International Journal on Artificial Intelligence Tools

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 05)
Accepted (DD MONTH YY)

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 Aesthetiscope—a computer program which portrays aesthetic impressions of text and renders an abstract color grid artwork reminiscent of early twentieth century abstract expressionism. Following Freud and Dewey's psychological interpretation of "aesthetic" and Jung's ontology of fundamental perceptions, 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 Aesthetiscope 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 capable of creative narrative understanding, and three color logics which employ psycho-semantic principles to render the aesthetic readings in color space. Evaluation of the Aesthetiscope revealed that the program is best at portraying intuition and feeling, and that overall, the Aesthetiscope is capable of creating the aesthetic of art based on an inspiration text in a non-arbitrary way.

Keywords: aesthetics, text, color psychology, reading, semantic interpretation, generative art

1 Introduction

In 1951, the 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, taken from a different memory in his personal experience. So the colors then have a very personal meaning for Kelly, and the gestalt, or whole, of the colors in 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.

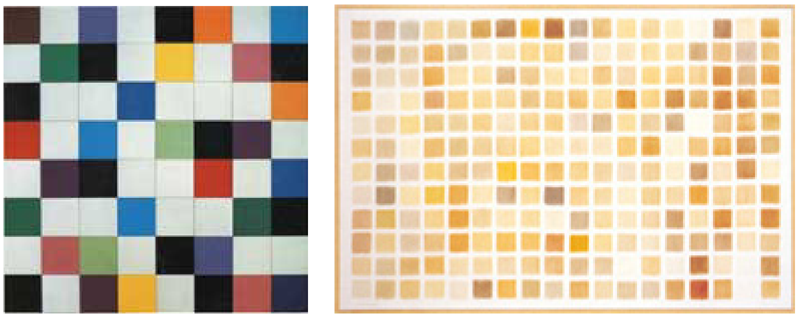

Fig 1. Left: Ellsworth Kelly: *Sixty-Four Panels: Colors for a Large Wall* (1951);
Right: herman de vries: *Terre Provençale* (1991).

But then consider the piece entitled *Terre Provençale* by Dutch artist herman de vries (Figure 1); each color square in its grid is a rubbing with earth from different locations in Provence, France. Whereas Kelly's piece could really only evoke its intended meaning for himself, *Terre Provençale* evokes more broadly than for a single person, potentially evoking specific meaning 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, try to decipher the code, and through this process, the viewer's imagination is stirred, and a resonance of memories, sensations, and emotions is evoked within the viewer.[1]

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 question, we faced many challenging questions. What are the various dimensions of text which contribute to an impression of the text? How should the composition of this 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 do 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 which we theorize in this paper, we have built and evaluated a generative art robot called the Aesthetiscope, which takes an input text such as a word, a poem, or song lyrics, and renders out of it a color grid meant to convey an aesthetic impression of the text which stirs sensations, memories, and emotions in the viewer. Figure 2 should give the reader an initial sense for the sorts of color grids which can be generated by the Aesthetiscope.

[1] As we will later clarify, our understanding of *the aesthetic* is away from Kantian formalism, and more in the experiential spirit of Freud and Dewey.

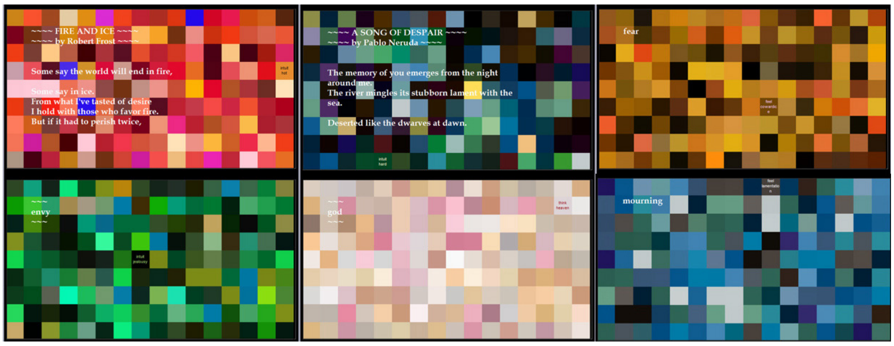

Fig 2. Aesthetic renditions generated by the Aesthetiscope, set with bias toward Feeling and Intuition modes, for the following texts (clockwise from upper left corner): a) the poem “Fire and Ice” by Robert Frost; b) the poem “A Song of Despair” by Pablo Neruda; c) the word ‘fear’; d) the word ‘mourning’;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

Our *artbot* works through the following mechanism. Based on theories of aesthetic and creative reading—that is, reading which more fully engages the imagination, feeling, and sensation (Rosenblatt, 1978; Moorman & Ram, 1994)— and based on Carl Jung’s theory that people interpret reality through a few fundamental modes, i.e., thinking, feeling, sensation, and intuition (1921), our artbot reads a narrative text in not one but five ways, reading rationally, emotionally, intuitively, culturally, and visually; the artbot uses various heuristics from color psychology to map those five textual interpretations into the world of colors; then finally, it blends the color palettes to fill a color grid. To account for the observation that some people prefer more emotional interpretations while others prefer more visual interpretations, all textual interpretations will contribute equally to the artwork in the blending process.

