as abstraction processes involved in the color cognition are described. In section III a schematic overview of the computational cognitive color perception is given. In section IV the color cognition and comprehension computations are presented. Their validity is verified by computer experiments in section V. This is followed by conclusions.

II. Perception and Abstraction in Color Cognition

A. *Perceiving existence of objects and scene*

We consider a scene consisting of several objects numbered from 1 to n that has been perceived by an observer. The likelihood an object has been perceived by the observer depends on the fulfillment of two conditions. One condition is the object occupies the observer's visual scope. Omitting color considerations, in [17] perception is modeled as a probabilistic event, obtained via the integral of a probability density that is given per unit solid vision angle. The probability density models visual attention paid by the observer for objects within his visual scope. In the present work a likelihood approach to perception is taken [18]. We consider the likelihood the object has been perceived due to its geometry, and we denote it by $\mathbb{L}_G$. The solid angle defining observer's visual scope is denoted by Ω_S. We are interested in an object entirely contained within the scope, and denote the angle subtended by the object by Ω. Accordingly we define $\mathbb{L}_G$ by (1)

$$\mathbb{L}_G = \int_0^{\Omega} \frac{1}{\Omega_S} = \frac{\Omega}{\Omega_S},\ \Omega \le \Omega_S \tag{1}$$

The validity of (1) can be verified considering the extreme cases. When $\Omega = \Omega_S$ an object fills the entire visual scope, so that the likelihood it has been seen due to its geometry is maximal, namely $\mathbb{L}_G = 1$. Conversely, in case Ω is smaller than human's acuity threshold, then $\mathbb{L}_G = 0$. The second condition for perceiving an object is that its color should differ from the color of its background. A model of human color difference assessment is given in [15] in the form of a three-dimensional space known as the *C.I.E. 1976 $L^*a^*b^*$ color space* (CIELAB). The dimensions are denoted L^*, a^*, and b^* and a color is specified by a 3-tuple of coordinates as to the dimensions. The CIELAB space is based on the standard observer model described in [19], the standard illuminants described in [20], and the experimentally obtained color matching functions given in [21]. The space is designed to be perceptually uniform. This means the magnitude of a difference between two colors is given by the Euclidian distance denoted by ΔE^*_{ab}; explicitly $\Delta E^*_{ab} = \sqrt{(L^*_2 - L^*_1)^2 + (a^*_2 - a^*_1)^2 + (b^*_2 - b^*_1)^2}$. In this work we consider the likelihood an object has been perceived due to color difference. We denote this likelihood by $\mathbb{L}_C$ and define it

$$\mathbb{L}_C = \frac{\Delta E^*_{ab}}{\Delta E^*_{ab\_\max}} = \frac{\sqrt{(L^*_2 - L^*_1)^2 + (a^*_2 - a^*_1)^2 + (b^*_2 - b^*_1)^2}}{\sqrt{(L^*_{\max} - L^*_{\min})^2 + (a^*_{\max} - a^*_{\min})^2 + (b^*_{\max} - b^*_{\min})^2}} \tag{2}$$

where $\Delta E^*_{ab\_max}$ denotes the maximal color difference in the uniform color space. The validity of (2) can be seen considering the extremities. In case an object has no color difference with its background the likelihood the object has been perceived due to color difference vanishes completely. Conversely, in case the color difference is maximal, i.e. $\Delta E^*_{ab} = \Delta E^*_{ab\_max}$ then the likelihood is also maximal. In the ensuing computer experiments $\Delta E^*_{ab\_max} = 375.6$, which is determined by the gamut of the computer screen. It is to note that for the case an object occludes, and/or is occluded by, multiple objects with different colors, which is the general case, then ΔE^*_{ab} in (2) should be replaced by an average color difference denoted by $\overline{\Delta E^*_{ab}}$, so that (2) becomes

$$\mathbb{L}_C = \frac{\overline{\Delta E^*_{ab}}}{\Delta E^*_{ab\_\max}} = \frac{\frac{1}{k}\left(\sum_{i=2}^{i=k} p_{\Delta E\_ij}\sqrt{\left(L^*_i - L^*_1\right)^2 + \left(a^*_i - a^*_1\right)^2 + \left(b^*_i - b^*_1\right)^2}\right)}{\sqrt{\left(L^*_{\max} - L^*_{\min}\right)^2 + \left(a^*_{\max} - a^*_{\min}\right)^2 + \left(b^*_{\max} - b^*_{\min}\right)^2}},\quad \sum_{i=2}^{i=k} p_{\Delta E\_ij} = 1 \tag{3}$$

In (3) the object considered has the index number 1, and k is the number of objects being occluded by or occluding object 1. The weighting factor $p_{\Delta E\_ij}$ denotes the relative portion of the geometric perception of object 1 that is associated with the occlusion between object i and object 1. Computation of $p_{\Delta E\_ij}$ is described in [18].

In the likelihood based approach to perception put forward in this work, perceptions of objects and scene are modeled by fuzzy neural tree (FNT) method described in an earlier publication [16]. In a fuzzy neural tree the output of i-th terminal node is denoted y_i and it is introduced to a non-terminal node; the output of i-th non-terminal node is denoted O_i and it is introduced to another non-terminal node. The detailed schemes of node connections from terminal node i to internal node j and from internal node i to another internal node j are illustrated in the earlier publication, where the connection weights between the nodes are denoted by w_{ij} in both connection cases. In the neural network terminology w_{ij} is the synaptic strength between the neurons. Both, terminal node and non-terminal node outputs have interpretation as likelihood. Accordingly a weight w_{ij} is shown as the likelihood parameter θ_i and the output of an inner node j that is denoted by O_j in the earlier publication is shown as $\mathbb{L}_j$ in the following equations and figures. Let us consider a non-terminal node j that has two inputs, which are the outputs of two previous nodes denoted by O_1 and O_2. As the two inputs to a neuron are assumed to be independent of each other, the fuzzy memberships at the inputs can be thought to form a joint two-dimensional fuzzy membership. In this case $\mathbb{L}_j$ is computed by

$$\mathbb{L}_j = \mathbb{L}_1(\theta_1)\,\mathbb{L}_2(\theta_2) = e^{-\frac{\theta_1^2}{2\sigma_j^2}(O_1-1)^2}\ e^{-\frac{\theta_2^2}{2\sigma_j^2}(O_2-1)^2} \tag{4}$$

where σ_j is a constant, maximizing satisfaction of the consistency condition of possibility theory. For the two-input case $\sigma_j = 0.299$. The likelihood parameters θ_1 and θ_2 are selected commensurate to the amount of information they convey according to Shannon's information theorem, and they must sum up to unity for defuzzification in the rule-chaining process from node to node. Therefore the likelihood parameters in (4) are given by

$$\theta_1 = \frac{1-O_1}{(1-O_1)+(1-O_2)},\ \theta_2 = \frac{1-O_2}{(1-O_1)+(1-O_2)} \tag{5}$$

so that (3) becomes

$$\mathbb{L}_j = e^{-\frac{1}{2\sigma_j^2}\left(\frac{1-O_1}{(1-O_1)+(1-O_2)}\right)^2(O_1-1)^2}\ e^{-\frac{1}{2\sigma_j^2}\left(\frac{1-O_2}{(1-O_1)+(1-O_2)}\right)^2(O_2-1)^2} \tag{6}$$

