

International Journal on Artificial Intelligence Tools
© World Scientific Publishing Company

RENDERING AESTHETIC IMPRESSIONS OF TEXT IN COLOR SPACE

HUGO LIU & PATTIE MAES

*Media Laboratory, Massachusetts Institute of Technology,
20 Ames Street 320D, Cambridge, MA, 02139, USA
{hugo, pattie}@media.mit.edu*

Received (05 May 2005)

Accepted (03 October 2005)

What is an artwork and how could a machine become artist? This paper addresses the provocative question by theorizing a computational model of aesthetics and implementing the Aesthetoscope—a computer program that portrays aesthetic impressions of text and renders an abstract color grid artwork reminiscent of early twentieth century abstract expressionism. Following Dewey’s psychological interpretation of “aesthetic” and Jung’s ontology of fundamental psychological functions, we theorize that a viewer finds an artwork moving and satisfying because it seduces her into rich evocations of thoughts, sensations, intuitions, and feelings. The Aesthetoscope embodies this theory and aims to generate color grids paired with inspiration texts (a word, a poem, or song lyrics), which can be received as aesthetic and artistic by a viewer. The paper describes five Jungian aesthetic readers which are together capable of creative narrative understanding, and three color-logics that employ psycho-semantic principles to render the aesthetic readings in color space. Evaluations of the Aesthetoscope revealed that the program is best at portraying intuition and feeling, and that overall, the Aesthetoscope is capable of creating the aesthetic of art based on an inspiration text in a non-arbitrary way.

Keywords: aesthetics; the text; Bauhaus color psychology; theories of reading; semiology; generative art.

1. Introduction

In 1951, American minimalist painter Ellsworth Kelly exhibited a piece called *Sixty-Four Panels: Colors for a Large Wall* (Figure 1). Each of the colors in Kelly’s 8x8 grid were, according to his account, sampled from a different memory in his personal experience. The colors had a very personal meaning for Kelly, and as a gestalt, the grid could be said to create for Kelly, an *aesthetic resonance*—a rich impressionistic evocation of his life. Of course, this grid of colors can only create its most meaningful resonance for Kelly himself; for others, the piece is more abstract and playful like a game, inviting its viewer to read a life into its colors.

Alternatively, consider a piece entitled *Terre Provençale* by Dutch artist Herman de Vries (Figure 1)—each color square in its grid is a rubbing of the earth from different locations in Provence, France. Whereas Kelly’s piece could really only

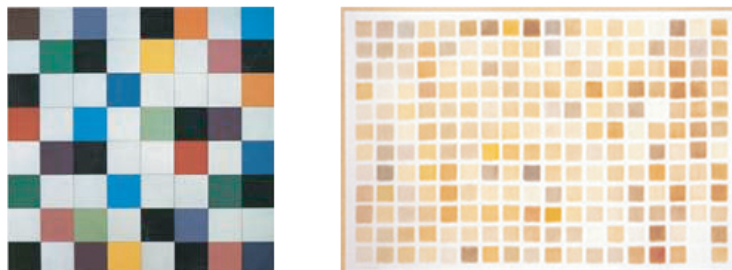
2 *Hugo Liu & Pattie Maes*

Fig. 1. Left: Ellsworth Kelly: *Sixty-Four Panels: Colors for a Large Wall* (1951); Right: Herman de Vries: *Terre Provençale* (1991). Figs. 1-12 in color at: <http://web.media.mit.edu/~hugo/>

evoke its original meaning for himself, *Terre Provençale* is more broadly evocative—its meaning is accessible for a whole community of people, namely, the residents of Provence, and to a lesser degree, all of mankind, who share a common experience with the various shades of yellow, brown, and red earth. Kelly and de Vries have both *enciphered* an aesthetic impression of something through color codes, but have set down differing rules for *decipherment*; Kelly’s cipher is a personal mystery, but de Vries’s cipher has not as much exclusivity. A viewer’s encounter with these pieces is *aesthetic* insofar as he is seduced by the code, he tries to decipher the code, and through this process his imagination is stirred—evoking a resonance of memories, sensations, and emotions within him.^a

The research described in this paper explores the question of how a machine might accomplish the same artistic feat as Kelly and de Vries—to likewise be able to use color grids to convey aesthetic impressions of some source material, which in the case of our research, is narrative text. In coming to answer this big question, we faced many other challenging questions. What are the semantic dimensions of text that contribute to the color impression of the text? How can the composition of a color impression account for the sensibilities of different people who engage, read, and value a text differently? How do colors evoke psychologically, and what sorts of things could they signal (emotions, visual memories, etc.)? How does the form into which colors are organized influence the aesthetic efficacy of the impression? And finally, how could we computationalize answers to these questions?

To test the computational models of aesthetic impression that we theorize in this paper, we built and evaluated a generative art robot called the Aesthetiscope. The Aesthetiscope takes as input a text such as a word, a poem, or some song lyrics, and renders out of that text a color grid that conveys an aesthetic impression which

^aAs we will clarify in Section 2, our understanding of *the aesthetic* is away from Kantian formalism, and more in the psycho-experiential spirit of Freud and Dewey. Notwithstanding, some readers will disagree with this paper’s semiotic standpoint toward aesthetic experience, which understands it as a language game.

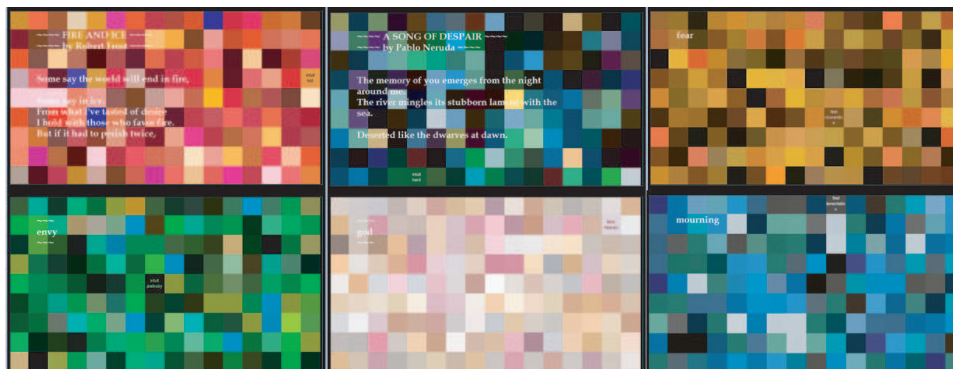


Fig. 2. Aesthetic renditions generated by the Aesthetoscope, set with bias toward Feeling and Intuition modes, for the following texts (clockwise from upper-left corner): a) the poem “Fire and Ice” by Robert Frost; b) the poem “A Song of Despair” by Pablo Neruda; c) the word ‘fear’; d) the word ‘mourning’; e) the word ‘god’; f) the word ‘envy’.

stirs sensations, memories, and emotions in the viewer. Figure 2 should give the reader an initial sense for the sorts of color grids that can be generated with the Aesthetoscope.

Our *artbot* works through the following mechanism. Based on theories of affected and aesthetic reading—that is, reading which fully engages the imagination, feeling, and sensation (Jakobson 1960; Rosenblatt 1978; Poulet 1980; Moorman & Ram 1994)—and based on Carl Jung’s (1921) theory that people interpret the world through four basal psychological modes, i.e., thinking, feeling, sensation, and intuition, our artbot “reads” a narrative text in five complementary ways: rationally, sentimentally, intuitively, culturally, and visually. The artbot uses psycho-semantic heuristics, such as those known to the Bauhaus school^b, to map those five textual interpretations into the world of colors; finally, the artbot blends the color palettes to fill a color grid. Some individuals are inclined toward romantic interpretations, while others lean to realistic interpretations; the Aesthetoscope can create artwork to account for both perspectives by selectively weighting the contributions of the five interpretive dimensions to the final color selection.

To generate the aesthetic interpretation of the input, five robotic readers skim the narrative text, brainstorming keyword evocations while reading. Each reader digests the text according to its aesthetic mandate, and each reader outputs a bag of keyword evocations to evidence its understanding. Take for instance, the text of Robert Frost’s poem, “Fire and Ice.”

Some say the world will end in fire,
Some say in ice.

^bWe computationalize the color psychology developed by Johannes Itten and Joseph Albers—two members of the Bauhaus who developed color curricula for the school’s Basic Course.

4 *Hugo Liu & Pattie Maes*

From what I've tasted of desire
 I hold with those who favor fire.
 But if it had to perish twice,
 I think I know enough of hate
 To say that for destruction ice
 Is also great
 And would suffice.

ThoughtReader imagines rational entailments about the text, producing rational reactions like **world**⇒**earth**, **ice**⇒**cold**, **fire**⇒**hot**. **CultureReader** imagines cultural evocations like **world**⇒**crazy**, **desire**⇒**fashion**, **hate**⇒**racism**. Currently, these cultural associations source from a corpus of popular culture magazines of the United States. **SightReader** extracts objects from the text for which there exists visual imagery in a large collection of 100,000 stock images—e.g. **fire**⇒**photos of fire**, **world**⇒**photos of world**. **IntuitionReader** makes psychologically immediate free associations like **fire**⇒**hot**, **fire**⇒**engine**, **fire**⇒**red**. **SentimentReader** makes emotional associations like **fire**⇒**arousing**, **desire**⇒**arousing**, **desire**⇒**pleasurable**. By allowing the text to evoke freely along these five interpretive dimensions, the artbot can be thought of as simulating the brainstorming process of a human artist—gathering together all the raw materials of inspiration from a text. We have however limited the current artbot to making commonsense associations—that is to say, these associations resonate with a whole culture of people, rather than being specific to the viewer's unique memories and experiences. In this sense, the aesthetic of the produced artwork is more in line with Herman de Vries's *Terre Provençale* than with Ellsworth Kelly's *Sixty-Four Panels*.

We feel that the subject of the present research bears significant implications for both the Aesthetic Theory and Artificial Intelligence communities. The literature of Aesthetic Theory is extensive and includes important semiotic analyses of art, music, and poetry, such as Roman Jakobson's (1960) structuralist critique. However, the literature's computable subset has largely been based in Kantian formalism, revolving around the mathematical or *exact* aesthetics of shapes (Birkhoff 1932; Stiny & Gips 1978a). The work described in this paper contributes a different computational perspective, exploring how the psychological dimensions of aesthetic experience might be modeled and implemented. Thusly, we do not claim that the present account of aesthetics is the true and definitive one. We hope only to enrich and reinforce the literature's understanding of aesthetic experience through the act of computation, because as Knuth famously argued, "The attempt to formalize things as algorithms leads to a much deeper understanding than if we simply try to understand things in the traditional way."^c For the Artificial Intelligence community, the prospect of programs that can reflexively stir up feelings, sensations, and emotions in their users expands the horizon for how A.I. programs might in the

^cDonald Knuth: 1973, Computer Science and Mathematics, *American Scientist*, 61(6): 709

future touch the lives of people. Even if this research fails to lay a generic foundation that could direct future research into computational aesthetics, we believe that the chronicle of our attempts documented here would still constitute an inspiring and salutary foray into one of the most dogged bastions of human intelligence—our art and emotion.

The rest of the paper is organized as follows. In Section 2, we present a computational theory of aesthetic impression, grounding our ideas within the literatures of traditional aesthetic theory and cognitive artificial intelligence. In Section 3, we overview the architecture of the Aesthetiscope implementation. Section 4 describes in detail the mechanics of our five-dimensional model of aesthetic textual interpretation. Section 5 discusses the psycho-semantics of rendering an evocative color grid in the Aesthetiscope implementation. Section 6 presents an evaluation for, and a further discussion of the Aesthetiscope. The paper concludes in Section 7.