To generate the aesthetic interpretation of the input, five robotic readers skim the narrative text, reading for evocations – things in the text which remind each reader of something else of concern to them. Each of the readers is reading from a different aesthetic standpoint (from a different Jungian mode of interpretation), and each outputs a bag of keyword evocations, which represents some measure of evidence for the way the reader has understood the text. 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.
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 (currently, the associations source from popular culture magazines in the United States) like “world→crazy”, “desire→fashion”, “hate→racism”; SightReader extracts from the text, objects for which visual imagery exists (in a large collection of 100,000 stock images) such as "fire→photos of fire", "world→photos of world"; IntuitionReader makes psychologically immediate free associations like "fire→hot", "fire→engine", "fire→red"; and 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 common sense associations, that is to say, these are associations which are meaningful to a community or culture of people rather than just to a single person; in this sense, our current approach is more in line with the aesthetic technique of herman de vries's *Terre Provençale* rather than 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. Within the Aesthetic Theory literature, a computational model and implementation of aesthetic evocation would put art criticism within the reach of direct scientific exploration and experimentation, and we suggest that this not be interpreted pejoratively 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."[2] For the Artificial Intelligence community, the prospect of being able to create programs which control how they affect feeling, sensation, and emotion in people could potentially open up a new realm of possibility for how A.I. programs might in the future touch the lives of people; even if this research fails to lay a generic foundation that could direct future computational aesthetic research, 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 computational narrative understanding. Section 5 discusses the psycho-semantics of rendering an evocative color grid in the Aesthetiscope implementation. Section 6 presents evaluation and further discussion of the Aesthetiscope. We conclude in Section 7.

2 A Computational Theory of Aesthetic Impression

In this chapter, we theorize the notion of creating an aesthetic impression for a narrative text. In Section 2.1 we make clear what we mean by the word 'aesthetic', describing *aesthetic experience* using the metaphor of an affected transaction between an artwork and its viewer, and attempting to articulate the principles of aesthetic's efficacy in affecting its viewer. Next, in Section 2.2, we discuss the aesthetic potential of a narrative text – what are the elements of meaning which can be read from a text that might participate in an aesthetic impression? Section 2.3 tackles the issue of how a user model of the viewer can prescribe how the various dimensions of aesthetic interpretation

[2] Donald Knuth, "Computer Science and Mathematics," *American Scientist,* 61, 6 (1973), 709.

might best combine together to form aesthetic impression customized to an individual. In Section 2.4, we discuss the role that colors and the grid format play in the conveyance of aesthetic impression. Section 2.5 summarizes these discussions into a concise thesis about the computation of aesthetic impression.

2.1. *Aesthetic transaction*

To most people, the words *aesthetics* and *aesthetic* evoke 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 idea that aesthetics refers to the formal study of art is perhaps the legacy of Immanuel Kant, who, in his Critique of Aesthetic Judgement (1790), proclaimed that judgement of beauty is more concerned with form than with function or content – so by that logic, a horse is beautiful because of its appearance as a horse, and not beautiful because of the symbolic significance of the horse to the viewer's life or memories. Hence Kant reinvigorated the formalist notion of aesthetics and the Platonist idea that judgements of beauty can be universal, objective and independent of the subject. Under this guise, aesthetics developed into the branch of philosophy that saw itself more often concerned with impersonal and socially formulated art criticism than with *joie de vivre*, or the impact of an artwork on an individual. But Kantian formalism is not how we view aesthetics.

We are concerned with aesthetics as an intimate and personal phenomenon. An artwork's *aesthetic* 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, as Freud put it (1919). For Freud, aesthetic was a much more intimate and narcissistic notion – something is aesthetic if it *moves* us, and we are moved only when we see the resemblance of aspects of ourselves, our memories, and our desires in things, and so Freudian aesthetics is about self-identification in artwork. Dewey, too, is important in having shaped aesthetic as a subjective idea. In *Art as Experience* (1934), Dewey puts forth the thesis that art has the character 'aesthetic' because art has the ability to *seduce* us 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 the true nature of things; in this state, we are highly perceptive, and receptive to sensations (seeing, listening, smelling, tasting), feelings, and our imaginations run wild.

Dewey also initiated a relativistic conceptualization of 'the aesthetic' not as a fixed property of an artwork, but as a transaction between artwork and viewer. 'The aesthetic' of an artwork then, exists as a potential energy of the artwork. Most of what our culture agrees upon as being aesthetic (art, or otherwise) has the potential to transact with a significant fraction of the participants in our culture, but this should not mean that it is in any sense more diminutive for an object to be 'aesthetic' for and to transact with just one person, *cf.* the colors from Kelly's *Sixty-Four Panels* transact with and affect himself in a way different than it affects other viewers. The transactional metaphor for aesthetic also suggests a particular computational model of aesthetic, which we sketch as follows: To model the potential for an aesthetic transaction, we actually need to model two entities: the artwork, and the viewer. Since, as Freud suggests, we are most readily seduced into aesthetic experience by seeing aspects of ourselves and our concerns in an artwork, the artwork's aesthetic potential might be computed as the contextual intersection between the artwork's message and the viewer's concerns. In addition to the basic transaction view of aesthetic experience, we enrich our model with two principles regarding the efficacy of this transaction: final resonance, and exclusivity.

2.1.1. 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, and so should an artwork which simply plays back photographs the viewer has taken. Clearly, there is some other criterion in the secret recipe of aesthetic. The missing piece, we suggest, is *intimacy*. Being obvious is not conducive to the aesthetic because the intimacy of the artwork-viewer transaction is violated. For a viewer to feel affected, she must feel that the experience is intimate and unique and that she has discovered herself or her concern in the art in an unexpected way; or in G.W.F. Hegel's words, art is aesthetic because it provokes thoughts and has intellectual import (*geistiger Gehalt*) (1835-8). The artwork may suggest and facilitate the viewer in making a particular meaningful discovery, but it is the viewer who must take that final step to identify herself in the artwork so that she can feel ownership over the discovery and thus be more intimately affected by the artwork; her initiating the discovery constitutes a possession ritual, according to Grant 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, the viewer's discovery of something extraordinary and personally meaningful in the artwork. In contemporary political art, the final resonance is often the discovery of the punch line of a joke. In early Twentieth century abstract art, the final resonance is the viewer's discovery of whatever her psyche mandates. And in the realm of the Aesthetiscope's color grids with which we are concerned, the initial resonance is the viewer finding the color grid pleasant but ineffable, while final resonance is the moment of the viewer's eventual discovery of the relationship of the colors to the purported subject matter being depicted. The colors obey a semantic code, for they have in them the capacity to signify many things (e.g., the color of a real object, or of a mood). Once the semantic code of the colors has 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) on Ludwig Wittgenstein's aesthetics).