The output neuron of a fuzzy neural tree is termed as *root* node, denoted by $\mathbb{O}$. The inner nodes providing the input to the root node are instances of $\mathbb{L}_j$ in (6). They are termed as *penultimate* nodes and denoted by $\mathbb{L}_k$. $\mathbb{O}$ is obtained via the weighted summation in (7), which represents the final defuzzification of the results from the neural tree computations

$$\mathbb{O}=\sum_{k=1}^{n} w_k \mathbb{L}_k \ , \quad \sum_{k=1}^{n} w_k = 1 \tag{7}$$

where n is the number of scene objects. In the absence of a priori preferences among scene objects, an important weight vector $(w'_1, w'_2, \ldots, w'_n)$ is the one that is aligned to the feature vector $(\mathbb{L}_1, \mathbb{L}_2, \ldots, \mathbb{L}_n)$. It maximizes the consistency of the defuzzification operation with the fuzzy logic principles, commensurately taking the fuzzy information from each input into account. The alignment means the influence of a root node's input on the node's output is proportional to the likelihood associated with the input, namely $w'_k = c\mathbb{L}_k$, $\forall k \in \{1, 2, \ldots, n\}$ where c is a scale factor and a constant. As to the perception model, the rationale is that the significance of an object's contribution to the perception of the scene is proportional to the perception of the object. Fulfilling the conditions of defuzzification, c is to be selected in such a way that the components of $\boldsymbol{w'}$ still sum up to unity as stipulated in (7). In this case (7) becomes

$$\mathbb{O}=\sum_{k=1}^{n} \mathbb{L}_k^{\,2} / \sum_{k=1}^{n} \mathbb{L}_k \tag{8}$$

Based on the above considerations, the FNT to compute perception based on likelihood is shown in figure 1.

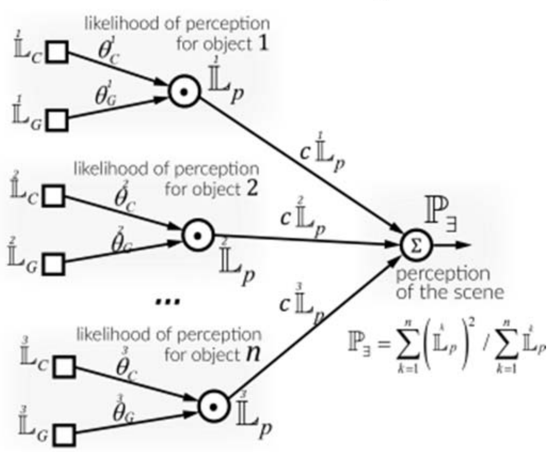


Fig. 1. Fuzzy neural tree for perception of a scene

Each inner node is associated to one scene object, and it has two inputs that are the likelihoods given by (1) and (3). Analog to (5) the likelihood parameters of the perception FNT are given by

$$\theta_C = \frac{1-\mathrm{L}c}{(1-\mathrm{L}c)+(1-\mathrm{L}_G)} \tag{9}$$

$$\theta_G = \frac{1-\mathrm{L}_G}{(1-\mathrm{L}_C)+(1-\mathrm{L}_G)} \tag{10}$$

The output of an inner node of the perception FNT represents the likelihood the existence of an associated scene object has been perceived. We denote the likelihood of perception by $\mathbb{L}_p$. Due to (6) $\mathbb{L}_p$ is obtained by

$$\begin{aligned} \mathbb{L}_p &= \mathbb{L}_1(\theta_C)\ \mathbb{L}_2(\theta_G) \\ &= e^{-\frac{1}{2\sigma^2}\left[\frac{1-\mathbb{L}_C}{(1-\mathbb{L}_C)+(1-\mathbb{L}_G)}\right]^2(\mathbb{L}_C-1)^2} \\ &\quad\ e^{-\frac{1}{2\sigma^2}\left[\frac{1-\mathbb{L}_G}{(1-\mathbb{L}_C)+(1-\mathbb{L}_G)}\right]^2(\mathbb{L}_G-1)^2} \end{aligned} \tag{11}$$

The root node output of the perception FNT models the perception of the scene's existence, which is the probability the scene is seen. It is denoted by $\mathbb{P}_\exists$. Referring to (8) $\mathbb{P}_\exists$ is obtained via aligned defuzzification of the objects' perception likelihoods given by

$$\mathbb{P}_\exists = \sum_{k=1}^{n}\left(\overset{k}{\mathbb{L}}_p\right)^2 / \sum_{k=1}^{n}\overset{k}{\mathbb{L}}_p \tag{12}$$

B. Perceived chromatic parsimony of objects and scene

Next to perceiving the existence of objects and scene, in an aesthetical judgment about the scene's color, the chromatic parsimony of the objects, and hence of the scene is subject to maximization [18]. *CIE Lab Chroma* C^*_{ab} is defined as the Euclidian distance of a color from the lightness axis L^* in the perceptually uniform CIELAB space $C^*_{ab} = \sqrt{a^{*2}+b^{*2}}$ [15]. The definition implies that the lightness components of the color and the nearest achromatic color have the same L^* value. For achromatic colors $C^*_{ab} = C^*_{ab\_min} = 0$ as these are located on the lightness axis, where a^*, b^* coordinates have the ues $a^*_{min} = b^*_{min} = 0$. One condition of chromatic parsimony of an object is defined in this work as the likelihood a color is achromatic, i.e. it has zero distance from the lightness axis in perceptually uniform color space. The parsimony is denoted by $\mathbb{L}_{C^*_{ab}}$ and given by

$$\mathbb{L}_{C^*_{ab}} = 1-\frac{C^*_{ab}-C^*_{ab\_\min}}{C^*_{ab\_\max}-C^*_{ab\_\min}} = 1-\frac{\sqrt{(a^{*2}-0^2)+(b^{*2}-0^2)}}{\sqrt{(a^{*\,2}_{\max}-0^2)+(b^{*\,2}_{\max}-0^2)}} = 1-\frac{\sqrt{a^{*2}+b^{*2}}}{\sqrt{a^{*\,2}_{\max}+b^{*\,2}_{\max}}} \tag{13}$$