2. A Computational Theory of Aesthetic Impression

In our experience, eighty-percent of the challenge in computing psychological aesthetics may be getting the representational framework right. This section presents a computational theory of psychological aesthetics which enframes the Aesthetiscope's implementation. In Section 2.1 we situate our meaning of the word 'aesthetic' within its historical meanings. In Section 2.2 we employ the metaphor of a transaction to illuminate the structure of *aesthetic experience*, and attempt to articulate principles governing the efficacy of aesthetic communication. Section 2.3 discusses the aesthetic potentials of a narrative text—what are the elements of meaning that can be read from a text which might contribute to an aesthetic impression of the text? Section 2.4 tackles the issue of how a 'user model'^d of the viewer could direct the synthesis of aesthetic impression customized to an individual. In Section 2.5, we discuss the role that colors and the grid format play in aesthetic efficacy. Section 2.6 summarizes these discussions into a concise thesis about the computation of aesthetic impression.

2.1. Which 'aesthetics'?

To many people, the word *aesthetics* evokes blurry meanings like “the beauty of things” or “the formal study of art,” but actually the idea has been approached with enormous precedent and rigor throughout intellectual history. The establishment of aesthetics as a formal discipline in the Western tradition is owed dually to Immanuel Kant (1790) and Alexander Baumgarten (1735). Kant in his “Critique of Judgement” posed perception of beauty as something that could be pure and universal and thus, formalizable. By Kant's neo-Platonist logic, a thing's true

^dThe term 'user model' is not entirely appropriate, as 'user' insinuates that the artwork is designed to provide a service—but we invoke it because it is standard nomenclature.

beauty is due only to its form, and not influenced by the thing's psychological significance in the viewer's life or memories. Kant's monolithic treatment of aesthetics has thrived to this day. In the twentieth century, some neo-Kantians have tried to compute the aesthetics of geometric objects (Birkhoff 1932) and paintings (Stiny & Gips 1978a) through mathematical and grammatical approaches; a popular field practicing numerical aesthetics today is the study of shape grammars.

Baumgarten is often credited with having coined the word *aesthetics* in his *Reflections on Poetry* (1735), in which he described aesthetics quite differently than Kant's more dominant interpretation. A philosophical descendent of Leibniz, Baumgarten was interested in the cognitive value of art—he posed aesthetics as the sensorial and imaginative experience of poetry, and he conceived of the ideas in poetry as “clear and confused” as opposed to rational ideas which were “clear and distinct.”^e Within Baumgarten's cognitive account of aesthetics, later theorists focused on two aspects of aesthetics—some implicated sociological conditions as the primary shaper of an individual's aesthetic cognition (Adorno 1970; Bourdieu 1984), while others felt that psychological conditions are more dominant (Freud 1919; Dewey 1934). Where the sociological and psychological camps agreed was to challenge Kant's notion that aesthetics could be universal.

The discussion in this paper is concerned with *psychological aesthetics*—aesthetics understood as an intimate and subjective phenomenon—although the influence of cultural conditions on cognition is partially accounted for. An artwork's *aesthetic*, then, is its capacity to affect a person in some manner. Two chief proponents of this perspective on aesthetics were Sigmund Freud and John Dewey, standing atop the experiential philosophies of Edmund Burke and David Hume. For them, aesthetic is just the opposite of what it was for Kant—it is not a matter of form, nor is it objective, but instead it is ‘related to the feelings’ of each subject's psyche (Freud 1919). For Freud, aesthetic was a much more intimate and narcissistic notion – something is aesthetic if it *moves* us, and we are moved most by things that mask and foil aspects of ourselves, our memories, and our desires—in other words, Freudian aesthetics is concerned with artwork-as-mirror. Dewey, too, is important in having shaped aesthetics as a subjective phenomenon. In *Art as Experience* (1934), Dewey put forth the thesis that art has the character ‘aesthetic’ because it can *seduce* viewers into *aesthetic experience*—a state of vulnerability, a state in which our censors are sublimated, we drop our callous social façade, and we become sensitized to myth and the true nature of things; in this state, we are highly perceptive, and receptive to sensations (seeing, listening, smelling, tasting), to feelings, and we let our imaginations run wild.

^eaesthetics: 2005, *Encyclopaedia Britannica*, retrieved November 23, 2005.

2.2. Aesthetic transaction

We carry forth the metaphor that Dewey initiated for conceiving ‘the aesthetic’ not as a fixed property of an artwork^f, but as a transaction between artwork and viewer. An artwork’s *a priori* aesthetic, then, is an aesthetics of potentiality. An artwork becomes aesthetic to a whole culture when it transacts successfully with the culture’s vanguards, who then endear it with a value cathexis so that it becomes significant to other cultural participants. In society, invocations of ‘aesthetic’ often refer to culturally-mediated aesthetics, but this should not diminish the significance of an artwork that can transact with but one individual, e.g. *Sixty-Four Panels* transacts with Kelley in a way different from how it moves other viewers.

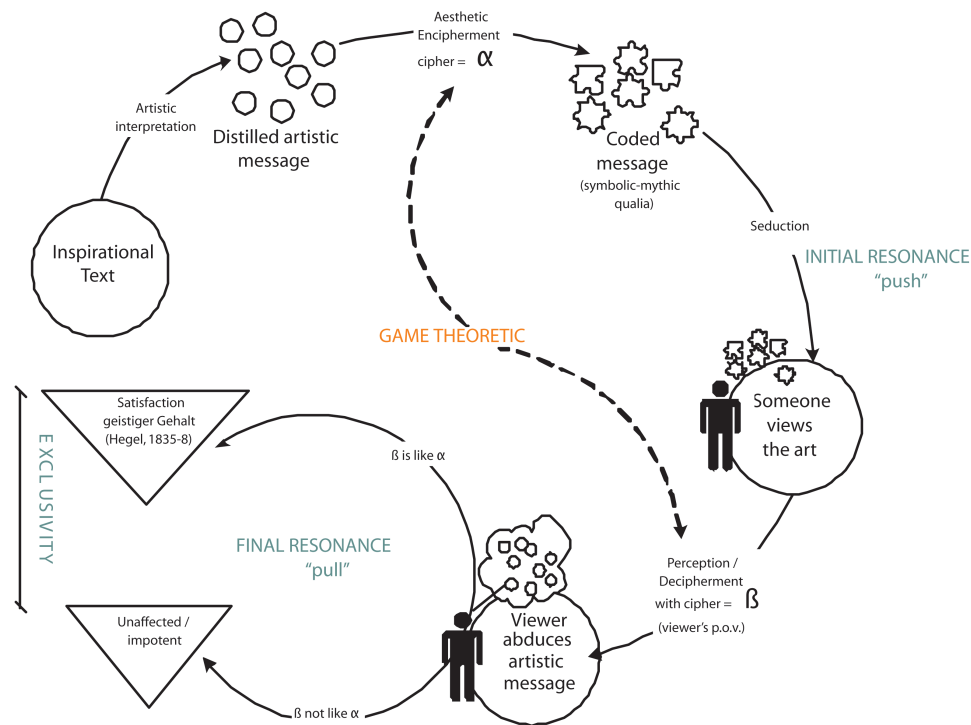


Fig. 3. An almost computational model of aesthetic transaction.

The transaction metaphor for aesthetic inspires a particular computational model of the aesthetic. Since, as Freud suggests, we are most readily seduced into aesthetic experience by seeing aspects of ourselves in an artwork, the transaction

^fHere, ‘artwork’ refers not just to art proper but more broadly to any object or idea.

can be conceived as occurring between the artwork's message and the viewer's concerns. Figure 3 depicts an almost computational model of aesthetic transaction, apropos the art genre of abstract expressionism in particular. We narrate the model as follows. An artist works with some inspirational 'text', which can be an idea, some imagery, or whatever. Exercising interpretation and judgement, the artist distills the text into a more concise artistic message. The artist chooses some artistic medium, e.g. a color grid. The expression of the message into the medium can be thought of as an aesthetic encipherment—the cipher being the artist's intent for and semantic appropriation of the medium. Next, the seduction phase involves motivating viewers to view and perceive the artwork. Viewers may be enticed toward an artwork by any number of factors—by its superficial beauty (e.g. shiny interesting things), by cultural endorsement (e.g. da Vinci's *Mona Lisa*), by intrigue or mythic qualities (e.g. sunset over a lake, the woods at dusk, a glimmering color grid), or by virtue of a thing being located in a place-for-perception (e.g. a museum). Marcel Duchamp, for example, successfully impelled viewers toward his *Fountain*, a urinal placed on its back, by virtue of the uncanniness of seeing such a crass and lowly object within a museum. When a viewer perceives an artwork, he deciphers it through an understanding of the medium's semantics and by adopting various interpretive standpoints. If the viewer deciphers the artwork as it was intended, he will receive the artistic message and feel satisfaction. If deciphering is unsuccessful or incorrect, it will leave the viewer unaffected. Decipherment should be viewed along a graded scale of efficacy, especially for abstract artwork. It should be stressed again that this mechanistic representation is a helpful stereotype of artistic production and consumption; symbolic computation after all deals primarily in stereotypes.

We enrich this process model with two principles governing the efficacy of this transaction: final resonance, and exclusivity.

2.2.1. *Final resonance principle*

A critic might point out the following flaw. If a viewer is concerned with himself and finds an artwork aesthetic insofar as he sees himself in the art, then a mirror should be art. Clearly, there is some other criterion in the secret recipe of aesthetic. The missing piece, we suggest, is *intimacy*. Insofar as experiencing art parallels reading a book, we should harken to theorist George Poulet's idealization of intimacy—"the I who 'thinks in me' when I read a book is the I of the one who reads the book" (Poulet 1980, 45). In contrast, semantic overttness and explication are not conducive to the aesthetic because they violate the intimacy between the artwork and viewer. For a viewer to feel affected and justified, she must perceive the experience to be psychologically intimate and unique and she must believe that she has discovered herself through the art; or harkening to Hegel (1838), art is aesthetic insofar as it provokes genuinely new thoughts and has intellectual import. The artwork should insinuate and facilitate meaning discovery, but it is the viewer who should be asked to take that final step to uncover significance in artwork. Her excavation of meaning

constitutes a ‘possession ritual’ (McCracken, 1988).

We refer to this idea as the *principle of final resonance*: an artwork may resonate with a viewer in many obvious ways, but what makes the artwork aesthetic is when the final resonance is the viewer’s move—a discovery of something extraordinary and personally meaningful in the artwork. A viewer would be most satisfied if she feels that her meaning is also the meaning intended by the artist (also *cf.* Section 2.2.2). In political art, the final resonance often comes as the discovery of the punch line of a joke. In early twentieth century abstract expressionism, the final resonance is the viewer’s discovery of whatever her psyche mandates, as that is the genre’s hope. Critiquing Duchamp’s Dadaist oeuvre, commentator Octavio Paz seemed to allude to final resonance: “the picture depends on the spectator because only he can set in motion the apparatus of signs that comprises the whole work” (Paz 1978, 80).

Within the realm of the Aesthetoscope’s color grids, the initial resonance is the viewer finding the color grid alluring but ineffable, while final resonance is the viewer’s eventual success in uncovering some meaningful mapping from the colors to the purported subject matter being depicted. The colors obey some semantic code, for they have in them the capacity to signify many things, e.g., a color could be the actual color of a real object, or it could depict a mood. Once the code and artistic intention for the colors have been broken, the viewer can feel the satisfaction of winning a game; in fact, it has been said that the essence of art ‘is a language game’ (Best (1985) commenting on Ludwig Wittgenstein’s aesthetics).

2.2.2. *Exclusivity principle*

Because we understand *aesthetic* to mean the resonance relationship between an artwork and a viewer, a color grid which evokes memories, sensations, and emotions for one viewer may completely fail to evoke in another viewer. *Sixty-Four Panels* can communicate an aesthetic impression of life to Kelly himself, but the piece will only tease other viewers and seem like a puzzle to them with pieces missing, because they cannot directly access Kelly’s memories and experiences. At best, viewers can only *aestheticize*, or ‘read into’, the abstract piece by invoking their psychic intuitions. *Terre Provençale* differed from *Sixty-Four Panels* in that its colors, sourced from shades of dirt from the earth in Provence, are amenable to more commonsensical interpretation—residents of Provence and all of humankind share rich memories of earth’s diverse shades. We suggest that as an evocation becomes increasingly commonplace (in opposition to unique and personally significant), the message’s intimacy declines and the power of the aesthetic is diminished (Liu 2004). Rare and challenging experiences are more endearing as rare things are often perceived to be more valuable. When an artwork impels a viewer to endear its message, it has won. We describe this idea as the *exclusivity principle*, and we state it as a play

on the famous aphorism of P.T. Barnum[§]—an artwork can be aesthetic for some of the people fully, it can be aesthetic for all of the people partially, but it cannot be aesthetic for all of the people fully.