2.1.2 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 anything in another viewer. *Sixty-Four Panels* succeeds strongly in delivering an aesthetic impression of life for Kelly because he associates the colors with his personal memories, but other viewers cannot share the same depth of the evocation that Kelly feels because they have not had his experiences. Certainly other viewers could still try to read personal meaning into Kelly's colors, but here we begin cross the boundary from something which is *aesthetic* to something which is *aestheticized* (and following this line of reasoning we could suggest that the character of all abstract art is that it is meant to be *'read into'* or aestheticized). *Terre Provençale* differed from *Sixty-Four Panels* in that its color grid, being sourced from shades of dirt from the earth in Provence, has the potential to behave aesthetically for a wider audience, either the residents of Provence who have seen those particular earth shades, or all humankind who share a common visual memory of earth's shades. However, we suggest that as an evocation becomes increasingly commonplace and shared by all people (as opposed to being unique and personally significant), intimacy of message is lost and the power of the aesthetic is diminished (Liu, 2004). Thus we come upon the following tradeoff which we call the *exclusivity principle* – we state it as a play on the famous aphorism of P.T. Barnum[3]—a particular color grid 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 a transactional framework on aesthetics, we can say that aesthetic is a transaction or resonance between artwork and viewer, and that the 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 qualified to receive than any arbitrary person; and 2) when the message of art is disguised under some semantic code such as colors, and the deciphering of the code which initiates the transaction, is a discovery reserved for the viewer; in other words, the core message of the aesthetic transaction is a 'pull' by the viewer rather than a 'push' by the artwork, and as a 'pull' the viewer is more affected because he feels greater ownership over the discovery of the art.

2.2. *Aesthetic potential of narratives*

Much of the AI narrative understanding literature subscribes to the dogma that there exists a single rational method of interpreting text, and that resultant interpretations and inferences can always be reconciled into a single consistent world model. One branch of research 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 says that textual understanding should be algorithmized simply as the rote invocation of inference rules. Moorman & Ram's revolt against the grain of the classical AI narrative understanding literature emboldens us in pursuing the idea of computationalizing an *aesthetic reading* of text.

Aesthetic reading is not reading purely for information. It is an affected and sensational reading, whereby the text's primary effect is to evoke aesthetic rumblings within the reader. Reading theorist Louise Rosenblatt states, "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, p. 25); but this notion of "living through" can be quite a complex amalgamate of perceptions and sensations. To view reading within our aforementioned transactional framework, Rosenblatt distinguishes between two types of transactions between a reader and a text – efferent and aesthetic. A reader

[3] 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."

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; and just as a person requires a pail to carry away some water from a river, efferent transactions imply that a reader brings a prior 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 brunt of the narrative's potential to affect; the reader allows herself to become affected and connected to the text—to receive sensations, moods, imaginations from the text.

Before we form an aesthetic impression of a text, we must first identify the raw materials, or the aesthetic potential, of a narrative. At this point, one strong caveat must be given about the approach of our research. There is no doubt that much of the aesthetic potential of a narrative is the potential to stir personal memories, personal imagery, and personal attitudes which are absolutely unique to each viewer. However, our present research does not pursue the personal potential of a narrative, for such a line of research requires access to memories, personal images, and attitudes which we are not prepared to gather. Rather, our approach is to pursue the common sense and collective potentials of a narrative. For example, what is the archetypal sentiment, sensation, thought, and intuition about a narrative text that is shared among our present culture? This is not to say, however, that there is no opportunity to customize the aesthetic impression for individuals, as we will see in the next subsection entitled "User model of the viewer".

Since there is 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 around the inspiration of Carl Jung's Modes of Interpretation (1921), a psychological theory that he put forth to account for all the different possible ways that people interpret the world – it is meant to be a complete description of the possible engagements. According to Jung's theory, there are 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 system. 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 our current scope of research.

Whereas objective reading relies primarily on the Thinking mode; aesthetic reading invites a reader to employ many, or all, of the Modes of Interpretation to engage with the text, each mode producing some evocations; and we can think of a weighted sum of all produced evocations as the aesthetic interpretation of the text. The full aesthetic potential of a narrative text then, can be computed by reading text under a multitude of interpretive lenses – Thinking a text, Feeling it, Sensing it, Intuiting it, and Culturalizing it. This multi-perspectival model of interpretation differs significantly from the monolithic rational interpretation dogma of traditional AI 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.

2.3. *User model of the viewer*

Once the five dimensions of interpretation have been established, the questions comes as to how to best combine, and in what proportions to combine the interpretations into a coherent aesthetic impression of the text. We suggest that the proportions of the Modes of Interpretation should ideally be sensitive and individualized to each viewer. Some viewers are biased to be more sensitive to the emotional aspects of a narrative text, while other viewers are biased to be more sensitive to the visual imagery embedded in a text. Actually, Jung, in proposing his Modes of Interpretation, anticipated that certain individuals are more inclined to engage the world along Thinking and Sensing, while others are more inclined to engage the world in a Feeling and Intuiting capacity. In fact, his Modes of Interpretation theory became the foundation for more contemporary personality type classification schemes like the Myers-Briggs Type Indicator (MBTI) (Briggs & Myers, 1976), which describe the perceptual biases of individuals.