In the ensuing computer experiments in this work $C^*_{ab\_max} = \sqrt{a^{*\,2}_{max} + a^{*\,2}_{max}} = 134$, which is determined by the gamut of the computer monitor. Clearly the parsimony becomes maximal, namely unity, for $a^* = b^* = 0$. The second condition to fulfill at the same time is an object's L^* component should be equal to scene's aesthetical reference lightness value denoted by L^*_r. The lightness of the illuminant forming the source for the color perception clearly is the maximal lightness value. In CIELAB space this value is $L^*_{max} = 100$. Pitch black color has the minimum lightness value. In CIELAB space this value is $L^*_{min} = 0$. The likelihood, the observer perceived an object's lightness to be the same as the scene's reference lightness, is denoted by $\mathbb{L}_{L^*}$ and given by

$$\mathbb{L}_{L^*} = 1-\frac{\left|L^*_r - L^*\right|}{L^*_{\max}-L^*_{\min}} = 1-\frac{\left|L^*_r - L^*\right|}{L^*_{\max}} \tag{14}$$

Based on the above considerations the fuzzy neural tree for color parsimony is shown in figure 2. Each inner node is associated to one scene object, and it has two inputs that are the terminal nodes given by (13) and (14). Referring to (5) the likelihood parameters of the color parsimony FNT are given by

$$\theta_{C^*_{ab}} = \frac{1-\mathbb{L}_{C^*_{ab}}}{\left(1-\mathbb{L}_{C^*_{ab}}\right)+\left(1-\mathbb{L}_{L^*}\right)} \tag{15}$$

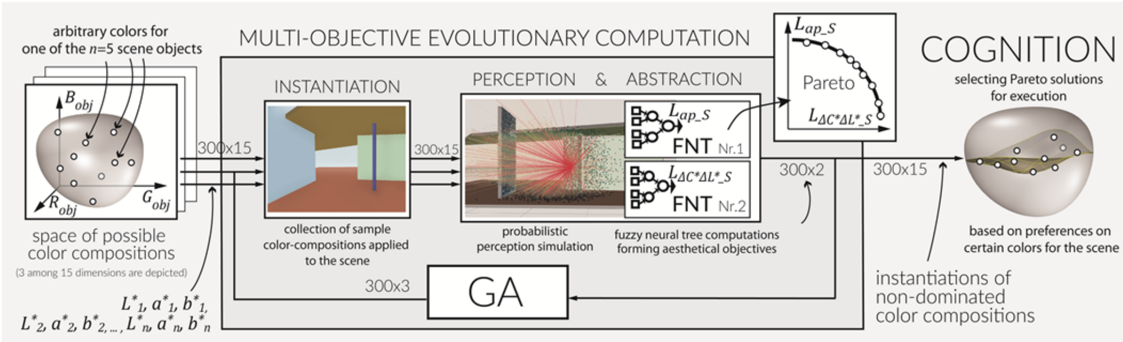


Fig.3. Scheme showing the computational cognition of establishing cognitive color perception

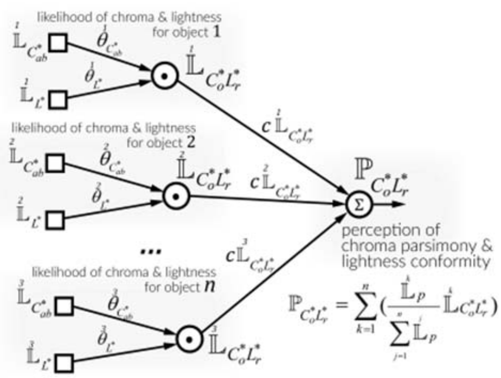


Fig. 2. Fuzzy neural tree for parsimony of CIELAB chroma of a scene

$$\theta_{L^*} = \frac{1-\mathbb{L}_{L^*}}{\left(1-\mathbb{L}_{C^*_{ab}}\right)+\left(1-\mathbb{L}_{L^*}\right)} \tag{16}$$

The output of an inner node of the FNT represents the likelihood the associated scene object is achromatic and while having the same lightness as the reference L^*_r. We term this quantity as the *likelihood of color parsimony and lightness conformity* and denote it by $\mathbb{L}_{C^*_oL^*_r}$. Due to (6) it is obtained by

$$\begin{aligned}\mathbb{L}_{C^*_oL^*_r} &= \mathbb{L}_1\left(\theta_{C^*_{ab}}\right)\mathbb{L}_2\left(\theta_{L^*}\right)\\ &= e^{-\frac{1}{2\sigma^2}\left[\frac{1-\mathbb{L}_{C^*_{ab}}}{(1-\mathbb{L}_{C^*_{ab}})+(1-\mathbb{L}_{L^*})}\right]^2\left(\mathbb{L}_{C^*_{ab}}-1\right)^2}\\ &\quad e^{-\frac{1}{2\sigma^2}\left[\frac{1-\mathbb{L}_{L^*}}{\left(1-\mathbb{L}_{C^*_{ab}}\right)+\left(1-\mathbb{L}_{L^*}\right)}\right]^2\left(\mathbb{L}_{L^*}-1\right)^2}\end{aligned} \tag{17}$$

The root node output of the color parsimony FNT models the perception of the chromatic properties of the scene. It is denoted by $\mathbb{P}_{C^*_oL^*_r\_S}$ and obtained via aligned defuzzification of the color parsimony and lightness conformity likelihoods given by

$$\mathbb{P}_{C^*_oL^*_r} = \sum_{k=1}^{n} w_k \overset{k}{\mathbb{L}}_{C^*_oL^*_r} = \sum_{k=1}^{n}\left(\frac{\overset{k}{\mathbb{L}}_p}{\sum_{j=1}^{n}\overset{j}{\mathbb{L}}_p}\overset{k}{\mathbb{L}}_{C^*_oL^*_r}\right) \tag{18}$$

In (18) the parameter n denotes the number of scene objects. One notes that In the equation the vector $\boldsymbol{w_p}$ is not aligned to the vector $\left(\mathbb{L}_{C^*_oL^*_{r}\_1}, \mathbb{L}_{C^*_oL^*_{r}\_2}, \ldots, \mathbb{L}_{C^*_oL^*_{r}\_n}\right)$ consisting of the objects' color parsimony likelihoods from (17); but it is aligned to the vector $(\mathbb{L}_{p\_1}, \mathbb{L}_{p\_2}, \ldots, \mathbb{L}_{p\_n})$ consisting of the objects' perception likelihoods from (11). This is done so that an objects' chromaticity influences the scene's chromaticity commensurate with its likelihood of perception. The rationale behind this modelling step is elucidated in [18].