To summarize, importing the principles of final resonance and exclusivity into the transaction framework, we can say that aesthetic efficacy of this transaction is strongest 1) when the viewer finds that the message of the art is one that he, given his experiences and values and perspective, is more competent to receive than an arbitrary person is (exclusivity); and 2) when the meaning of art is not worn on the surface but is something that must be excavated from the artwork by the viewer's own wits (final resonance).

2.3. *Aesthetic potential of narratives*

Much of the A.I. narrative understanding literature subscribes to the dogma that there exists an object method of interpreting text, and that resultant interpretations and inferences can always be reconciled into a single consistent world model. One research program notably departing from this dogma is concerned with *creative reading* (Moorman & Ram 1994). According to the cognitively motivated theory of creative reading, textual understanding involves imagination, the suspension of disbelief, and the projection of inexact memories onto read situations; in contrast, dogma would prescribe that textual understanding be algorithmized as the rote invocation of inference rules. Moorman & Ram's cognitivism goes against the grain of the classical A.I. narrative understanding literature and emboldens us in pursuing the computation of an *aesthetic reading* of text.

Aesthetic reading means reading not purely for information. It is an affected and sensuous reading, whereby the text's primary effect is to evoke aesthetic rumblings within the reader. Reading theorist Louise Rosenblatt offers this interpretation—“In aesthetic reading, the reader's attention is centered directly on what he is living through during his relationship with that particular text” (Rosenblatt 1978, 25) where this notion of “living through” can be a complex amalgamate of perceptions and sensations. To view reading within our aforementioned transaction framework, Rosenblatt distinguishes between two types of transactions involving a reader and a text – efferent and aesthetic. A reader may have an *efferent* transaction with a text, meaning that the reader is reading in order to carry something away (usually information) from a text. Just as a person requires a pail to carry away some water from a river, efferent transactions imply that a reader brings a mindset to the task of reading, and uses the mindset to scoop away something from the text. Objective reading, or reading to retain information, is well-described by efferent transaction. In contrast, aesthetic transaction is one in which a reader interacts with a text not through the narrow peephole of a mindset, but feels the full force of

[§]P.T. Barnum, the circus entrepreneur, is credited with the saying “You can fool some of the people all of the time, all of the people some of the time, but you cannot fool all of the people all of the time.”

the narrative's potential to affect; the reader allows herself to become affected and to receive sensations, moods, imaginations from the text.

Before formulating an artistic message, raw materials out of which the aesthetic message is crafted—such as imagery, moods, and symbolism—must be 'read out' of the text. These raw materials constitute the aesthetic potential of the text. While a robotic artist such as the Aesthetoscope could theoretically excavate raw materials in accordance with a viewer's personal memories (embodied in weblogs and emails), our present research pursues only a commonsensical excavation of raw materials—the Aesthetoscope excavates sentiments, sensations, and imaginations out of an inspirational text according to typical interpretation. This is not to say, however, that there is no opportunity to customize the aesthetic impression for individuals—in the next subsection, we discuss aesthetic customization based on a viewer's psychological and perceptual tendencies.

Since there are a great diversity of ways in which a reader may interpret a text under the aesthetic mode, a computational model must be sophisticated enough to account for *all* of them. Thus we develop a model of aesthetic reading that is inspired by Carl Jung's^h theory of psychological functions (1921). The theory accounts for all the different ways that people engage and interpret the world by giving an ontology of four fundamental modes of interpretation: Thinking, Feeling, Sensation and Intuition. To these four modes, we added a not-so-fundamental fifth, Culturalizing, which incorporates Roland Barthes' thesis (1964) that people also interpret the world through the optics of our culture's values. Also, for practical considerations, our work means the Sensation mode to refer solely to the remembrance of visual images, since the other senses are not within the research's scope.

Whereas objective reading relies primarily on the Thinking mode; aesthetic reading invites a reader to employ many, or all, of the five dimensions of interpretation to engage with the text, each mode producing its own set of evocations. The full aesthetic potential of a narrative text then, can be computed by reading text under a multitude of interpretive lenses—thinking about a text, feeling it, sensing it, intuiting it, and culturalizing it. This multi-perspectival model of interpretation differs significantly from the monolithic rational dogma of traditional A.I. narrative understanding in that it produces not so much a coherent understanding as just a creative brainstorm around a text. It is the same creative brainstorm process that an artist might engage in to expose the potentialities of a narrative text if he were to use the text to inspire the creation of a color grid artwork.

^hJung's psychoanalytic theories have largely fallen out of favour with the psychoanalytic establishment, shunned for their mystical and religious aspects. His theory of archetypes and collective unconscious remain highly influential in popular culture, carried forth by intellectuals like Joseph Campbell, and by widely used personality inventories.

2.4. *User model of the viewer*

Once the five dimensions of interpretation have been established, the question turns to how best to combine, and in what proportions to combine, the output of these interpretations into a coherent aesthetic impression of the text. Ideally, the proportions of the five modes should be customized to suit each viewer. Some viewers lean toward perceiving the emotional aspects of a narrative text, while other viewers are more sensitive to the visual imagery suggested by a text. Contemporary personality inventories offer a model of the viewer, which could be used to determine the relative contributions of the Aesthetiscope's five modes to the overall impression. The mapping in from personality schemes like Keirsey-Bates Temperament Sorter and the Myers-Briggs Type Indicator (MBTI) (Briggs & Myers 1976) are facilitated by the fact that these popular schemes are based on Jung's four psychological functions.

Currently, our Aesthetiscope implementation only allows for the manual adjustment of the proportional contributions of each Jungian mode to the final palette of colors. However we could imagine employing schemes such as MBTI to determine the proportions of contribution. MBTI is today's most widely used personality inventory and represents a person's personality along four scales: (I)ntrovert-(E)xtrovert, i(N)tuition-(S)ensing, (F)eeling-(T)hinking, and (P)erceiving-(J)udging. N, S, F, and T map directly into Aesthetiscope's modes. I and E were proposed as psychological attitudes, but we make no use of this. Perceivers rely more on intuiting and sensing, while Judgers rely more on feeling and thinking. Thus, P and J can be interpreted as meta-parameters over Jung's original four modes—P boosts the contributions of N and S, while J boosts the contributions of F and T. The contribution of Aesthetiscope's Culturalizing mode would need to be assessed separately as the degree to which an individual is initiated into the symbolism, imagery, and perspective of popular culture.

2.5. *Colors and form as vehicles of aesthetic impression*

Five-dimensional aesthetic reading of a narrative text yields five interpretations, which manifest as five sets of textual reactions. To realize this stew of text as an aesthetic, the stew should be wrapped in some aesthetic code so that viewers can themselves claim the experience of uncovering the artwork's meaning. In this work, we consider colors as the codifying realm for conveying aesthetic impression.

Colors are a superb medium of portraiture for the aesthetic character of a text, since color space is a complete micro-consciousness of pathos, just like taste and smell. Colors touch every part of our lives. They not only impart meaning to things, but also absorb meaning from things. Thus, they induce meaning not only through psychophysiology, but also by their symbolic and emotional power. Johannes Itten (1961) oversaw the teaching of color theory in the Basic Course at the Bauhaus School. His *Art of Color* captured many insights regarding common subjective experiences of colors, and in particular, how colors evoke emotion and portray mood. That colors are multi-purposable to represent visual sense, evoke symbolism, and

portray mood make them semantically ambiguous and thus an ideal artistic medium. Aesthetic, conceptualized as a language game, thrives in spaces of connotation; the ambiguity of colors affords viewers the opportunity to read many sorts of meaning into them.

Mapping the outputs of each Jungian mode into color space is also a most practical way of unifying the outputs of various interpretations into a gestalt. For example, consider the problem of unifying the visual and affective perceptions of the word “sunset.” In color space, this unification is trivial: remembered visual swatches of past seen sunsets can be epitomized as a color palette and this palette can simply be blended with the palette produced by sentimental entailments of the word “sunset” such as “warmth, fuzzy, beautiful, serenity and relaxation.” Our goal of conveying the text’s singular, complex aesthetic character to the perceiver is facilitated by the expectation that the viewer will blend these colors together in the mind’s eye, and attend to their undeconstructed gestalt rather than to each square individually. In this manner, the aesthetic character is not a simple sum of individual color squares, but rather, it becomes that Spirit which lives in-between the color squares.

If we mean color to be the sole vehicle of aesthetic impression, then we must carefully control the form that the colors take. The form of a grid of squares may be a particularly appropriate way to present the colors because a grid is a homogenous form which does not pretend to be carrying information in and of itself. Grids also have a great heritage in twentieth-century art, appearing in the works of artists like Sophie Taeuber, Jean Arp, Piet Mondriaan, Paul Klee, and Ellsworth Kelly. Grids add to the seductiveness of the artwork because as Rosalind Krauss described, “The grid’s mythic power is that it makes us able to think we are dealing with materialism while at the same time it provides us with a release into belief” (Krauss 1979, 12).

According to the Semiotician Roman Jakobson’s (1960) theory of communication, all communications serves one of six functions, describing—the sender, the receiver, the message, the message’s context, the communication channel, and the communication code. For Jakobson, communication was most poetic when it explicated the message alone. The invocation of color grids here can be said to have such an isolated communicative focus.

2.6. Summary

“Aesthetic” means the capacity of an artwork to sublimate the rigidity of a viewer, thrusting him into rumblings of imagination, sensation, feeling, and thoughts. Aesthetic is not a static property of artwork, but rather, an ephemeral transaction between artwork and viewer. The efficacy of a transaction is determined by how well a viewer can decipher and receive the artwork’s intended message—exclusivity and final resonance were introduced as two principles that model aesthetic efficacy. The text that inspired the artwork is modelled as the sum of all the artistic ways in which it can be interpreted. Based on Jung’s theory of psychological functions, we

presented five dimensions of artistic interpretation: thinking, culturalizing, seeing, intuiting, and feeling. We proposed that computational readers implementing each of these interpretive modes be applied to the narrative text to produce a range of interpretations. Applying a model of the viewer’s psychology, these five interpretations can be blended together to create an aesthetic impression.

3. Aesthetiscope’s implementation

In this section, we first describe the Aesthetiscope’s presentation and capabilities (3.1). Then, we present the architecture of its implementation (3.2).

3.1. *Presentation and capabilities*



Fig. 4. The Aesthetiscope installed in a “living room of the future” at MIT, generates color grid artwork to provide an “aesthetic pairing” for a book of poetry or a song playing over the room’s stereo.

The Aesthetiscope is currently installed in a “living room of the future” at the MIT Media Laboratory, and is projected onto one of the room’s walls (Figure 4). The grid of color squares is 16 wide by 9 tall (golden ratio aspect), flanked by black striping on top and bottom. There is a “glimmer” effect added to the colors in the grid, causing their Munsellian Values to wax and wane according to various periodicities. Finally, the glimmering of the color grid refreshes at not more than 24 frames per second, to affect a cinematic quality about the piece.

We intend for the Aesthetiscope not simply to stand alone as a showpiece but also to play a supporting role for other activities in the room. By visualizing the aesthetic character of a poem being read (this activity can be detected by our context-aware room), or of the lyrics to a song being played over the room’s stereo

system, we can imagine how the *pairing* of the Aesthetiscope's color grid with the poem or song might enhance the bandwidth of an aesthetic encounter, just as the tasteful pairing of food and wine enhances the experience of both.