Currently, our Aesthetiscope implementation only allows for the manual adjustment of the proportional contributions of each Mode of Interpretation to the final colors. However, although not currently implemented, we could imagine employing schemes such as Myers-Briggs as a representation of the viewer's user model, and leveraging this representation to make decisions about how to proportion and combine the contributions of the five dimensions of textual interpretation to constitute an aesthetic impression which is capable of most greatly impacting and affecting that viewer. MBTI is the 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. The first three dimensions source from Jung and the fourth is unique to MBTI. To use MBTI as the Aesthetiscope's user model, we would propose the following algorithmic mapping scheme. N-S and F-T have an explicit correspondence to the Aesthetiscope's four out of the five dimensions. P-J, according to MBTI, specifies a person's orientation toward the outer world; a Perceiving person is more sensitive to external opinion, while a Judging person is not as easily influenced by external opinion, and is more likely to be guided by the self's opinions. Although not a perfect mapping is possible from P-J into the Aesthetiscope, we would propose that P be mapped to a high contribution for CultureReader, while J be mapped to a low contribution for CultureReader; the rationale being that a Perceiving person, more sensitive to external opinion, is more likely to be a cultural participant and to inherit the attitudes and viewpoints of the containing culture. Conversely, a Judging person is more likely to ignore external perspectives like culture's interpretation. The I-E MBTI dimension is ignored in our proposed mapping scheme.

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

Five-dimensional aesthetic reading of a narrative text yields five interpretations, which we assume are expressed as five corpora of reactions. A user model exists as an MBTI personality profile or a manual proportioning of the contributions of each of the five interpretive dimensions. By applying the user model to the five corpora of reactions, we know the textual ingredients which are to constitute the aesthetic impression, and the proportions in which they will combine. To realize this stew of text as an aesthetic, the stew should be codified in some aesthetic code so that a viewer can have the experience of uncovering the artwork's meaning in an act of final resonance. 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.

Mapping the outputs of each Mode of Interpretation 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 eventuality that the human eye will blend these colors together, 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. Aesthetic thrives in spaces of connotation, and what the ambiguity of colors affords a viewer is the opportunity to discover a personal meaning in the colors.

If we mean colors to be the sole vehicles of aesthetic impression, then we must carefully control the form that the colors take. The form of a grid of squares is 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 as the subject of works of artists like Sophie Taeuber, Jean Arp, Piet Mondriaan, Paul Klee, and Ellsworth Kelly. There is also the idea that grids assist in the seduction of the viewer to experience and become affected by the artwork, because as Rosalind Krauss wrote, "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, p. 12).

2.5. *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. Whether or not a transaction will be efficacious is commensurate to asking to what degree a model of the artwork's message and a model of the viewer's tastes intersect. A narrative text, the subject of the artwork's aesthetic impression, is modelled as the sum of all the artistic ways in which it can be interpreted. Based on Jung's Modes of Interpretations theory, we give five dimensions of artistic interpretation: Thinking, Culturalizing, Seeing, Intuiting, and Feeling; we propose that computational readers implementing each of these interpretive modes be applied to the narrative text to produce textual interpretation. While in the present research we do not consider a viewer's unique experiences in creating his user model, the viewer is described by a pseudo-personal categorical model, representing his perceptual tendencies, and taking input from a Myers-Briggs Type Indicator (MBTI) personality inventory.

To produce an aesthetic impression which is best received by a viewer, we suggest that the artwork's aesthetic impression be adjusted to account for the viewer's perceptual tendencies and MBTI. The user model is used to determine the relative contributions of each of the five interpretations of the text to the final aesthetic impression. When this has been determined, the weighted sum of these interpretations is mapped from the textual domain into color space, and rendered in as a mythical color grid. The rationale for expressing the aesthetic impression through colors is that colors represent an aesthetic code which the user must break in the midst of ambiguity. By codifying the artwork's message through an aesthetic code, a viewer is invited to discover the truth of the artwork, and we theorize this as the efficacious *final resonance* of the aesthetic transaction, a moment of a-ha! We also theorize that by tweaking the constitution of the aesthetic impression to meld with the viewer's perceptual biases (e.g. Rational-Sensorial versus Intuitive-Sentimental), the viewer will feel greater rapport and intimacy with the structure of the message, since perceptual bias enhances the exclusivity of the artwork (insofar as personality types can be conceptualized as cultural clans), thus further enhancing the aesthetic transaction.

3 Aesthetiscope's implementation

In this chapter, 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*

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 (Fig. 3). The grid of color squares is 16 wide by 9 tall, flanked by black striping on top and bottom. There is a "glimmer" effect added to the colors in the grid, as their Values (i.e. Value, as in the Munsellian Hue-Value-Chroma system for colors) wax and wane according to various periodicities. Finally, the glimmering of the color grid refreshes at 24 frames per second, to complete the cinematic quality of the piece.

Fig 3. 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.

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 artwork generated consists of warm yellows, oranges, and reds. The artistic explanation mode flashes 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 of the generated artwork. 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 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 4.