III. Schematic Overview of Cognitive Color Perception Modeling

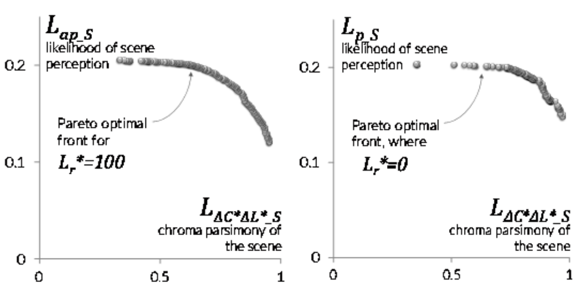
Establishment of cognition in computational form is initiated in this work by multiobjective evolutionary search. The resulting non-dominated parameters form the basis for color cognition, as they are considered to be representatives of the smooth continuum of relationships among them. The relationships are stored in an auto-associative radial basis function network (RBF), which provides the actual continuum. This is the basis of computational cognition considered in the present work, and its formation is schematically shown in figure 3. The inputs are color components assigned to the scene objects and specified in red, green, blue (RGB) color space, and they are converted to $L^*a^*b^*$ color coordinates as indicated in the figure. The evolutionary search is guided by consecutive instantiation, perception and abstraction processes that are repeated multiple times, providing the feed-back that drives the evolutionary process. The result of the search is a set of solutions that are conforming to the abstractions in Pareto optimal sense. In contrast to conventional multiobjective optimization occurring without reference to cognition, in computational cognition, the selection of a solution among the Pareto optimal ones for execution is not merely based on the objective function values of the solutions, but it also include consideration of the solutions' detailed features. In the present case of color cognition, these features are the detailed components of a color, such as amount of red, green and blue light that constitute the color. In the figure, instantiation refers to the assignment of possible colors to the objects of a scene.

IV. Aesthetical Color Cognition due to Multiobjective Search & Autoassociation

A. Multiobjective Evolutionary Search for Color Cognition

As described in [18] a certain color composition of a scene should be termed as aesthetical, when it fulfills the following condition for a certain $L^*_r \in \{L^*_r{:}0 \leq L^*_r \leq 100\}$. There exists no other color composition for this scene that at the same time yields a greater scene perception $\mathbb{L}_{p\_S}$ AND greater color parsimony $\mathbb{L}_{C_O^* L_r^* \_S}$. In the special case that $L^*_r \approx 100$, then the aesthetics is of *beautiful* kind; when $L^*_r \approx 0$ then it is of *sublime* kind. The condition described is the non-dominance criterion used in Pareto-based multi-objective optimization, when (12) and (18) are considered as two objectives subject to simultaneous maximization. The optimization should be carried out by a stochastic optimization algorithm due to the nonlinearity involved in the objectives. In this work we use an evolutionary algorithm to find Pareto front of aesthetical color compositions for a certain scene. The details of the scene will be addressed in the ensuing sections. For now we concentrate on the general role Pareto optimal solutions play in cognition formation.

Pareto front for the aesthetical perception problem is shown in figure 4a for the case of beautiful color compositions, and in figure 4b for the case of sublime compositions. As to multi-objective optimization, selection of a solution among the Pareto solutions is generally due to considerations in objective function domain exclusively. In contrast to this, in cognition, selection among non-dominated solutions is due to preferences for specific combinations of decision variable values. That is, cognitive considerations concern the decision variable domain. Selection of one of the Pareto solutions found by an appropriate search method due to preferences in the decision variable domain is defined as *computational cognition* in this study [22]. Cognition can lead to comprehension, when the relationship pattern inherent to the restricted number of non-dominant solutions is generalized in such a way, that it encompasses a theoretically infinite number of hitherto unknown non-dominated solutions. The presence of this continuum of best solutions manifests itself, when a point on the continuum is reached as result of feed-forward treatment of a presented stimulus, i.e. without invoking explicit perception and abstraction processes recursively, like they take place in optimization for instance.



(a) (b)

Fig. 4. Pareto front of beautiful color compositions, $L^*_r=100$ (a); of sublime compositions, $L^*_r=0$ (b)

B. Cognitive Perception by Autoassiciative Radial Basis Function Network

From the Pareto optimal solutions, computational color comprehension is obtained by establishing the autoassociative radial basis function (RBF) network seen in figure 5. In the following this network is referred to as cognitive perception network. Autoassociation implies that the input and output vectors are identical as seen in the figure. The vectors consist of chromaticity coordinates associated with each object of the scene. Due to the present computer implementation, the chromaticity is expressed by red (R), green (G), and blue (B) coordinates in the standard RBG (sRGB) color space [23]. The subscript at each coordinate indicates the scene object it belongs to. The network is trained by data pairs respectively formed by each Pareto solution vector and a clone of it. Clearly, due to the training a non-dominated vector at the model input is to produce an exact copy of itself at the model output retaining the same non-dominated position in objective function space. The essential point is that for a dominated input vector, as the network has exclusively been trained by non-dominated solutions, it produces an output that is also nearly non-dominated; and due to the naturally high number of hidden layer neurons of the network, the difference between input and output vector is desirably small at the same time. Such behavior emulates the manifestation of comprehension by human. Through the network training the explicit abstract conditions to be fulfilled have been converted into relations among the decision variables, so that their presence became

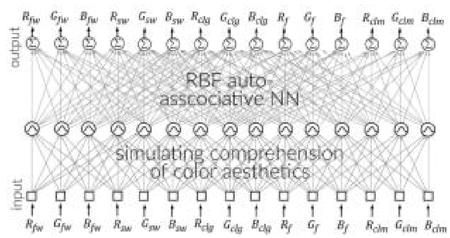


Fig. 5. Structure of the radial basis function network emulating cognitive color perception

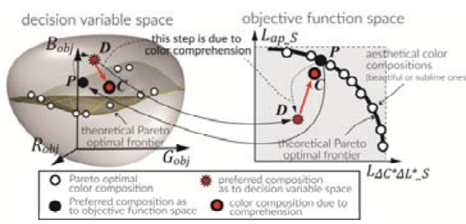
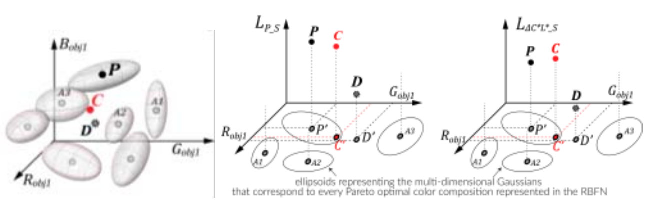