Other capabilities of the interface are as follows. The narrative text that is at the heart of the dynamic artwork's message can be displayed as an overlay to the color grid, or it may be hidden. Artistic explanation is a feature that if turned on, flashes textual clues into the squares of the color grid which reveal the rationale for the colors. For example, for a rendition of a "sunset", with the aesthetic impression biased toward Feeling and Intuition, the generated artwork consists of warm yellows, oranges, and reds. The artistic explanation mode blits phrases like **feel warmth**, **intuit beauty**, **feel hug**, and **feel romantic** into the squares of the color grid. Currently the Aesthetiscope does not automatically customize its artwork to the MBTI of a viewer, but instead offers a menu with five sliders for Think, Culturalize, See, Intuit, and Feel, each from 0% to 100%, allowing a user to manually set the interpretive biases. Finally, to background the Aesthetiscope into the aesthetic integration of the room, the piece can be set to automatically visualize whatever book of poetry is laid on a radio frequency sensing coffee table, and whatever song is played in the room's jukebox.

3.2. Implementation overview

The Aesthetiscope is implemented in 11,000 lines of Python code, and a process model of its implementation architecture is depicted in Figure 5. To achieve implementations of the five modes of reading and translation of text into color, a great many different tools and representations were opportunistically employed. Marvin Minsky has labeled this opportunistic approach as 'scruffy AI', as opposed to 'neat AI'. We do not suggest that this juggernaut of an implementation is authoritative; it does, however, indicate the great range of semantic resources needed to model an underlying complex phenomenon such as aesthetics.

The implementation architecture can be viewed as taking five stages of processing, as shown in the rightmost-column in Figure 5. The first two stages, Text Parsing and Aesthetic Reading, are concerned with digesting the input narrative text, passing those digested pieces through the different interpretive lenses of five Readers, and collecting together the understandings of the input produced by each Reader. In the Text Parsing phase, the input narrative text is first digested with the MontyLingua surface semantic parser (Liu 2002). We chose a surface semantic parse, also known as a shallow parse, because the parse mechanism is more robust on genre-generic raw English text than many deep semantic parsers, and because it produces output in a representation required by the five Readers. MontyLingua performs the following textual digestion tasks: semantic tokenization, part-of-speech tagging, rule-based chunking, morphological lemmatization, and phrase attachment/linking. It outputs both a structured parse and a back-off parse. The structured parse is a linear sequence of syntactic frames, one for each independent clause, and taking the form,

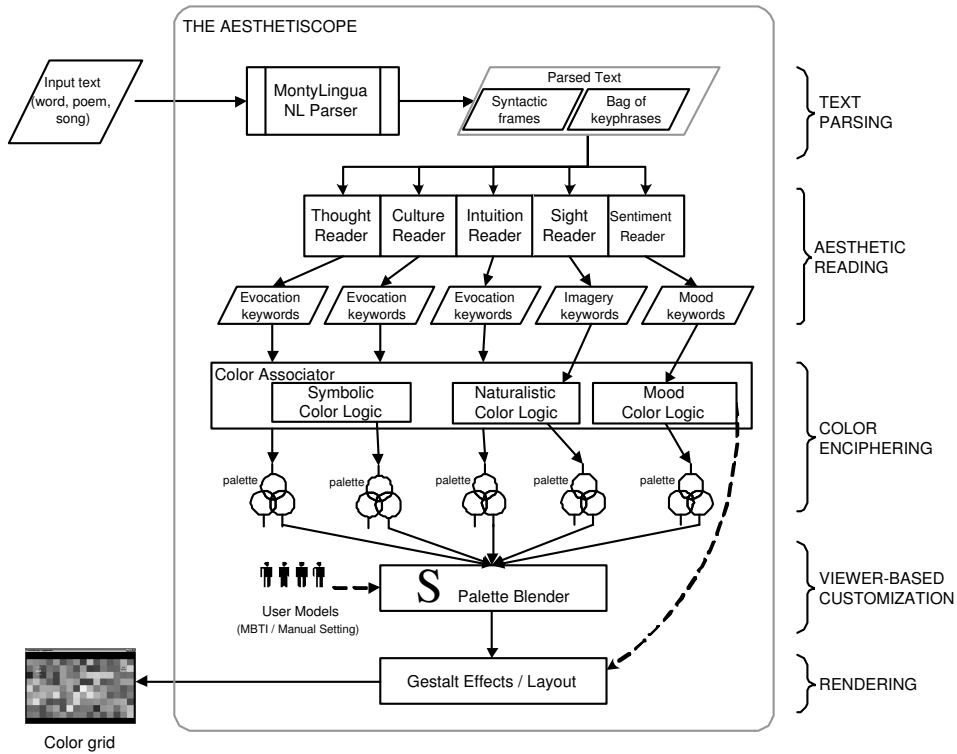


Fig. 5. Process architecture of the implemented Aesthetiscope.

e.g. (this has been simplified):

“Some say the world will end in fire” \implies
 FRAME1: {VERB: say, SUBJECT: some, OBJ1: FRAME2};
 FRAME2: {VERB: end, SUBJECT: world, OBJ1: in fire}

The unstructured back-off parse just extracts from the text a “bag” of important keyphrases, sans a “stop list” of very common semantically confounded words, e.g.:

“From what I’ve tasted of desire” \implies **taste, desire**

ThoughtReader, **SentimentReader**, and **CultureReader** know how to exploit the structured output, while **SightReader** and **IntuitionReader** only utilize the backoff output. In stage two, Aesthetic Reading, the pieces of the text digested by the parser are evaluated through the different interpretive lenses of five Readers, each Reader generating as a by-product of its understanding a bag of *evocation keywords*—as if to imagine that each Reader, while reading the text, had evoked in its mind a set of concepts, e.g. (only top few keywords from each Reader’s actual output are shown):

The poem “Fire and Ice” by Robert Frost \implies

ThoughtReader: earth, cold, hot

CultureReader: crazy, fashion, racism

SightReader: photos of fire, photos of world, photos of ice

IntuitionReader: hot, engine, red, freezing, summer

SentimentReader: arousing, pleasurable, passionate

While the entirety of Section 4 is devoted to a deeper exposition of the internal workings of the Readers, we will say here that the design decision to represent the individual Reader outputs as bags of keywords is intended to make computation easier. A bag of keywords may be a reductive form to evidence *understanding*, but the homogeneity of the keyword form allows for much more uniform translation of the interpretation’s output into color space, and also allows the contributions of the interpretations to be weighted and combined easily without further conflict arbitration between interpretations. Also, representing understandings with bags of keywords is consistent with the spirit that aesthetic is impressionistic in nature—bits and pieces of partial understandings and influences from sight, thought, feeling, intuition, and culture swirl together in a signature proportion (i.e. the aesthetic sensibility of the artist) to shape an artwork.

The latter three stages of Aesthetoscope’s processing are Color Enciphering, Viewer-Based Customization, and Rendering. Color Enciphering translates the evocation keywords outputted by of the Readers into color palettes. We are conscious to call this process *encipherment* to reflect that we are operationalizing the Final Resonance Principle’s (Section 2.2.1) suggestion that color space be viewed as an aesthetic code which invites a viewer to uncover its underlying significance for himself. Viewer-Based Customization takes the color palettes consequent to each Reader interpretation and decides in what proportion to blend the palettes to produce a single palette. Currently the percentage contribution of each Reader is set manually with graphical slider bars in the Aesthetoscope graphical user interface, but it is also reasonable to automate this customization based on the input of a particular user’s MBTI personality profile, as discussed in Section 2.4. Finally, in the last stage, Rendering, the palette is coordinated around some gestalt parameters, e.g. to dim all the colors, to fade all the colors, to lay out the colors in the grid to maximize contrast or to minimize it. Instructions for what gestalt operations, if any, are to occur, source from the “Mood Color Logic” module in the Color Enciphering layer. If **SentimentReader** makes a contribution above a certain threshold of all Reader contributions (50%, in the current implementation), then the mood keywords outputted by the **SentimentReader** will drive the gestalt operations on the final palette. This harkens to Jung (1921), who had described feeling as a judgmental faculty, which would override sensation and intuition when judgment dominated over perception. In Section 5, an expanded discussion of the evocation keyword to color-space mapping process is given. Finally, the final color palette is rendered in the 16 wide by 9 tall color grid and the artwork is complete!

4. The Aesthetic Readers

This section dives into rationale behind, and implementation mechanics of, the five evocative Readers at the heart of Aesthetiscope’s aesthetic reading. We preface this discussion with some general observations.

The choice of these five Readers is in the spirit of aesthetic reading because together, they intend to uncover all the different ways that a text can result in artwork. Jung proposed that four fundamental modes of interpretation—Thinking, Feeling, Intuiting, and Sensing – were a sufficient vocabulary to describe all the different ways that a person might engage the world, and so by proposing five Readers to read a text (inspired by Barthes, we added the `CultureReader` to Jung’s model), we hope to anticipate most of the ways that a hypothetical artist might appropriate a text to create art. In the interest of facilitating computation, we have left out the influence of an artist’s personal memories and experiences and imagery in creating the artwork, in favor of commonsensical interpretation, e.g. Aesthetiscope would express “dirt” as brown and yellow, recalling common sense, rather than idiosyncratic personal experience. The Exclusivity Principle (Section 2.2.2) tells us that a side effect of creating art using common sense rather than personal experience is that the artwork loses a certain intensity of aesthetic appeal. However, under our framework, exclusivity can be restored to some degree by Viewer-Based Customization under the premise that Aesthetiscope can make its artwork customized to particular personality types.

The five Readers, while focused on different interpretations, are not completely orthogonal and will tend to overlap in some interpretations. For example, both `ThoughtReader` and `IntuitionReader` will react to the text “fire” with the evocation keyword “hot” perhaps because this evocation is both rational, and intuitive. Also, in the absence of Jung giving precise criteria for what constitutes the boundaries of thinking, feeling, sensing, and intuiting, we can only claim that our implementation adheres to the spirit of these ideas. Undoubtedly there are a myriad of other ways we might have implemented these Readers. One common aspect of the five implemented Readers is that they employ *associative* or *contextual* reasoning, which is semantically broad rather than deep. We feel that the nature of associations makes them very suitable for brainstorming the aesthetic potential of a text.

The remainder of this section discusses the mechanics and implementation of `ThoughtReader` (4.1), `CultureReader` (4.2), `SightReader` (4.3), `IntuitionReader` (4.4), and `SentimentReader` (4.5).

4.1. *ThoughtReader*

We interpret rationality –dealing with information in an explicit, structured, and logical manner– as a quintessential essence of Jung’s Thinking mode, even though the acts of sentimental interpretation of text, and recognizing imagery in text also arguably engage thinking. From this, we chose the ConceptNet commonsense reasoning system (Liu & Singh 2004) as a framework well-suited for computing rational

evocations of an input text. ConceptNet is a semantic network containing more than 100,000 commonsense concept nodes (e.g. `lemon`, `swim`, `eat sandwich`), interconnected by 1.6 million semantic edges (e.g. `EffectOf(be hungry, eat sandwich)`). Each edge is a common sense fact. ConceptNet is a machine-computable common sense representation, automatically mined from the 800,000 common sense facts in the Open Mind Common Sense (OMCS) Knowledge Base (Singh *et al.* 2002); each fact is expressed as an English sentence. ConceptNet is an ideal source of rational reasoning because the knowledge in OMCS represents some form of common consensus between 15,000 web contributors to the project about how people, things, and events affect each other in the everyday world. For the interested reader, (Liu & Singh 2004) contains examples of the types of common sense inferences made by ConceptNet. Alternative large-scale rational reasoning platforms which we have also considered for **ThoughtReader** include the Cyc Project (Lenat 1995), and the ThoughtTreasure Project (Mueller 2000). ConceptNet and Cyc are the largest publicly available common sense reasoning platforms, and would be to some extent interchangeable as **ThoughtReaders**. Figure 6 depicts the I/O process model for **ThoughtReader**'s implementation using ConceptNet.