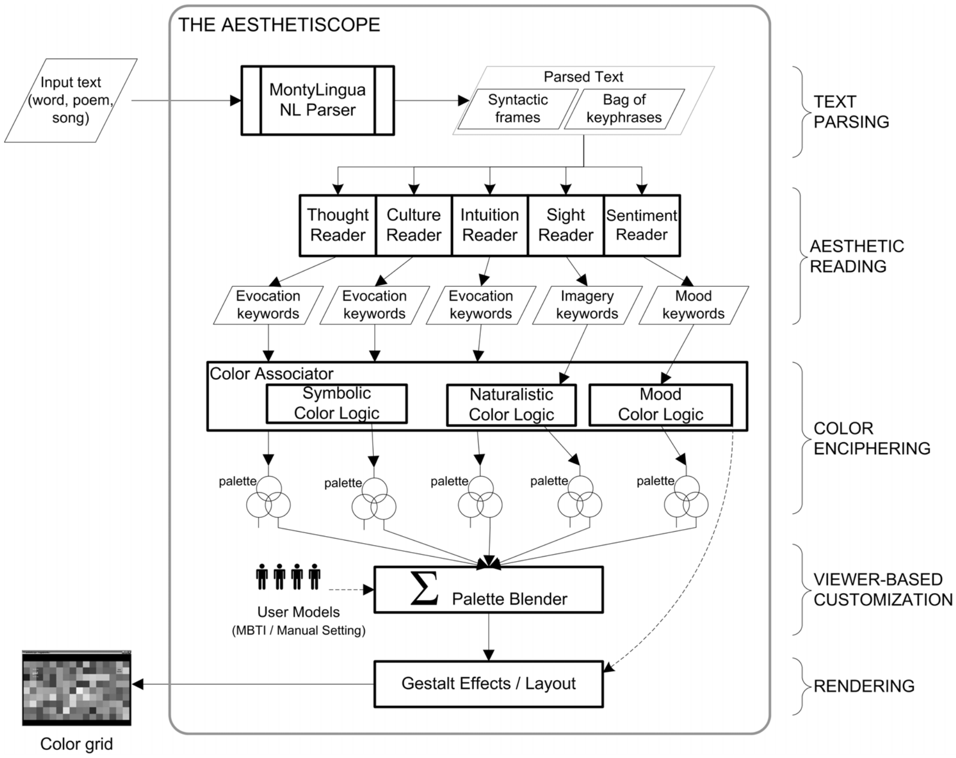

Fig 4. Input-Output process model of the Aesthetiscope implementation architecture.
NB: Detailed schematics of the five Readers are not shown here.

The implementation architecture can be viewed as taking five stages of processing, as shown in the rightmost-column in Figure 4. 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, 2003). 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, e.g. (this has been simplified):

“Some say the world will end in fire” ➔
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” ➔ “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 passed 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 ➔
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 facile. 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 interpretation 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 Aesthetiscope's processing are Color Enciphering, Viewer-Based Customization, and Rendering. Color Enciphering translates the evocation keywords outputs 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 2.1.1) suggestion that color space be viewed as an aesthetic code which invites a viewer to decipher it and uncover its underlying significance so that the final resonance is initiated by the viewer. 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 Aesthetiscope 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 2.3. 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 past a certain threshold (50%, in the current implementation) of all Reader contributions, then the mood keywords outputted by the SentimentReader will drive the gestalt operations on the final palette. 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 (golden ratio aspect) color grid and the artwork is complete!

4 The Aesthetic Readers

This chapter dives into the design decisions taken by, 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 ways of perceiving the world – by Thinking, Feeling, Intuiting, and Sensing – were a sufficient vocabulary to describe all the different ways that a person might interact with a 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 read a text and find inspirations for a color grid artwork like the Aesthetiscope. Harkening to an aforementioned caveat, 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 driving interpretation with archetypal common sense, or collective experience as a human and as a cultural participant (e.g. Aesthetiscope would express "dirt" as brown and yellow, recalling common sense, rather than idiosyncratic personal experience). The Exclusivity Principle (Section 2.1.2) 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 – the cachet that a viewer feels in receiving an artistic message meant just for him or only an exclusive few like him who are "in-the-know"; for instance an avant-garde receives the newest clothes hot from the fashion designer with greater aesthetic intensity than if the clothes were already known to many people. 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 computational 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 alternate ways we might have implemented these Readers. One common aspect of the five implemented Readers is that their mechanisms tend toward *associative* or *contextual* reasoning, which does not engage very cognitively deep reading; however, we feel that the nature of associations makes them very suitable for brainstorming the aesthetic potential of a text.

The remainder of this chapter 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 the 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 selected 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 100,000 common sense concept nodes (e.g. "lemon", "swim", "eat sandwich"), interconnected by 1.6 million semantic edges (e.g. "EffectOf("be hungry", "eat sandwich")"). Each edge represents 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 ideal as a 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 5 depicts the I/O process model for ThoughtReader's implementation using ConceptNet.

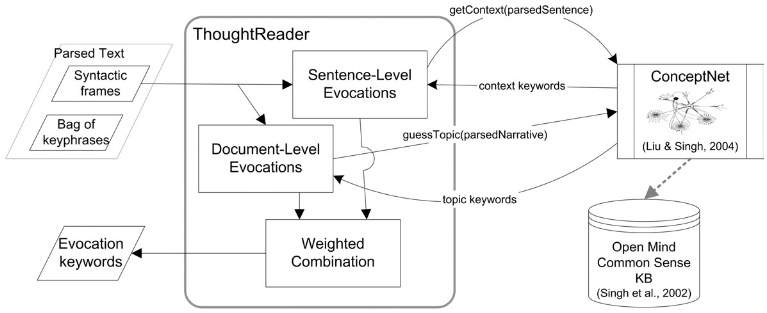

Fig 5. Input-Output process model of ThoughtReader.

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 which 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):

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 each sign correlates to some set of *signifieds*, and the nature of the correlations is dependent upon the value system of each culture. For example, the sign "*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 (in our case, toward *signs*) 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.

A cultural model, for WWTT, 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 directed toward 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 are not any off-the-shelf alternative systems specialized to the purpose of acquiring a cultural model automatically from a text corpus, other than the alternative of re-implementing something similar to WWTT from scratch.

Figure 6 depicts the I/O process model for CultureReader. We have been exploring the idea that in the future, the Aesthetiscope should be able to load the cultural models possessed by the viewer, dynamically. However, for Aesthetiscope's current implementation, we use only one cultural model, that for American popular culture, acquired automatically by WWTT from a 500kilobyte text corpus we compiled together, consisting broadly 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 garners the reactions to each sentence in the input text, and also reactions to the narrative as a whole. The reactions are then weighted and summed into a single bag of evocation keywords.

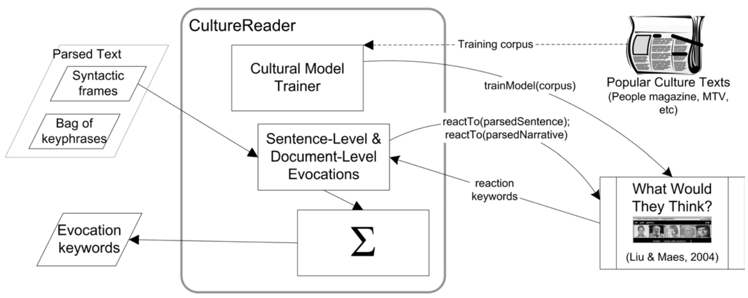

Fig 6. Input-Output process model of CultureReader.