Fig. 6. Effect of stimulating the RBF network with a dominated solution at its input manifesting comprehension

implicit. The network is bound to produce solutions conforming to the desirable conditions. Based on the cognition network, the effect of comprehension during the pursuit of an aesthetical color composition is shown in figure 6. The Pareto frontier implies a hypersurface in the decision variable space [22], which is the color space in the cognitive perception case. It is to note that despite the emulative character of the computational cognition, the validity of the description in the figure extends to its biological counterpart. Namely from a desired solution D that is located at some distance from the Pareto frontier, comprehension produces solution C that is near to the Pareto front nearby P. This corroborates with the common observation that human beings with established cognition minimally display irrational behaviour. From this explanation one notes that manifestations of human cognition generally satisfy a second condition; namely the resulting solution should minimally differ from the corresponding stimulus in the decision variable domain, i.e. in color space. Indications of this characteristic of cognition can be found in human behavior. For instance an architect generally makes relatively small modifications to his design at every creative step during the design process, yet with every step he aims to maximally improve the design's performance with respect to the objectives. The cognitive perception network by RBF emulates this behavior due to the following inherent property of the computational method. We consider a suboptimal response D that is obtained by modifying a few x values of a non-dominated solution P. This is shown in figure 7a, in the three-dimensional space formed by a subset of the decision variables. One notes that we consider the difference between D and P to be exclusively with respect to the three decision variables that form the space in figure 7a, while for the remaining 12 decision variables R_{sw}, G_{sw} ..., B_{clm} solution D and P are considered to have identical values. Figure 7b shows the same solutions in the space formed by two exemplary decision variables and the first objective, given by (2). Figure 7c shows them in the space formed by the same two exemplary decision variables and the second objective, given by (6). Stimulating the cognition network by D then the solution C produced at the network output is bound to be similar to P in terms of their $\boldsymbol{x}$ data vectors as seen in figure 7b and 7c. This behavior is due to the multi-dimensionality of the dataset, where the autoassociative relation is established via combinations among all input variables as seen in figure 5. This is illustrated in figure 7a by the ellipsoids representing the multidimensional Gaussian basis functions of the cognition network. Point D is represented chiefly by the basis function belonging to point P, and two aesthetical solutions $A2$ and $A3$ that are the Pareto points with greatest affinity to D in the decision variable domain. As many decision variables values are identical to P, only slight movement away from D in the multi-dimensional space suffices to produce a solution with significantly different objective function values. With respect to the objective function domain, the response C will be close to the location of P on the Pareto front as shown in figure 6. This is due to the local representation nature of the radial basis functions, which ensures that the population members that are near to the original point P in the multi-dimensional response space are commensurately more effective in representing C [1].



(a) (b) (c)

Fig. 7. Cognitive color perception by radial basis functions, yielding solution C from D, where D has several multiple identical color components as P

V. Computer Experiments

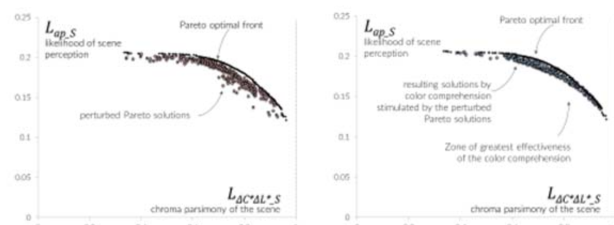
The above theoretical considerations are verified by means of computer experiments. An architectural scene is considered consisting of the following five objects: a wall oriented approximately perpendicular to the central perception direction and referred to as *frontal wall*; a wall oriented laterally to the perception direction and referred to as *side wall*, a floor, a ceiling and a column. For this scene a Pareto front of aesthetical color compositions is obtained using a non-dominated sorting based multiobjective genetic algorithm [24]. The decision variables are 3 chromaticity coordinates per object, which coincide with the variables of the network in figure 5. The problem is maximizing (12) and (18) where L_r*=*100* in (14). This means *beautiful* color combinations are sought. In the genetic search the population size is 300, and the algorithm parameters are set to standard values. Among the resulting 300 Pareto solutions 26 are used for the cognitive perception network training, where the sigma of the radial basis functions is set to $\sigma = 1.04$. The selection of the 26 training samples is based on having sufficient diversity in the data for effective network training. Using the trained cognitive perception network, two sets of experiments are carried out in the following subsections.

A. Experiment Set Nr. 1: Behavior for randomized stimuli

The first set concerns the general behavior of the cognitive perception network for partly or totally random stimuli shown in figures 8-11. The figures show the objective function space, where the Pareto optimal solutions are displayed by means of black colored, small dots.

a.) Injecting random color values to the existing Pareto solutions

Figure 8a shows modified versions of the 300 Pareto solutions that are referred to as *perturbed Pareto solutions* in this figure and subsequent figures. In figure 8, for each Pareto solution, the green color component of the front wall, the red component of the floor, and the green component of the

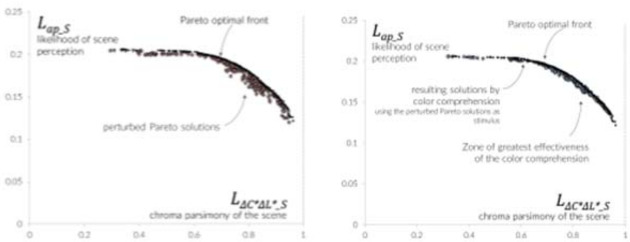


(a) (b)

Fig. 8. Pareto solutions where the color components G_{fw}, R_f and G_{clm} in figure 5 have been replaced by random values (a); resulting solutions by color comprehension (b)

column are replaced by a random value. The corresponding variables are denoted by G_{fw}, R_f and G_{clm} in figure 5. The random values are generated within the bounds given by the minimum and maximum values in the Pareto set for the respective variable. From the figure it is seen that the random injection causes most of the Pareto solutions to deviate from their original location on the Pareto front, and the deviation is most severe in the region of high chroma parsimony. The 300 perturbed solutions are used as input to the cognitive perception network, and the 300 solutions produced by the network are shown in figure 8b. From the figure it is seen that those solutions that deviated significantly from the Pareto front are brought close to the front, so that all solutions resulting from the cognitive perception network are nearly non-dominated.

Figure 9 shows a second experiement based on another set of perturbed solutions. This time for each solution, the green component of the front wall, red component of the floor, and blue component of the column are replaced by a random value. That is, compared to before, instead of the green component of the column the blue one is modified. The corresponding variables are denoted by G_{fw} , R_f and B_{clm} in figure 5. Again the random values are generated within the bounds specified by the minimum and maximum values in the Pareto set for the respective variable. Comparing figure 8a and 9a it is seen that the random injection causes less deviation from the Pareto front in the second experiment. This means that in this design problem, in order to reach an aesthetical scene, there is somewhat more tolerance about the blue component of the column compared to its green one. Using the perturbed stimulus, the output from the cognitive perception network is shown in figure 9b, demonstrating the method's effectiveness also for this case. One notes from figures 8b and9b that the region along the Pareto front, where the comprehension is most effective, remains the same as in the first experiment.