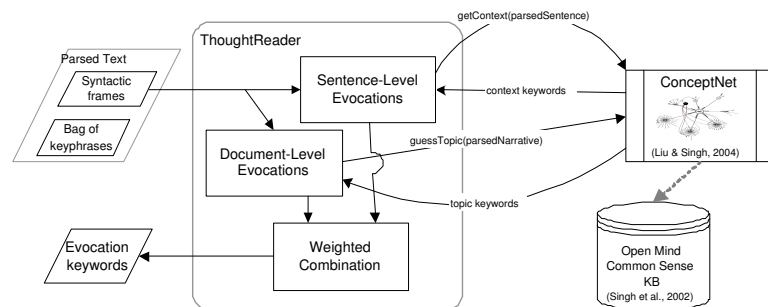


Fig. 6. Input-Output process model of **ThoughtReader**

ConceptNet is both a semantic network of common sense knowledge, and also a reasoning toolkit. It reasons contextually, by the method of spreading activation (Collins & Loftus 1975) away from seed concept nodes fed to it as input. **ThoughtReader** computes rational evocations of a narrative text at two different levels of granularity. It computes rational evocation keywords in reaction to each sentence, but the bigger picture about a narrative should not be missed either, so **ThoughtReader** also computes document-level evocations, which are the topic keywords that best summarize the contents of the narrative text. **ThoughtReader** interfaces with ConceptNet through two calling functions. First, `getContext(parsedSentence)` is called for every sentence of the input text, and the return value is a rank-ordered list of keywords, e.g. (actual top results shown):

20 *Hugo Liu & Pattie Maes*

```

ConceptNet.getContext( MontyLingua.parse("the boy threw the Frisbee
                        to the dog") )
    ⇒
    "frisbee", "play", "run after ball", "throw", "park"

```

Second, `guessTopic(parsedNarrative)` is called once for the whole input text, and the return value is a rank-ordered list of the most important topic keywords in the text, e.g.:

```

ConceptNet.guessTopic( MontyLingua.parse(FireAndIceByRobertFrost) )
    ⇒
    "fire", "desire", "ice", "know", "world", "perish", "stop", "kill"

```

`ThoughtReader` merges the sentence-level keywords and the document-level topic keywords into a single evocation keywords list to output. The document-level topic keywords are given greater weight in the combination process.

4.2. *CultureReader*

Semiotician Roland Barthes' structuralist theory of culture declared that, in its essence, each culture can be represented as a sign system (1964), where signifiers correlate to *signifieds*, and the nature of the correlations is dependent upon the value system of each culture. For example, "*sex*" signifies something negative and taboo in a religious culture, but not in a more socially progressive culture.

Using this simple representation of culture, we have begun to compute cultural models for some broad cultural groups like American pop culture, Roman Catholic culture, and the culture of the American feminist movement. We do so using the What Would They Think? (WWTT) system (Liu & Maes 2004), which is capable of compiling together a model of a person or group's attitudes toward various subjects by automated analysis of a corpus of texts compiled on the person or group. WWTT employs reinforcement-based machine learning to acquire a cultural model from a text corpus exemplifying the viewpoints of the desired group.

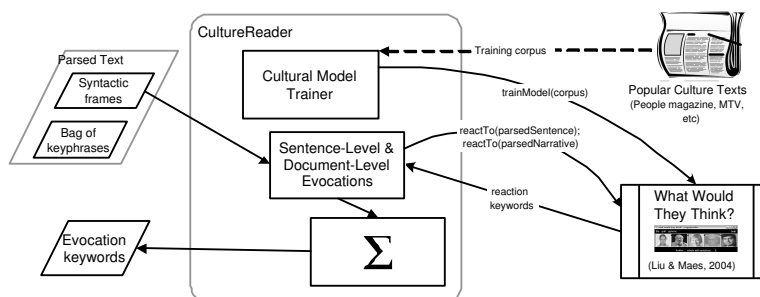


Fig. 7. Input-Output process model of `CultureReader`

A cultural model is a system of attitudes, either hierarchically consistent and organized, or just a bag of attitudes at its crudest. An attitude is represented computationally as a topic-affect pair, and can be thought of as some feeling about some topic. WWTT is equipped with a topic spotter and a textual affect sensor, and attitudes are learned from the text by detecting that certain topics are consistently talked about from a particular affective stance; for example, “movie stars” in American pop culture, signifies “wealth,” “glamour,” “good,” “popular”, etc., and this affective stance is one of high arousal, high pleasure.

We suggest that a system like WWTT fulfills the spirit of a Reader whose objective is to read through a cultural lens and produce reactions from the position of a cultural participant. To our knowledge, there is no alternative system specialized to the purpose of acquiring a cultural model automatically from a text corpus. Figure 7 depicts the I/O process model for **CultureReader**. We have been exploring the idea that in the future, the Aesthetoscope should be able to load the cultural models possessed by the viewer, dynamically. However, for Aesthetoscope’s current implementation, we use only one cultural model—that for American popular culture, acquired automatically by WWTT from a 500kilobyte text corpus consisting of news articles from a variety of popular periodicals such as People Magazine, MTV News, etc. Once WWTT has acquired the American pop culture model, **CultureReader** passes text to WWTT and receives keyword reactions from it. As with **ThoughtReader**, **CultureReader** reacts to each sentence in the input text, and also to the narrative as a whole. The reactions are weighted and summed into a single bag of evocation keywords.

It was necessary to modify WWTT to accommodate our required output format. WWTT normally reacts by emoting a numerical affect score obeying the third-dimensional PAD (pleasure-arousal-dominance) model for affect (Mehrabian 1995b). We modified WWTT so that in lieu of a score, WWTT would emote affective keywords the system learned during the cultural model training phase. So for example, given the stimulus **movie stars**, rather than emoting a numerical score equivalent of high-arousal and high-pleasure, the system would emote the keywords **wealth**, **glamour**, **good**, **popular**, which are the original affect keywords associated with **movie stars** in the text corpus.

4.3. *SightReader*

In Jung’s original four fundamental modes, perceivers inclined toward Sensing were those who relied heavily on the five senses –sight, sound, smell, taste, and touch– to interpret the world. In our current research, we only explore sight because our artwork deals with colors, and the mapping from visual imagery to colors was the most direct (though the other senses could demonstrate interesting synaesthetic mappings to color, or mappings mediated by affect).

To create a corpus of visual memories, we collected 100,000 images from several keyword-annotated stock photography collections, and for each keyword, we sam-

pled out the color palette epitomes from the photo collection. So, for example, **taxi** would have the color epitome of some yellows (sourcing from photos of New York City taxis), **wedding** would have black (the groom), white (the bride, the cake), and some colors (the flowers), etc. Of course, the constitution of the stock photo collection should be considered culture-specific because weddings in Asia have a lot of red, and taxis have no consistent color in many parts of the world.

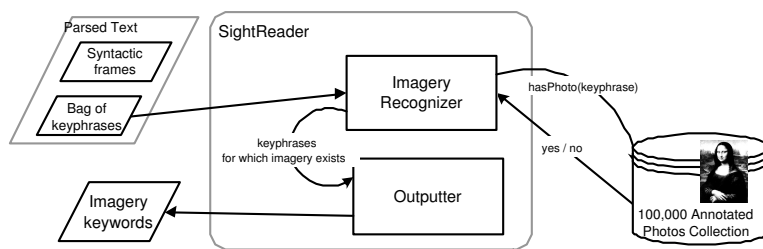


Fig. 8. Input-Output process model of *SightReader*

SightReader's implementation is direct and lightweight (Figure 8). It utilizes only the bag-of-keyphrases parse of the input text. A recognizer filters to keep just the subset of the keyphrases for which color epitomes exist in the photo database. These keyphrases are formatted by the outputter from x to **photos of x** , e.g. from **taxi** to **photos of taxi**. In the color enciphering stage of processing, all phrases with the **photos of x** syntax will be mapped into color epitomes.

4.4. *IntuitionReader*

Intuition can be difficult to characterize because the word has been historically appropriated to refer to many qualities of a person. Some, like F.W.J.v. Schelling and Arthur Schopenhauer, have used the word in opposition to intellectual intelligence, to suggest that it is a form of understanding which is metaphysically sourced. We interpret intuition and intuitive agency more in line with Henri Bergson and the consciousness psychologist George Mandler, and feel that this interpretation is also most in the spirit of Jung's intention. Bergson called intuition 'immediate consciousness', and "the direct vision of the mind by the mind—nothing intervening, no refraction through the prism, one of whose facets is space and another, language" (Bergson 1946, 32). Mandler (1980) distinguished between "remembering" and "knowing," characterizing remembrance as a form of recognition based on the explicit retrieval of an episodic memory and its surrounding context, and characterizing knowing as recognizing by familiarity, without conscious retrieval of memories, and with only the sense or feeling of intuition. Intuitive agency, then, can be summarized as psychologically immediate, indeed, instantaneous and reflexive responses to a situation.

One of the ways in which experimental psychology has tried to capture or measure the instantaneous knowledge that people have around concepts is by recording how they freely associate in response to a stimulus. Psychologists Nelson, McEvoy & Schreiber (1998) have compiled together decades worth of research into a corpus of free association norms. For example, in their corpus, the concept “traffic” triggers “car,” “light,” “jam,” “sucks,” “stop,” “noise,” etc. We acknowledge that this measurement is specific to a certain population of people and a certain temporal period. We also give the caveat that *IntuitionReader* is not as a delicate and sensitive has the human faculty for intuition. A human intuition for a narrative would be a convergent response to the *whole* of the narrative; however, this would seem to demand full story understanding, which is an unsolved and Turing-complete problem in Artificial Intelligence. Our *IntuitionReader* lacks this *sensitivity to gestalt* because Nelson, McEvoy & Schreiber’s corpus of free association norms only enables us to respond to each individual concept contained within a text; the input narrative is not treated with the integrity due to the whole but rather, as a loose bag of concepts. To some, this sort of reading will feel to be wildly divergent and psychotic rather than nuancefully convergent and intuitive, but given the difficulty of full story understanding, and the uniqueness of the psychological free norms corpus as a candidate corpus of intuition, we will proceed with these caveats in mind, taking *IntuitionReader* *cum grano salis*.

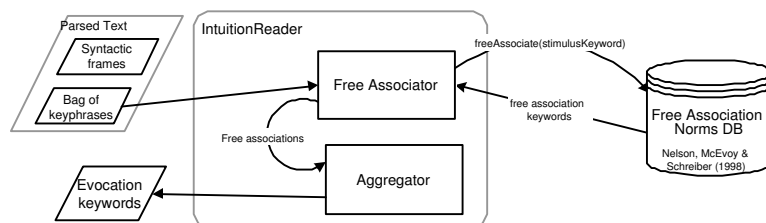


Fig. 9. Input-Output process model of *IntuitionReader*

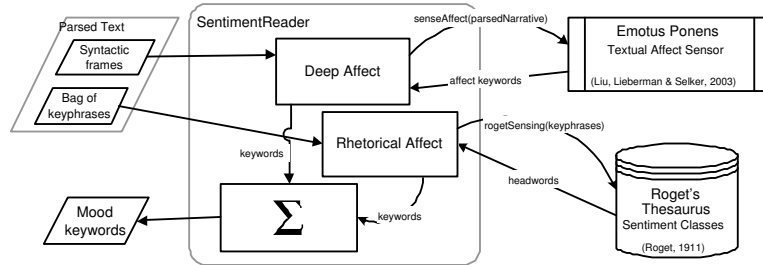
Figure 9 depicts the process model of *IntuitionReader*. We use the free association norms resource more or less at its face value, and the process of intuition in our implementation is closer to spotting for visual imagery than it is to understanding a story coherently. Inputting the narrative text as a bag of keyphrases, a Free Associator passes each keyphrase to the database, and harvests all of the weighted free association keywords that result. An Aggregator merges all the weighted free associations into a single list of evocations, where hopefully, the most common ideas sewn into the narrative subtext can emerge as top evocations.