It was necessary to modify WWTT in the following manner, to accommodate our required output format. WWTT normally reacts by emoting a numerical affect score obeying the third-dimensional PAD (pleasure-arousal-dominance) affect model of Albert Mehrabian (1995b). We modified WWTT so that in lieu of a score, WWTT would react by emoting 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. This modification is meant to accommodate the computational contract to output evocation keywords as a Reader’s interpretation.

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 are only exploring sight, and we are taking sight to be an ambassador for all the senses that Jung intended. We chose to deal with 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). Also, the choice is most facile because there exists large annotated corpora of photography and images in digital form, and this is to be a boon if we are to teach the computer to emulate the faculty of seeing.

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

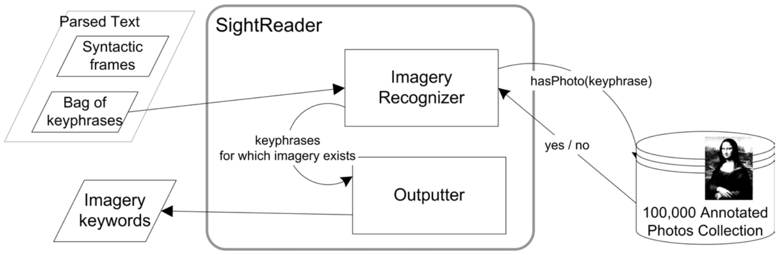

Fig 7. Input-Output process model of SightReader.

SightReader’s implementation is direct and lightweight (Figure 7). It utilizes only the bag-of-keyphrases parse of the input text. A recognizer filters out a subset of the keyphrases for which photos and hence color epitomes exist in the photo database. And 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, p. 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 have compiled together decades worth of research into a corpus of free association norms (1998). For example, in their corpus, the concept "traffic" triggers "car," "light," "jam," "sucks," "stop," "noise," etc. Of course, we should acknowledge that this measurement is specific to a certain population of people during a certain temporal period; nonetheless, we believe this corpus of free associations to be of high quality for the purposes of building an evocative reader which aims to respond "intuitively" to a text. Of course, we must give the caveat that IntuitionReader does not capture the whole spirit of intuition. For example, when we think of intuition, we think of it as a delicate and sensitive faculty. The intuitive consideration of a text should carefully account for all the subtleties of a text, and in general, an intuitive evocation for a narrative should be a convergent response to the *whole* of the narrative; however, this is outside of the scope of our present research capability, as it seems to demand full story understanding, which is an unsolved 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.*

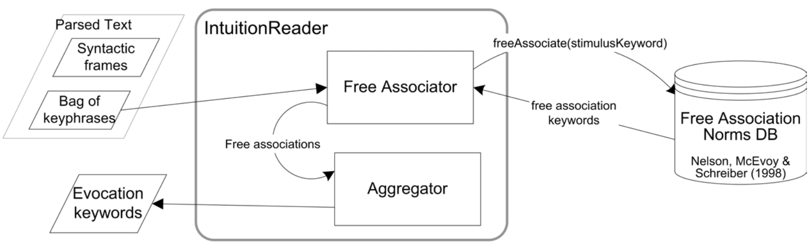

Fig 8. Input-Output process model of IntuitionReader.

Figure 8 depicts the process model of Intuition Reader. 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 which 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 which 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 model-based 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 arbitrary words, e.g. "accident" might be assigned a 75% affinity toward the fear emotion. Pennebaker, Francis, & Booth's Linguistic Inquiry and Word Count computer program (2001) is a good example of this approach, but as with other statistical language models, classification is only successful if the input text is of the same genre as the corpus used for training. Finally, knowledge-based approaches such as Liu, Lieberman & Selker's Emotus Ponens system (2003) 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.

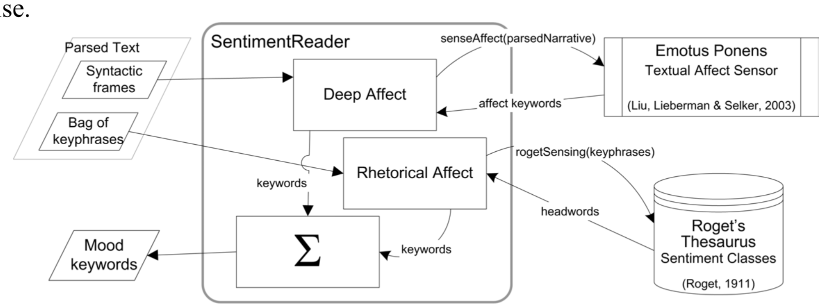

Fig 9. Input-Output process model of SentimentReader.

Figure 9 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 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 will 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 the next section, 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 chapter, we discuss how evocation keywords are mapped into color space. There are three calculi of color logic in the Aesthetiscope 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 render input text into color space in unique fashion, as illustrated in Figure 10. In Sections 5.2-5.3, we take on the task of describing how each calculus maps keywords into color space, and Section 5.4 describes how colors are blended together into a single palette. But before we exposit the technique of psych-semantic color rendering (it is more than just semantic because we are motivated to influence the psychological state of a person), we will briefly recapitulate, in Section 5.1, the context and motivation for our approach which we began in Section 2.4.