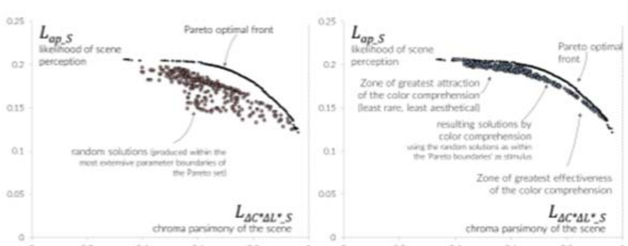


(a) (b)

Fig. 9. Pareto solutions where the color components G_{fw} , R_f and B_{clm} in figure 5 have been replaced by random values (a); resulting solutions by color comprehension (b)

b.) Cognition for random color stimuli generated within the Pareto set boundaries

In a third experiment, instead of merely three, all 15 components of the input vectors are randomized, abandoning the previous strong resemblance of the cognition stimulus to the Pareto optimal solutions. The only information used from the Pareto set are the 15 pairs of minimum and maximum values forming the set boundary in the decision variable domain, as the random vectors are generated within these boundaries. The random solutions are shown in figure 10a. As one should expect, among the random vectors almost none is close to the Pareto front. Using the random stimulus, the

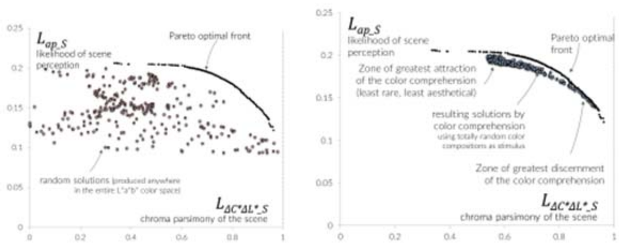


(a) (b)

Fig. 10. Color vectors, where all 15 variables in figure 5 have been replaced by random values within the region demarked by the Pareto set (a); resulting solutions by color comprehension (b)

output from the cognitive perception network is shown in figure 10b. From the figure one notes that all of the random stimuli are brought close to the Pareto front by the cognitive perception network, manifesting the robust nature of the computational aesthetical comprehension. This is an indication that the computational color cognition presented does corroborate with the common observation, that humans with established cognition rarely take suboptimal decisions irrespective of present contingencies. Considering the density of solutions produced by the color comprehension one notes that from most of the random stimuli the cognitive perception network produces solutions in the zone along the Pareto front, where chroma parsimony is rather low, while the scene perception is rather high. The cause of this behaviour lies in the nature of the aesthetical problem. One notes that it is more difficult to reach solutions having low chromaticity while minimally sacrificing perception intensity, compared to reaching solutions with high perception and some moderate chromaticity parsimony. This is because the conflict is more 'severe' for the former condition compared to the latter. The severity is in the following sense. When an object has a certain color with low chroma, the region of other colors in the color space that also have a low chroma while yielding a high color difference with the first color at the same time, is relatively small. This is a property of the shape of the perceptually uniform CIE Lab space, and it originates in the pronounced metamerism of human vision nearby the color white [25]. In contrast, when the stipulation of low chromaticity is relaxed, then there exist a large amount of color combinations producing equally high contrast with a second color. Therefore solution density is generally significantly lower at the front's extremity of high color parsimony, compared to the extremity of high scene perception. This character of the aesthetical problem should show up even more clearly, the less 'Pareto-like' a stimulus for cognition is. Before coming to that, it is noteworthy to mention that in contrast to the previous two experiments, although the solutions in figure 10b are near to the front, around the knee point of the front they do not touch the front. This indicates that cognitive perception network created solutions that, from decision vector viewpoint, may have little in common with the original Pareto solutions. This is remarkable, since this shows the cognitive perception network structure implies a vast flexibility in reaching the goals at hand, and that it did in deed grasped the goals. This indicates the explanatory potential of the modelling, providing a unique reproduction of the original, goal oriented character we observe in human creativity.

In a fourth experiment the stimuli for the cognitive perception network are totally random color vectors, i.e. the boundaries for the random generator are taken to be utmost extensive, namely they are the boundaries of the RGB color space. These entirely random stimuli are shown in figure 11a. The figure confirms the fact mentioned in the previous paragraph, that encountering a solution with high chroma parsimony is a rarer event compared encountering a solution with high scene perception. The result produced by the cognitive perception network is shown in figure 11b. As in the previous experiment all solutions are close to the Pareto front, while the majority of solutions occur nearby the region of high scene perception and moderate chroma parsimony. One should stress the remarkable character of this result. Although the stimulus to color comprehension contains no information at all about Pareto optimal solutions, the quality of the result in terms of objective function values produced by comprehension is almost as high for the stimuli that closely resemble Pareto solutions, seen in experiments one and two. Certainly, from practical viewpoint this behavior may not valuable, since in general cognition is particularly concerned with the exact decision variable combination. However, from a theoretical viewpoint it substantiates the earlier indication, that the cognitive perception modelling approach presented indeed does corroborate with the common manifestations of human cognition. The result gives an explanation, how it is possible that any stimulation of human brain, even background noise, is subject to conversion to a reasonable response and exclusively a matter of the presence of a sufficiently developed cognition. For instance, a trained pianist is able to perform a piece while thinking of unrelated matters or not thinking at all. This behavior is due to the enormous number of connections existing in the cognitive perception network, even considering a moderate problem having 15 input variables, and the extraordinary generalization capability of RBF network. Due to the many connections, there are an unimaginable number of possible ways the network is able to imbue the non-dominance property into a stimulus, and as all the samples used to establish the model are Pareto points, then any stimulus to the network will be represented by means of membership degrees from Pareto points. Therefore a response at the cognitive output is bound to conform to the Pareto front [22]. This property of the model might uniquely shed some light on the enigmatic coincidence of immediacy and approximate adequacy that characterizes a creative act, as it occurs for instance during the early phase of design or art creation.

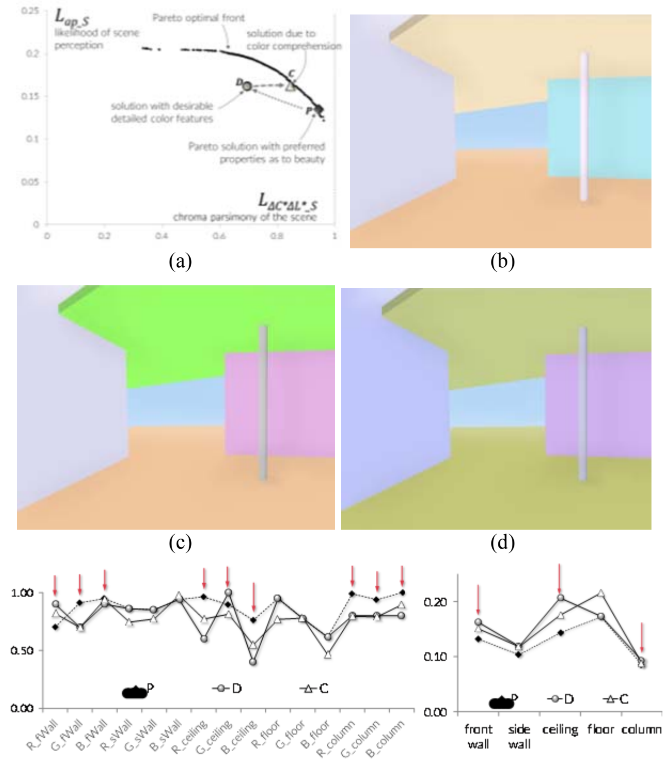


(a) (b)