4.5. *SentimentReader*

An evocative Reader that demonstrates Jung’s Feeling mode of perception is one which is presumably able to empathize with the sentiment contained in and expressed by the text; in other words, *SentimentReader* can be thought to implement textual affect sensing. In the computational literature, there are three main approaches taken to the affective classification of text: the keyword-based approach, the statistical language modelling approach, and the knowledge-based approach. Classifying text by spotting for overtly emotional *mood keywords* like “distressed”, “enraged,” and “sad” is a hand-crafted approach taken by systems like Clark Elliott’s Affective Reasoner (Elliott 1992). While effective at capturing the affect apparent at language’s surface, it does not consider the deep semantics being communicated; for example, a keyword-based approach can register negative affect in the utterance “I had a *terrible* day” yet it would miss the affect in the utterance “I got fired today,” whose affect is more subtextual than it is explicit. Classifying affect using statistical language models (e.g. Deerwester *et al.* 1990) trained up on manually classified text corpora can work quite well on lengthy texts; however, the approach is limited by the fact that only coarse classifications, preferably binary, like happy-unhappy, or inflammatory-uninflammatory are shown in the literature to work well. Blending the keyword-based and statistical approaches are classifiers which work on lexical affinity—the assignment of probabilistic affinities toward particular affect classes to non-mood words, e.g. “accident” might be assigned a 75% affinity toward the fear emotion. Pennebaker, Francis, & Booth’s (2001) Linguistic Inquiry and Word Count computer program is a good example of this approach. Finally, knowledge-based approaches such as Liu, Lieberman & Selker’s (2003) Emotus Ponens system use background semantic knowledge to make inferences about a text’s deep semantic structure rather than its surface semantics. Emotus Ponens parses a story into events and evaluates the affective connotations of those events (thus it is sensing the affect of the deep structure of text). For example, “getting into an accident” connotes fear, anger and surprise.

Figure 10 presents the I/O process model of *SentimentReader*. In implementing *SentimentReader*, we opted to make a full-coverage classifier by combining the deep affect sensing of Emotus Ponens with the surface or rhetorical affect sensing of a keyword-based approach. Because a major genre of input narrative we hope to handle are poetic texts, we opted for Peter Roget’s lexical sentiment classification system (1911) on the rhetorical affect end because of its extensive treatment of poetic language. Roget’s 1911 English Thesaurus features a 10,000 word affective lexicon, grouping words under 180 affective *headwords*, which can be thought of as very fine-grained and well nuanced affect classes.

The Deep Affect component feeds a structured parse of the input text to Emotus Ponens, and receiving as a result, a weighted list of affect words (from an ontology of 100 affect words, adapted from Roget’s affective headwords) characterizing the deep affect in the text. The Rhetorical Affect component feeds the unstructured parse

Fig. 10. Input-Output process model of **SentimentReader**

of the input text to Roget’s Thesaurus and computes a weighted list of headwords which best characterizes the text. The outputs of the Deep and Rhetorical Affect components are combined (in the current implementation, they are combined with equal weight), and outputted as **SentimentReader**’s evocations. We should note here that all the evocation keywords source from an ontology of 180 Roget affective headwords. This fact is important and relevant to how these mood evocations are mapped into color space, which is the topic of Section 5.

5. Psycho-Semantic Color Rendering

Section 4 detailed how a narrative text could be computationally mined for its aesthetic potentialities in the five categories of Thought, Culture, Sight, Intuition, and Sentiment, and outputted as vectors of evocation keywords. In this section, we discuss how evocation keywords are mapped into color space. There are three calculi of color logic in the Aesthetoscope implementation: naturalistically sampled colors (e.g. colors of a tree taken from a photo), mood colors (e.g. colors for love and fear), and symbolic colors (e.g. apples are red, the sky is blue). Using various combinations of these calculi, each of the five aesthetic Readers renders input text into color space in unique fashion, as illustrated in Figure 11. Sections 5.2-5.4 describe how each calculus maps keywords into color space, and Section 5.5 describes how colors are blended together into a single palette. But before we present the technique of psycho-semantic color rendering (it is more than just semantic because we are motivated to influence the psychological state of a person), we briefly recapitulate, in Section 5.1, the context and motivation for our approach which we began in Section 2.5.

5.1. Colors as a coding scheme

Why did we choose to render aesthetic impressions of text as color grids? We were not simply propelled by the fact that colors have a long established role in art proper, or that the colors have absorbed a stereotype for “being pretty things.” Our motivation stems from a theoretical framework for understanding aesthetic as a transaction. In Section 2.2.1, we posited the Final Resonance Principle— the

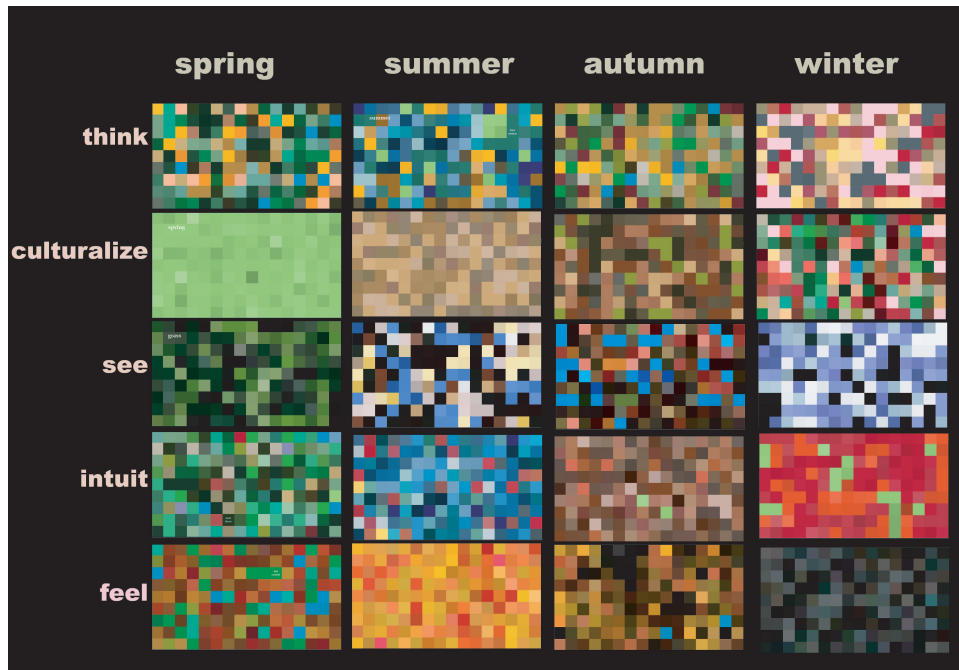


Fig. 11. Aesthetoscope's aesthetic impressions of the four season keywords (columns) rendered through the monadic optics of each **Reader** taken alone (rows).

suggestion that aesthetic is more potent when it is not on the surface but if it must be uncovered by a viewer, harkening to Dewey's suggestion that an experience with art must engage a person into active perception. So we view colors as a particular way of enciphering an artistic message, say, some evocation keywords. Our hypothesis is the following: If people are generally competent mapping from texts to colors and back via the three logics of natural colors, mood colors, and symbolic colors, then Aesthetoscope will encipher evocation keywords into these colors, and invite, *as an aesthetic game*, viewers to decipher the significance of the colors.

We do not suggest that Wittgenstein was right to claim that the heart of all art is a symbolic deciphering game, because calling it a game implies that the artistic creator and viewer are conscious that it is a game, but we do claim that the power of art has always been to cause people to perceive and to perturb them with personal evocations; structuring this process as a symbolic game is an advent of modern art. Clive Bell, an art theorist who was in some sense, anti-representationalist and anti-reductionist in his view of art, described the essence of art as 'significant form', commenting on one painting that "line and colour are used to recount anecdotes, suggest ideas, and indicate the manner and customs of an age: they are not used to provoke aesthetic emotion" (Bell 1914, 18). Even though Bell's aesthetic is non-symbolic and non-representationalist, the 'lines and colors' he describes are aesthetic

precisely because they encode experience and memory (albeit liminally and unconsciously), just as Aesthetiscope's colors encode evocations that a viewer might have from reading a narrative.

5.2. *Naturalistically sampled colors*

The mapping of a fire to its actual colors as seen in the world exemplifies a calculus that would appeal most to a Jungian Sensing individual. We term this type of text-to-color mapping *naturalistic*. In Aesthetiscope's architecture (Figure 5), we show that the output of **SightReader** feeds directly into "Naturalistic Color Logic," because **SightReader** represents the influence of visual memories in aesthetic impression. Naturalistic Color Logic takes imagery keywords and maps them to the actual colors of an imagery, sampled from photos, e.g., **photos of sunset** returns a color palette consisting of strokes of warm hues scattered throughout large swatches of deep purple.

The Naturalistic Color Logic module has a large knowledgebase of palettes for the most common things and events in the world. It is a corpus of what we term the *color epitomes* of things. To implement this knowledgebase, we collected 100,000 low-resolution images (approximately 300x400pixels) from a few large online stock photography collections. The images were already annotated with keywords. For each keyword, we computed the color epitomes for all the photos in the database with that keyword as its primary annotation. We assumed that objects of interest were foregrounded in the image, so we employed epitomic appearance and shape image analysis (Jojic, Frey & Kannan 2003) to isolate foreground objects and to subtract away potential sources of color noise, such as the recurrence of a blue sky, buildings, and roads. We also disqualified all black and white formatted photos, for obvious reasons. Once areas of interest were identified in photos, level histograms were computed for those areas using Hue-Saturation-Lightness channels, a baseline histogram (computed as the summation histogram of all photos in the collection) was subtracted, and centroid colors were identified. Then, actual pixels from the photos were sampled by searching for the nearest neighbour centroid colors in HSL color space. In cases where no satisfactory color epitome could be converged upon, those keywords were disqualified from the knowledgebase. The final knowledgebase has color epitomes for 4,000 keywords, from an initial seed set of 15,000 annotations. We observe that abstract keywords (e.g. love) represent the bulk of keywords for which color epitomes could not be computed, and most of the 4,000 keywords in the knowledgebase refer to concrete things (e.g. taxi, tree, bear). We have given the caveat that this corpus of color epitomes is culturally dependent, the culture being determined as the representational bias of American stock photography collection compilers, e.g. taxis are yellow because urban photos depict primarily New York City in the photo corpus.

5.3. *Mood colors*

Just as Jungian Sensing-inclined individuals might prefer to map imagery into naturalistic colors, so might Jungian Feeling-inclined individuals prefer to read mood into a color grid presentation. As for the story of the artist, the centrality of color as a medium for conveying emotion can be seen prominently in abstract expressionist pieces of Mark Rothko and Josef Albers, who both focused on the emotional entailments of color interactions. Color's unalienable connection to emotion is also strong in the paintings of Paul Cezanne and Henri Matisse. Concerning Matisse's use of color, Susan Sidlauskas wrote, "color is the armature upon which emotion is structured in all its multiplicity, scope, and unseen, but sensed, potential. Cezanne caused color to pulse, occlude, unmask, dramatize, insinuate, unsettle, and solidify" (Sidlauskas, 2004).

In conveying emotion, colors interact richly with one another, and interplay also with form and subject matter, as in Cezanne's sophisticated application of color. However, such interactions are beyond the scope of our present research, where we focus on the psychological mood of colors as the primary communication. Emotion-to-color mapping is primarily a culturally dependent phenomenon, as colors are tied to the metaphors and myth of each culture; for example, white signifies peace and purity in the Occident, but in some Asian cultures, it signifies death and mourning. That being said, modern sensibilities for emotion-to-color semantics are arguably converging as an artefact of globalisation. Also, there is, to a certain extent, as Johann Wolfgang von Goethe wrote about in his *Theory of Colours* (1840), a neurological and physiological universality to our responses to colors. For example, red is physiologically received as being more arousing. In China, pure red is the color of congratulation, whereas in the Occident, pure red is the color of danger, and although the evoked emotions differ in their valence, they share the property of both being high arousal emotions, according to Mehrabian's Pleasure-Arousal-Dominance model of colors (1995b).