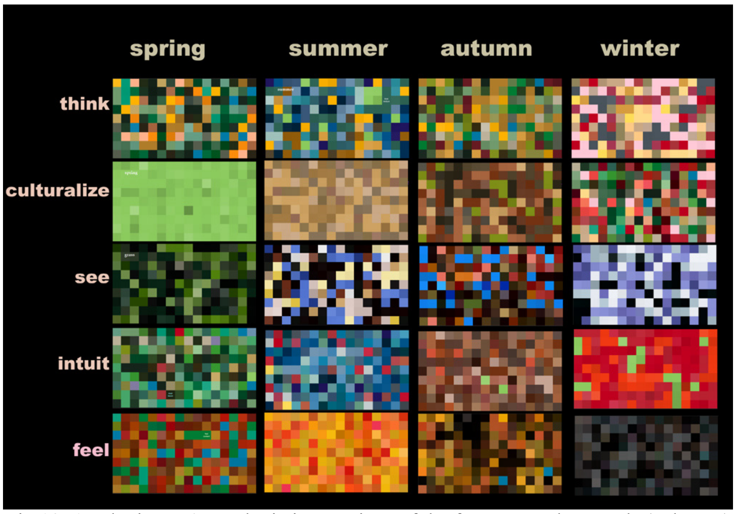

Fig 10. Aesthetiscope's aesthetic impressions of the four season keywords (columns) rendered through the monadic optics of one Reader taken at a time (rows).

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 2.1.1, we posited the Final Resonance Principle – the 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 Aesthetiscope will encipher evocation keywords into these colors, and invite, *as an aesthetic game*, viewers to perceive 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' and suggested, in discussing 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, p. 18) Even though Bell's sensibility of visual arts is that aesthetic is non-symbolic, 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 is a logical calculus which might most appeal to someone for whom the effect of visual memories is strong (perhaps a Jungian Sensing individual. We term this type of text-to-color mapping *naturalistic*. In Aesthetiscope's architecture (Figure 2.4), 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 amassed a large knowledge base 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 together 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 and 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). Recall that 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, and 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 writes, "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 the emergence of a global cultural bricolage. 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 are different between these cultures, 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, are 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: Eva Heller's *Wie Farben Wirken* (1989)*,* John Gage's *Color and Culture* (1993)*,* Johnannes Itten's *The Elements of Color* (1970)*,* and Josef Albers's *The Interaction of Color* (1963). These texts give explicit guidance for the emotional sign value of colors. A sampling of the guidance we used to construct our mapping (by hue) :

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 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 so they 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 4, 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 is greater than 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, or a bee, or a smiley face, or 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. The symbolic color palette is closer to kitsch than it is to subtlety. All colors are pure and stereotyped; these colors are 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," and "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 2.3 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 selected. As Figure 11 illustrates, biasing the Aesthetiscope toward certain readings can dramatically affect the final artwork.

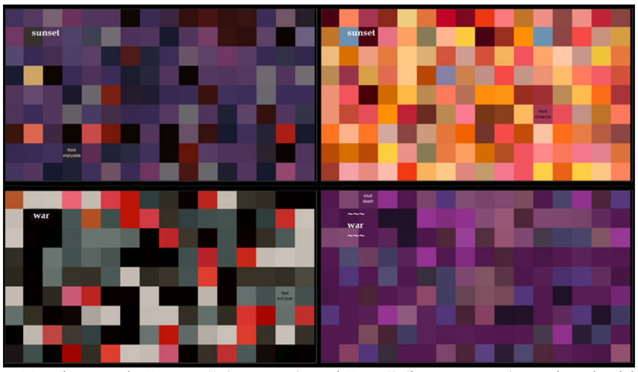

Fig 11. The words "sunset" (top-row) and "war" (bottom-row) rendered with a Thinking-Seeing bias (left-column) versus with an Intuiting-Feeling bias (right-column).

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 meant to reduce the brutal noisy appearance associated with uniform distributions.

6 Evaluation

Since our initial implementation and installation of the Aesthetiscope, we have heard many ideas from psychologists, designers, colorists, and hundreds of real people on how to improve the Aesthetiscope, and since then the piece has undergone a few iterations of redesign. In a companion paper (Liu & Maes, 2005), we reflect more upon these redesigns. We have also received a few suggestions on how best to evaluate the Aesthetiscope, as that seemed particularly problematic because aesthetic was such a subjective and relative matter. The visual artists we spoke with expressed doubt that such a thing as aesthetic efficacy could ever be proven in a controlled experiment, that it should only be studied ethnographically. One human-computer interaction specialist encouraged us to just issue a survey to see how people liked the Aesthetiscope 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 measured the signalling efficacy of each of the five reading dimensions. The second evaluation measured the aesthetic efficacy of a golden combination of the five reading dimensions which 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 Aesthetiscope 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") etc. chosen dispassionately by the examiner but with care to maintain diversity. Because some words were potentially unknown to Aesthetiscope, the examiner discarded 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 can 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. Results are shown in Table 1.

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

	Plausibility – 100 Poems/Songs					*Plausibility – 100 Evocative Words*				
	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

The results 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 offer them a choice. Taking the text from the 100 poems/songs, and 100 words, we overlayed 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 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. Twenty-six undergraduate students at MIT (perhaps in hindsight a skewed sampling for an evaluation on aesthetics) 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 79.2% across all judges; in the words category, the Golden Setting was identified as the artwork with an accuracy of 74.0% across all judges. Kappa statistics could not be calculated because each volunteer only 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 Conclusions

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 with the likes of artists like Ellsworth Kelly and herman de vries. We proposed a computationally minded theoretical framework for understanding the aesthetic quality of art as a type of transaction between an artwork's message and a model of the viewer, and we implemented an artbot called Aesthetiscope to realize and test this theory of art. 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.

Acknowledgements

First, we thank Walter Bender for inspiring the Aesthetiscope with his and Jon Orwant's incredibly mythical and addictive Color Deducto game, and for a thorough initiation into the color literature. We are also grateful to Prof. Judith Donath for inspiring a signalling approach to aesthetic, and Profs. Glorianna Davenport, Jeffrey Huang, and John Maeda for fruitful discussions and suggestions within the scope of this work. This research was supported by the sponsors of the MIT Media Laboratory, and in particular we thank British Telecom for their continued support for this line of research. Finally, we highly appreciate the anonymous reviewers of our FLAIRS paper for their helpful feedback.