Fig. 11. Color vectors, where all 15 variables in figure 5 have been replaced by random values within the full gamut of the color space (a); resulting solutions by color comprehension (b)

B. Experiment Set Nr. 2: Detailed Analyses of the Color Comprehension

The second set of experiments concerns analyses of the comprehension mechanism in detail, namely verifying its effect on the decision variables. More precisely, the theoretical considerations illustrated in figure 7 are verified, that for a relatively small modification of a Pareto solution there is an additional condition the cognitive perception network should fulfill, next to producing a solution that is nearby the Pareto front in objective function space. The additional condition is that the resulting solution should be minimally differing from the corresponding stimulus in the decision variable domain, i.e. in color space, corroborating with the commonly observed behavior of designers mentioned in the previous section. To investigate the fulfillment of this condition by the computational cognitive perception network, three experiments are carried out, and the results are shown in figure 12-14. It is to note that the colors in ensuing figures depicting the architectural scene subject to cognitive perception analyses are bound to appear somewhat incorrect in printed copies of the paper. This is due to the



(a) (b) (c) (d) (e) (f)

Fig. 12. Objective function space (a); selected Pareto solution *P* (b); desired solution *D* (c); solution by cognitive perception *C* (d); differences among *P*, *D*, and *C* as to parameter space (e); as to perception likelihoods (f)

restricted color space gamut of printing process compared to that of electronic representation on computer screen. One also notes that the virtual camera from which the pictures in the figures were rendered uses the same position and orientation vectors that were used during the establishment of the cognitive color perception for the experiments.

One of the Pareto optimal color compositions is selected and denoted by *P* in figure 12a. The corresponding scene is shown in figure 12b. Nine color properties of *P* are modified as marked by the red arrows in figure 12e, yielding solution *D* that is shown in figure 12c. The modification concerns the color of the frontal wall from cyan to purple, the ceiling from beige to an intensely saturated green, and the column from a purplish white to steel grey. Stimulating the cognitive perception network with solution *D* yields solution *C* shown in figure 12d. Referring to figure 12e one notes that solution *C* is similar to solution D with respect to the frontal wall color and the color of the column, while the color of the floor and ceiling are quite strongly altered, for *C* to be located near to *P* on the front as seen in figure 12a. This behaviour can be elucidated from the perception plot in figure 12f. In order to let *C* be near to *P* on the beauty frontier, the intense perception of the *D*'s ceiling is lowered and the floor perception is increased by cognition. As both objects have highest likelihood of perception in the scene, the mutual balance among their perceptions is most important to maintain the beauty of the scene, while the other objects' perceptions remain less affected.

A second Pareto optimal color compositions is selected and denoted by *P* in figure 13a. The corresponding scene is

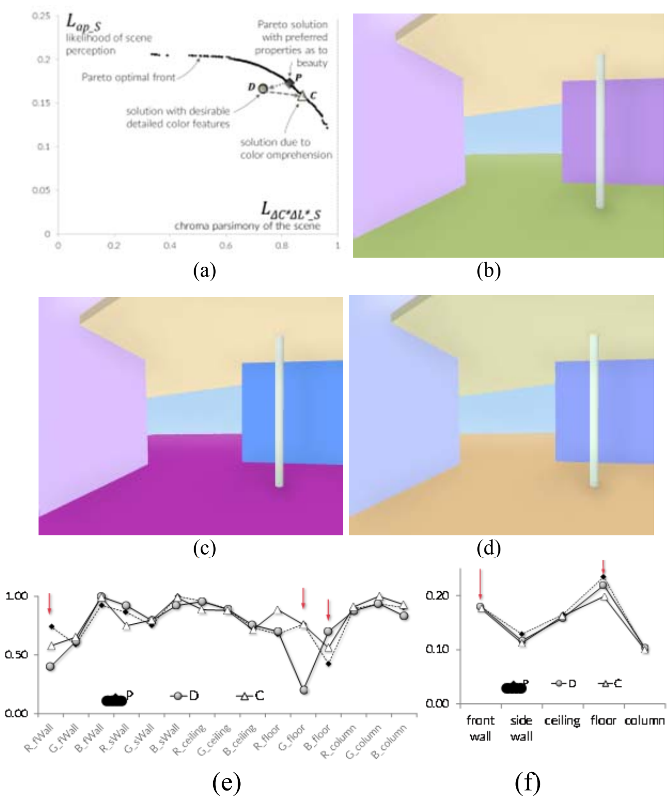


Fig. 13. Objective function space (a); selected Pareto solution *P* (b); desired solution *D* (c); solution by cognitive perception *C* (d); differences among *P*, *D,* and *C* as to parameter space (e); as to perception likelihoods (f)

shown in figure 13b. Three color properties of *P* are modified as marked by the red arrows in figure 13e, yielding solution *D* that is shown in figure 13c. The modification concerns the color of the frontal wall from purple to blue, and the floor from green to a saturated purple. Bringing *D* back to the Pareto the front as *C* shown in figure 13d, comprehension is able to largely accommodate the demand for blue frontal wall by taking some red out from ceiling and sidewall and adding some red to the floor. However, comprehension rejects the demanded diminishment of the green component of the floor. The floor should have sufficient green component, otherwise the floor would not have enough contrast with the two walls that are now both lacking in red, whereas originally the side wall was lacking in green, not in red. Compared to the comprehension event in figure 12, change in object perception is only minimally occurring in figure 13.

A third Pareto optimal color compositions is selected and denoted by *P* in figure 14a. The corresponding scene is shown in figure 14b. A single color properties of *P* is modified as marked by the red arrow in figure 14e, yielding solution *D* that is shown in figure 14c. The modification concerns the color of the frontal wall from green to turquoise. One notes that in this case *D* is very near to the Pareto front in contrast to the previous two experiments. Comprehension yields solution *C* in figure 14d located on the Pareto front, accommodating partly the requested color change for the frontal wall, by minimally modifying several color components of the other objects.