The Mood Color Logic module implements a mapping from the select ontology of mood keywords outputted by `SentimentReader` into color space; the mapping is dependent upon the sensibilities of the global cultural bricolage of the contemporary period. This ontology, as introduced in Section 4.5, consists of 180 sentiment *headwords* (categories) devised by Roget in his 1911 Thesaurus. Mappings into color space are achieved heuristically by a handcrafted annotation system we devised, with our interpretation of emotions guided strongly by four texts: Johannnes Itten's *Art of Color* (1961), Josef Albers's *The Interaction of Color* (1963), Eva Heller's *Wie Farben Wirken* (1989), and John Gage's *Color and Culture* (1993). These texts give explicit guidance for the emotional tint of colors. A sampling of the guidance we used to construct our mapping (by hue) is as follows:

RED: arousal, danger, love, exciting, struggle, sin
ORANGE: warmth, friendly, happy, festive
YELLOW: cowardice, sickness, gold, treason, caution

GREEN: nature, youth, envy, spring, growth, corruption, organic
 BLUE: stable, distant, solid, true, loyal, shy, calm, forever
 PURPLE: submission, mystery, passion, metaphysical, royal
 WHITE: pure, light, peace, innocent, joyful, divine, spirit
 BLACK: absence, death, silence, gravity, privacy

Additional guidance from the cross-cultural ethnographic color surveys of Brent Berlin and Paul Kay (1969), and Goethe's color theory helped us to strategically select emotion-to-color mappings which have the greatest potential for cross-cultural recognition. Based on this guidance we annotated Roget's 180 sentiment headwords using terms organized into the following ontology of dimensions, which is an extension of the color space proposed by Albert Munsell (1905):

HUE: e.g. green, brown, blue, purple, red
 TEMPERATURE: e.g. hot, warm, cool, cold
 CHROMA: e.g. colorless, off-primary, primary
 SATURATION: e.g. low, medium, high
 VALUE: e.g. dimmest, dim, medium, bright
 HARMONY: e.g. discordant, harmonious

These dimensions are not orthogonal and thus overlap each other in dominion; however, they provide a broad descriptive vocabulary with which we can characterize colors flexibly. A sample annotation for a Roget headword is given below:

`Inexcitability = harmony-harmonious, temperature-cool,
 hue-blue, chroma-colorless, saturation-medium, value-dimmest`

NB: The color space for our annotations include some guidance for gestalt blending in the color grid like color harmony, and global sensibilities like color temperature and chromaticity. As shown in Figure 5, these gestalt effects are saved and applied to the whole blended palette (after the five Reader's palettes are merged) in the Rendering stage, if and only if the `SentimentReader`'s contribution to the whole artwork surpasses a certain threshold. To operationalize a "discordant" versus an "harmonious" layout, we computationalize a basic prescription from Albers's theory that the hardness of an edge between two color squares be measured as the value difference between the squares; the more hard edges, the more discordant, generally speaking.

5.4. *Symbolic colors*

If naturalistic color logic appeals to the senses, and mood color logic appeals to feeling, then symbolic color logic appeals to the intellect. What color is a school bus, a bee, a smiley face, a traffic light, or the sun? Yellow. Not because they actually are, but because yellow is integral in our culturally iconified notions of

these things. The symbolic imagery and colors of things is reinforced into us by culture, through cartoons, language-learning flashcards, and illustrated children's stories, to name a few vehicles of acculturation. The symbolic color palette is closer to kitsch than it is to subtlety—the colors are pure, stereotyped, linguistic.

The three remaining Readers—**ThoughtReader**, **CultureReader**, and **IntuitionReader**—are rendered into color space partially through the Symbolic Color Logic module—rationality and culture are strongly symbolic, and intuition has at least some symbolic component. They are also rendered partially through Naturalistic Color Logic and Mood Color Logic. The rule used to guide this in the implementation is: Naturalistic Color Logic's role as renderer grows proportionally with the contribution of **SightReader** to the artwork; Mood Color Logic's role as renderer grows proportionally with the contribution of **SentimentReader** to the artwork; and the absence of Sight and Sentiment's dominance implies that Symbolic Color Logic dominates.

Because **ThoughtReader**, **CultureReader**, and **IntuitionReader** can return arbitrary keywords, e.g. **traffic light**, **wealth**, there needs to be a mechanism to force these to map into color space. Here, we use ConceptNet's **PropertyOf** and **PartOf** relations to perform, iteratively if necessary, semantic expansion on these arbitrary keywords until a color word can be arrived at. For example, ConceptNet knows that a **traffic light** has the properties: **red**, **yellow**, **green**; and that **wealth** has the property **desirable** which we can in turn map into color space using Mood Color Logic.

5.5. *Blending palettes*

The five Readers' color palettes are joined statistically. In Sections 2.4 and 3.2, we describe how an MBTI personality inventory user model might in the future be used to drive the proportions for palette blending, but currently, blending is dictated by manually setting the percentage contribution of each Reader (from 0 to 100%) to the artwork. These contribution percentages create a probability distribution with which the final color palette is sampled. As Figure 12 illustrates, biasing the Aesthetiscope toward certain readings can dramatically affect the final artwork.

After the final palette is selected, gestalt considerations from Mood Color Logic may be applied, dictating overall color harmony, chroma, and temperature. Other than those considerations, colors are laid out randomly, and subjected to local color clustering optimizations performed in windows of 3x3 squares. This is meant to reduce the brutal noisy appearance associated with uniform distributions.

6. Evaluation

Since our initial implementation and installation of the Aesthetiscope, we have received many ideas from psychologists, designers, colorists, and hundreds of real people on how to improve the Aesthetiscope, and since then the piece has undergone multiple iterations of redesign. In a companion paper (Liu & Maes 2005), we

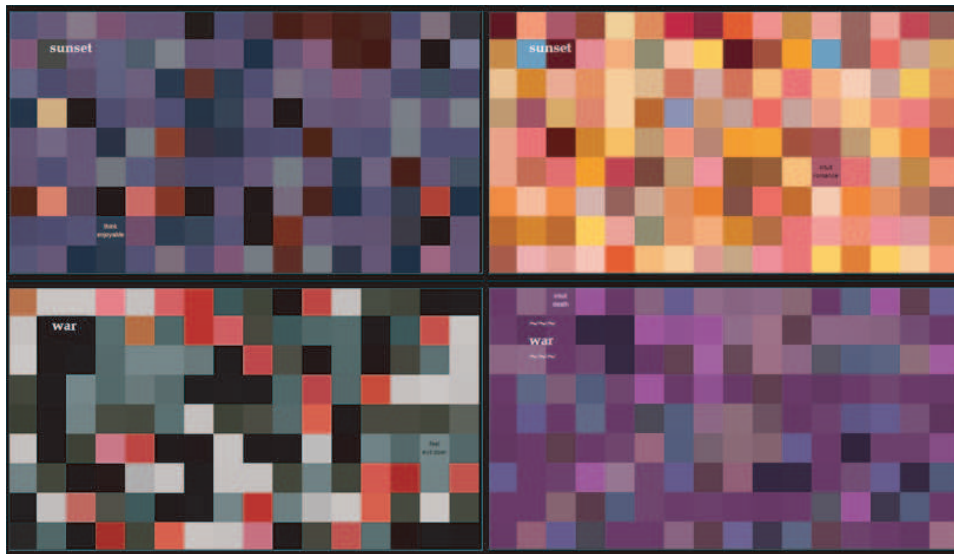


Fig. 12. The words “sunset” (top-row) and “war” (bottom-row) rendered with a Thinking-Seeing bias (left-column) versus with an Intuiting-Feeling bias (right-column).

reflect more upon these redesigns. We have also received suggestions on how best to evaluate the Aesthetoscope, as that seemed particularly problematic because aesthetics is often such a slippery matter. The visual artists we spoke with expressed doubt that aesthetic efficacy could ever be proven in a controlled experiment; they suggested that it be studied ethnographically. One human-computer interaction specialist encouraged us to just issue a survey to see how people liked the Aesthetoscope regardless of its innards. In light of the fact that this paper has focused on aesthetic transactions, their efficacy, and the communication of meanings through the color code, we opted for an information-theoretic set of two evaluations. The first evaluation aimed to measure the signalling efficacy of each of the five reading dimensions. The second evaluation aimed to measure the aesthetic efficacy of a golden combination of the five reading dimensions that seemed to perform best across all viewers.

6.1. *Signalling efficacy of single reading dimensions*

In the first evaluation, four human judges, all graduate students in science, art, or architecture, were asked to score Aesthetoscope renditions of 100 commonly known assorted poems and songs (e.g. Browning’s “How Do I Love Thee?”, first passage of “The Raven,” “I Know Why the Caged Bird Sings,” “I Can’t Get No Satisfaction”, Lenin’s “Imagine”, “Good Vibrations”), most in the range of 150-400 words, and 100 evocative common words (e.g. “God,” “money,” “power,” “success,” “crime”) chosen dispassionately by the examiner but with care to maintain diversity. Because some words were potentially unknown to Aesthetoscope, the examiner discarded

Table 1. Results of depth evaluation of aesthetic impressions from five reading dimensions.

	<i>Plausibility – 100 Poems/Songs</i>					<i>Plausibility – 100 Evocative Words</i>				
	think	culture	see	intuit	feel	think	culture	see	intuit	feel
Judge1	2.3	2.2	3.6	3.6	3.8	3.0	2.6	3.1	4.0	3.5
Judge2	2.0	2.3	3.3	3.3	3.8	2.5	1.8	2.9	3.5	3.6
Judge3	1.8	1.9	3.1	2.6	3.5	1.9	2.0	2.3	3.6	4.0
Judge4	2.5	2.3	3.7	3.4	4.3	2.6	2.5	2.6	3.5	4.5
Avg Score	2.2	2.2	3.4	3.2	3.8	2.5	2.2	2.7	3.6	3.9
Avg StdDev	±0.9	±0.7	±0.6	±0.8	±0.7	±1.1	±1.2	±1.6	±1.0	±0.8
Kappa (avg)	0.31	0.33	0.51	0.40	0.56	0.48	0.42	0.68	0.70	0.75

unknown words and replaced them until 100 known words were arrived at. Image sets of the text laid over the color grid rendition (so judges could refamiliarize themselves with the text) were precomputed for these 200 renditions. Each set contained five images, each image visualizing one of the reading dimensions. Judges were asked to score each of the 1000 total images on the following instruction: “How plausibly does this artwork communicate the thoughts|cultural notion|imagery|free intuition|feelings you had of this text?” Scores were recorded on a standard Likert 1-5 scale (1=not plausibly, 5=very plausibly). Kappa coefficients, a commonly used measure of inter-rater agreement in classification tasks, were calculated between every pair of judges, and the average scores computed. We relaxed the definition of agreement as two judges giving Likert scores with difference 0 or 1.

The results (Table 1) suggest that renditions from Think and Culturalize were fairly poor insofar as they fell short of employing colors to manifest the judges’ Think and Culturalize readings of the text. Renditions from See were fairly plausible in the poems/songs task, but very inconsistent on the word task; its very high average standard deviation of 1.6 on words suggests that it completely failed to visualize some abstract words, e.g. “power,” while succeeding perfectly on words corresponding to concrete things. Intuit and Feel performed the best, and were consistently plausible in their renditions. Standard deviations trended higher on the word task, while the average scores were on par with the poems/songs task – this indicates that each reading was more brittle on just the one word input; however, when a rendition was successful, it was more intensely successful on the one-word input than for poems/songs. The average Kappa statistics (0=pure chance, 1=perfect agreement) indicate a fair to good agreement amongst the judges, with the greatest convergence of opinion around Feel, and demonstrating greater agreement in the word task than in the poems/songs task. These results are promising, but reveal that Think and Culturalize lead to weak renditions; however, because these categories also saw the lowest inter-rater agreement scores, we could conclude that either 1) these are difficult dimensions to computationalize for a general public, and we should try to personalize these models; or 2) these are dimensions not generally amenable to expression in color space, and perhaps colors are not strong enough stand alone signals for these dimensions, perhaps form is also required.

6.2. Aesthetic efficacy

From the first evaluation, we learned that the strengths of the aesthetic readings and renderings lied in See, Intuit, and Feel. In this second evaluation, we wanted to test the aesthetic efficacy of Aesthetiscope— that is to say, can Aesthetiscope produce a satisfying color impression of a text in a non-arbitrary manner? To avoid complication for which we are not currently prepared, we do not try to correlate personality types with customized presentations of Aesthetiscope, but rather, we have chosen to use a Golden Setting, a manual setting of Think10%-Culturalize10%-See40%-Intuit50%-Feel70% which seems to be, from our experience, the most winning combination. Because a viewer's satisfaction with Aesthetiscope's renditions can be hard to normalize and the self-assessment can be difficult for viewers, we present viewers with a choice. Taking the text from the 100 poems/songs, and 100 words, we overlaid each text over its own Golden Setting rendition, and also over the Golden Setting rendition based on another random text within the same category (poems/songs and words are separate categories). This randomized rendition should control for, *inter alia*, the form of Aesthetiscope's presentation, and should help to isolate measurement to just the ability of the Golden Setting to judiciously and aesthetically express the gestalt of the text. Since the Golden Setting mixes influences, the gestalt artwork is harder to decompose into component signals, when viewed at-a-glance. Fifty-one¹ undergraduate students from MIT and Harvard University were each asked to make twenty at-a-glance (under ten seconds) binary judgements on randomly selected items in each of the two task categories: poems/songs, and words. The instruction was: "this text inspired which of these two artworks?"

The results were as follows: in the poems/songs category, the Golden Setting was identified as the artwork with an accuracy of 75.2% across all judges; in the words category, the Golden Setting was identified as the artwork with an accuracy of 80.7% across all judges. Kappa statistics could not be calculated because each volunteer judged a randomly selected subset of the available renditions. With these results, we gain a measure of confidence that Aesthetiscope's color renditions produce an aesthetic in the vein of art, and its aesthetic is demonstrably and non-arbitrarily tied to, and inspired by a reading of a text.

7. Conclusion

This paper addressed the provocative question of how a computer might become a visual artist, rendering aesthetic impressions of a text as an abstract color grid, and standing atop the shoulder of artists like Ellsworth Kelly and Herman de Vries. We proposed a computational framework for understanding the aesthetic quality of art as a type of transaction between an artwork's message and a model of the

¹Since an earlier reviewed draft of this manuscript, we have extended the study from its original participation size of twenty-six MIT undergraduates. For some, this sampling may still be too skewed or too small to draw definitive conclusions from.

viewer. We implemented an artbot called the Aesthetiscope to realize and test this mechanistic account of art experience. The Aesthetiscope renders aesthetic impressions of text as a 16x9 grid of colors, and emulates the creative process of a visual artist. First, evocative readings of an inspiration text (word or poem) –reading along five Jung-inspired dimensions of Think, Culturalize, See, Intuit, and Feel–expose the aesthetic potential of the text. Second, various psycho-semantic color logics map these evocative readings into color palettes. Third, the five color palettes are merged into the final artwork by considering that certain reading dimensions appeal more heavily to certain personality types. We theorize that the reason for mapping textual evocations into color space is to encipher them, and that giving the viewer something to discover, unwrap, and learn is commensurate to seducing them into perception and experience, which Dewey characterizes as the essence of aesthetic.

In evaluating the Aesthetiscope with human judges, we found that Think and Culturalize tended to be ineffective at producing meaningful color grids for viewers, while Feel produced the most consistently meaningful renditions. In a second evaluation of Aesthetiscope-generated artworks against a randomized control, we successfully demonstrated that the aesthetic of Aesthetiscope’s artwork is not arbitrary, but demonstrably inspired by a reading of text. We view our contribution as a salutary and successful foray in applying Artificial Intelligence tools to the subject area of aesthetics, which is traditionally considered to be a bastion of human emotional intelligence. By demonstrating within our prototype system that computers can control the aesthetics they project, we would like to embolden more research in this vein, which promises to extend the reach of how Artificial Intelligence will be able to touch and enrich all our lives in the future.

Acknowledgments

We thank Walter Bender for inspiring the Aesthetiscope with his and Jon Orwant’s incredibly mythical and addictive Color Deducto game. We are also grateful to Judith Donath, Glorianna Davenport, Jeffrey Huang, and John Maeda for fruitful discussions and suggestions within the scope of this work. This research was supported by fellowships from American On-Line and British Telecom, and by the sponsors of the MIT Media Laboratory. Finally, we acknowledge the anonymous reviewers of our FLAIRS paper and of this manuscript for their helpful feedback.

References

1. ADORNO, T.: 1970/1998, *Aesthetic Theory*, University of Minnesota Press.
2. ALBERS, J.: 1963, *Interaction of Color*, Yale University Press.
3. BARTHES, R.: 1964/1967, *Elements of Semiology (Translated by Annette Lavers & Colin Smith)*, London: Jonathan Cape.
4. BAUMGARTEN, A.G.: 1735/1954, *Reflections on Poetry*, trans. K. Aschenbrenner and W.B. Holther, Berkeley: University of California Press.

5. BERGSON, H.: 1946/2002, *The Creative Mind: An Introduction to Metaphysics* (Trans: Mabelle L. Andison), New York: Citadel Press.
6. BELL, C.: 1914/1987, *Art*, Oxford: Oxford University Press.
7. BERLIN, B. & P. KAY: 1969, *Basic Color Terms*, Berkeley and Los Angeles: University of California Press.
8. BEST, D.: 1985, *Feeling and Reason in the Arts*. London: George Allen and Unwin.
9. BIRKHOFF, G.D.: 1932, *Aesthetic Measure*, Harvard University Press, Cambridge, MA.
10. BOURDIEU, P.: 1984, *Distinction: A Social Critique of the Judgement of Taste*, London: Routledge.
11. BRIGGS, K.C., & I.B. MYERS: 1976, *Myers-Briggs Type Indicator*, Palo Alto, CA: Consulting Psychologists Press
12. COLLINS, A. M., AND E. F. LOFTUS: 1975, A spreading-activation theory of semantic processing, *Psychological Review* 82: 407-428.
13. DEWEY, J.: 1934, *Art as Experience*, New York: Perigee Books.
14. DEERWESTER, S., S.T. DUMAIS, G.W. FURNAS, T.K. LANDAUER, R. HARSHMAN: 1990, Indexing by Latent Semantic Analysis, *Journal of the Society for Information Science*, 41(6), 391-407.
15. ELLIOTT, C.: 1992, *The Affective Reasoner: A Process Model of Emotions in a Multi-agent System*, PhD thesis, Northwestern University, May 1992. The Institute for the Learning Sciences, Technical Report No. 32.
16. FREUD, S.: 1919/1990, The 'uncanny', in Dickson, A. (Ed.), *The Penguin Freud Library, Vol. 14: Art and Literature*: 335-6, Penguin, Harmondsworth.
17. GAGE, J.: 1993, *Color and Culture: Practice and Meaning from Antiquity to Abstraction*, University of California Press.
18. GOETHE, J.W.v.: 1840/1970. *Theory of Colours*, trans. C. L. Eastlake, Cambridge, MA: MIT Press.
19. HEGEL, G.W.F.: 1838/1975, *Aesthetics: Lectures on Fine Art*, trans. T. M. Knox, 2 vols, Oxford University Press (originally lectured 1835-1838).
20. HELLER, E.: 1989, *Wie Farben Wirken: Farbpsychologie, Farbsymbolik, Kreative Farbgestaltung*, Rowohlt Verlag, Reibek bei Hamburg.
21. ITTEN, J.: 1961, *The art of color: the subjective experience and objective rationale of color*, New York: Van Nostrand Reinhold.
22. JAKOBSON, R.: 1960, Closing Statements: Linguistics and Poetics, in: T.A. Sebeok, *Style In Language*: 350-377, MIT Press.
23. JOJIC, N., B. FREY & A. KANNAN: 2003, Epitomic Analysis of Appearance and Shape, *Proceedings of ICCV 2003*.
24. JUNG, C. G.: 1921/1971, *Psychological Types*, trans. by H. G. Baynes, Princeton, NJ: Princeton University Press.
25. KANT, I.: 1790/1929, Critique of judgement, in T. M. Greene (Ed.), *Kant selections*: 375-432, New York: Scribner.
26. KRAUSS, R.: 1979/1985, *The Originality of the Avant-Garde and Other Modernist Myths*, Cambridge, MA: MIT Press.
27. LENAT, D.: 1995, Cyc: a large-scale investment in knowledge infrastructure. *Communications of the ACM* 38(11). ACM Press.
28. LIU, H.: 2004, Articulation, the Letter, and the Spirit in the Aesthetics of Narrative. *Proc. of the 2004 ACM Workshop on Story Representation, Mechanism, and Context*.
29. LIU, H.: 2002, MontyLingua: Commonsense-Enriched NLP, Toolkit and API. Accessed at: <http://web.media.mit.edu/hugo/montylingua/>
30. LIU, H., H. LIEBERMAN AND T. SELKER: 2003, A model of textual affect sensing using real-world knowledge, *Proceedings of the 7th International Conference on Intelligent*

36 *Hugo Liu & Pattie Maes*

User Interfaces, IUI 2003, 125-132, ACM Press.

31. LIU, H. AND P. MAES: 2004, What Would They Think? A Computational Model of Attitudes. *Proc. of the 2004 ACM Conference on Intelligent User Interfaces*, 38-45. ACM Press.
32. LIU, H. AND P. MAES: 2005, The Aesthetiscope: Visualizing Aesthetic Readings of Text in Color Space. *Proceedings of FLAIRS2005*: 74-79, AAAI Press.
33. LIU, H. AND P. SINGH: 2004, ConceptNet: A Practical Commonsense Reasoning Toolkit. *BT Technology Journal 22(4)*, 211-226. Kluwer Academic Publishers.
34. MCCracken, G.: 1988, *Culture & Consumption*, Indiana University Press.
35. MANDLER, G.: 1980, Recognizing: The judgment of prior occurrence, *Psychological Review*, 87: 252-271.
36. MEHRABIAN, A.: 1995b, Framework for a comprehensive description and measurement of emotional states, *Genetic, Social, and General Psychology Monographs*, 121: 339-361.
37. MOORMAN, K. & A. RAM: 1994, A function theory of creative reading. *The Psychgrad Journal*. Technical Report GIT-CC-94/01, Georgia Institute of Technology.
38. MUELLER, E. T.: 2000, *ThoughtTreasure: A natural language/commonsense platform*. Accessed on 11 November 2005 from <http://www.signiform.com/tt/>
39. MUNSELL, A.H.: 1905, *A Color Notation*, Boston.
40. NELSON, D.L., C.L. McEVOY & T.A. SCHREIBER: 1998, *The University of South Florida word association, rhyme, and word fragment norms*. Accessed on 11 November 2005 from <http://www.usf.edu/FreeAssociation/>
41. PAZ, O.: 1978, *Marcel Duchamp, or the Castle of Purity*, New York, Viking Press.
42. PENNEBAKER, J.W., M.E. FRANCIS, & R.J. BOOTH: 2001, *Linguistic Inquiry and Word Count (LIWC): LIWC2001*, Mahwah, NJ: Erlbaum Publishers.
43. POULET, G.: 1980, Criticism and the Experience of Interiority, in J.P. Tompkins (Ed.) *Reader-Response Criticism*.
44. ROGET, P.: 1911, *Roget's Thesaurus of English Words and Phrases*. Accessed on 11 November 2005 from <http://gutenberg.net/etext/10681>.
45. ROSENBLATT, L.: 1978, Efferent and Aesthetic Reading, *The Reader, The Text, The Poem: A Transactional Theory of the Literary Work*: 22-47, Carbondale: Southern Illinois UP.
46. SIDLAUSKAS, S.: 2004, Emotion, Color, Cezanne (The Portrait of Hortense), *Nineteenth-Century Art Worldwide: 3(2)*, AHNCA Press.
47. SINGH, P., T. LIN, E. T. MUELLER, G. LIM, T. PERKINS, W. L. ZHU: 2002, Open Mind Common Sense: knowledge acquisition from the general public. *Proceedings of ODBASE'2002*.
48. STINY, G., J. GIPS: 1978a, *Algorithmic Aesthetics: Computer Models for Criticism and Design in the Arts*, Berkeley: University of California Press.