The three experiments show that computational comprehension corroborates to its biological counterpart in that it

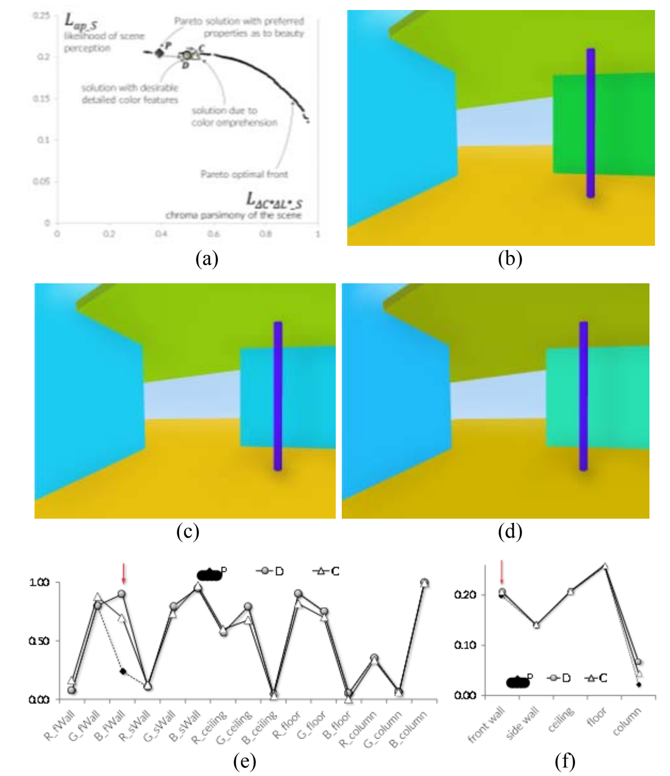


Fig. 14. Objective function space (a); selected Pareto solution *P* (b); desired solution *D* (c); solution by cognitive perception *C* (d); Differences among *P*, *D,* and *C* as to parameter space (e); as to perception likelihoods (f)

appropriately adjusts the dosage of change it exerts to a stimulus, namely commensurate with the already present affinity of the stimulus to the goals at hand.

CONCLUSIONS

Computational human's cognitive color perception is presented. The research reflects several commonly observed complex properties of cognition phenomenon. In particular the ability of an experienced human designer is computationally reproduced through the computational cognitive perception, obtaining aesthetical color compositions without explicit involvement of reasoning and search processes. The theoretical considerations are verified by means of computer experiments, showing the effectiveness of the work, and providing also a unique insight into the cognition phenomenon. In this way aesthetics of objects is shown to be possible subject to computation. Next to its practical value for diverse design and industrial applications, by means of the work the insight gained into the computational cognitive perception contributes to the underlying theoretical bases of such implementations.

REFERENCES

[1] O. Ciftcioglu and M. S. Bittermann, "Architectural design by cognitive computing," presented at the IEEE Congress on Evolutionary Computation - CEC 2015, Sendai, Japan, 2015.

[2] E. Burke, A philosophical enquiry into the origin of our ideas of the sublime and beautiful, 1757.

[3] I. Kant, Kritik der urteilskraft - analytik der ästhetischen urteilskraft. Darmstadt Wissenschaftliche Buchgesellschaft (published in 1983), 1790.

[4] T. W. Adorno, Aesthetic theory. London: Athlone Press, 1997.

[5] R. Scruton, The aesthetics of architecture: Princeton University Press, 1980.

[6] W. Tatarkiewicz, History of aesthetics vol. 1-3, vols. 1-2, 1970; vol. 3, 1974.

[7] M. Nishiyama, T. Okabe, I. Sato, and Y. Sato, "Aesthetic quality classification of photographs based on color harmony," presented at the IEEE Conference on Computer Vision and Pattern Recognition - CVPR, 2013.

[8] K. B. Schloss and S. E. Palmer, "Aesthetic response to color combinations: Preference, harmony, and similarity," Atten. Percept. Psychophys, vol. 73, pp. 551-571, 2010.

[9] R. Arnheim, Art and visual perception: The new version: University of California Press, 1974.

[10] D. E. Berlyne, Studies in the new experimental aesthetics. Steps towards an objective psychology of aesthetic appreciation. Washington D.C.: Hemisphere, 1974.

[11] E. H. Gombrich, A sense of order. London: Phaidon, 1995.

[12] R. L. Solso, Cognition and the visual arts. Cambridge, MA: MIT Press, 1997.

[13] J. Schmidhuber, "Driven by compression progress: A simple principle explains essential aspects of subjective beauty, novelty, surprise, interestingness, attention, curiosity, creativity, art, science, music, jokes. ," SICE Journal of Control, Measurement, and System Integration, vol. 1, pp. 21-32, 2009.

[14] H. Leder, B. Belke, A. Oeberst, and D. Augustin, "A model of aesthetic appreciation and aesthetic judgments," British Journal of Psychology, vol. 95, pp. 489-508, 2004.

[15] C. I. d. l'Eclairage, "Joint iso/cie standard: Colorimetry — part 4: Cie 1976 l*a*b* colour space ", ed. Vienna, Austria: CIE Central Bureau, 2007.

[16] O. Ciftcioglu and M. S. Bittermann, "A fuzzy neural tree based on likelihood," presented at the 2015 IEEE International Conference on Fuzzy Systems - FUZZ-IEEE 2015, Istanbul, Turkey, 2015.

[17] M. S. Bittermann, I. S. Sariyildiz, and Ö. Ciftcioglu, "Visual perception in design and robotics," Integrated Computer-Aided Engineering, vol. 14, pp. 73-91, 2007.

[18] M. S. Bittermann and O. Ciftcioglu, "Visual perception with color for architectural aesthetics," presented at the IEEE World Congress on Computational Intelligence - WCCI 2016, Vancouver, Canada, 2016 (under review for publication in this conference).

[19] C. I. d. l'Eclairage, "Iso 11664-1:2007(e)/cie s 014-1/e:2006: Joint iso/cie standard: Colorimetry — part 1: Cie standard colorimetric observers ", ed. Vienna: CIE Bureau, 2007.

[20] C. I. d. l''Eclairage, "Iso 11664-2:2007(e)/cie s 014-2/e:2006: Joint iso/cie standard: Colorimetry — part 2: Cie standard illuminants for colorimetry ", ed. Vienna: CIE Bureau, 2007.

[21] C. I. d. l'Eclairage, "Iso 11664-3:2012(e)/cie s 014-3/e:2011: Joint iso/cie standard: Colorimetry - part 3: Cie tristimulus values ", ed. Vienna, Austria: CIE Central Bureau, 2011.

[22] O. Ciftcioglu and M. S. Bittermann, "Generic cognitive computing for cognition," IEEE Congress on Evolutionary Computation - CEC 2015, Sendai, Japan, 2015.

[23] I. E. Commission, "Iec 61966-2-1:1999," in Multimedia systems and equipment - Colour measurement and management - Part 2-1: Colour management - Default RGB colour space - sRGB ed. Geneva, Switzerland: IEC, 1999, p. 51.

[24] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multi-objective genetic algorithm: Nsga-ii," IEEE Transactions on Evolutionary Computation, vol. 6, pp. 182-197, 2000.

[25] G. Wyszecki and W. S. Stiles, Color science, 2nd ed. New York: Wiley, 1982.