References

1. Louise Rosenblatt: 1978, Efferent and Aesthetic Reading. *The Reader, The Text, The Poem: A Transactional Theory of the Literary Work.* Carbondale: Southern Illinois UP, 22-47.
2. Kenneth Moorman & Ashwin Ram: 1994, A function theory of creative reading. *The Psycgrad Journal.* Technical Report GIT-CC-94/01, Georgia Institute of Technology
3. C. G. Jung: 1921/1971, *Psychological Types*, trans. by H. G. Baynes, Princeton, NJ: Princeton University Press.
4. Kant, I. (1790/1929). Critique of judgement. In *T. M. Greene (Ed.), Kant selections* (pp. 375-432). New York: Scribner. (Original work published 1790)
5. Freud, S. (1919/1990), "The 'uncanny'", in *Dickson, A. (Ed.), The Penguin Freud Library, Vol. 14: Art and Literature*, Penguin, Harmondsworth, pp. 335-6 (first published in German in 1919).
6. Dewey, J. (1934) *Art as Experience*. New York: Perigee Books

7. G. W. F. Hegel, Aesthetics: *Lectures on Fine Art* (trans. T. M. Knox, 2 vols., 1975) (originally lectured 1835-1838)
8. Grant McCracken (1988). *Culture & Consumption*. Indiana University Press.
9. David Best, *Feeling and Reason in the Arts*. London: George Allen and Unwin 1985
10. Hugo Liu: 2004, Articulation, the Letter, and the Spirit in the Aesthetics of Narrative. *Proc. of the 2004 ACM Workshop on Story Representation, Mechanism, and Context*.
11. Roland Barthes: 1964/1968, *Elements of Semiology*. publ. Hill and Wang
12. Briggs, K. C., & Myers, I. B. (1976). *Myers-Briggs Type Indicator*. Palo Alto, CA: Consulting Psychologists Press
13. Heller, E. *Wie Farben Wirken: Farbpsychologie, Farbsymbolik, Kreative Farbgestaltung*, Rowohlt Verlag, Reibek bei Hamburg, 1989
14. John Gage. *Color and Culture: Practice and Meaning from Antiquity to Abstraction*, University of California Press, 1993
15. Itten, Johannes. *The Elements of Color*. New York: Van Nostrand Reinhold. 1970.
16. Joseph Albers: 1963, *Interaction of Color*. Yale UP.
17. J.W.v. Goethe: 1840/1970. *Theory of Colours*, trans. C. L. Eastlake. Cambridge, MA: MIT Press.
18. Brent Berlin and Paul Kay: 1969, *Basic Color Terms*. Berkeley and Los Angeles: University of California Press.
19. Rosalind Krauss: 1979/1985, *The Originality of the Avant-Garde and Other Modernist Myths*. Cambridge, MA: MIT Press.
20. Liu H: *MontyLingua: Commonsense-Enriched NLP, Toolkit and API* (2003) — http://web.media.mit.edu/hugo/montylingua/
21. Hugo Liu & Push Singh: 2004, ConceptNet: A Practical Commonsense Reasoning Toolkit. *BT Technology Journal 22(4).* pp. 211-226. Kluwer Academic Publishers.
22. Singh, P. et al. (2002). Open Mind Common Sense: Knowledge acquisition from the general public. In *Proceedings of ODBASE'02*. LNCS. Heidelberg: Springer-Verlag
23. Lenat, D. B. (1995). CYC: A large-scale investment in knowledge infrastructure. *Communications of the ACM, 38(11)*:33—38.
24. Mueller, E. (2000). *ThoughtTreasure: A natural language/commonsense platform.* Retrieved from http://www.signiform.com/tt/htm/overview.htm
25. Hugo Liu & Pattie Maes: 2004, What Would They Think? A Computational Model of Attitudes. *Proc. of the 2004 ACM Conference on Intelligent User Interfaces*, pp. 38-45
26. Collins, Allan M. and Loftus, Elisabeth F. (1975). "A spreading activation theory of semantic processing." *Psychological Review, 82(6)*, 407-428
27. Mehrabian, A. (1995b). Framework for a comprehensive description and measurement of emotional states. *Genetic, Social, and General Psychology Monographs, 121*, 339-361.
28. Henri Bergson (1946/2002). *The Creative Mind: An Introduction to Metaphysics*. (Trans: Mabelle L. Andison). New York: Citadel Press.
29. Mandler, G. (1980). Recognizing: The judgment of prior occurrence. *Psychological Review, 87*, 252-271.

30. D.L. Nelson, C.L. McEvoy & T.A. Schreiber: 1998, *The University of South Florida word association, rhyme, and word fragment norms*. http://www.usf.edu/FreeAssociation/
31. Elliott, C. *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.
32. S. Deerwester, S. Dumais, G. Furnas, T. Landauer, and R. Harshman, Indexing by latent semantic analysis, *Journal of the American Society for Information Science, 41* (1990), pp. 391-407
33. Pennebaker, J. W., Francis, M. E., & Booth, R. J. (2001). *Linguistic Inquiry and Word Count (LIWC): LIWC2001*. Mahwah, NJ: Erlbaum Publishers.
34. Hugo Liu, Henry Lieberman, Ted Selker: 2003, A Model of Textual Affect Sensing using Real-World Knowledge. *Proceedings of ACM IUI' 2003*. pp. 125-132.
35. Peter Roget: 1911, *Roget's Thesaurus of English Words and Phrases*. Retrieved from gutenberg.net/etext/10681
36. Bell, Clive (1914/1987). *Art*, Oxford: Oxford University Press.
37. N.Jojic, B.Frey & A.Kannan, "Epitomic Analysis of Appearance and Shape", *Proceedings of ICCV 2003*.
38. Sidlauskas, Susan. (2004). Emotion, Color, Cezanne (The Portrait of Hortense), *Nineteenth-Century Art Worldwide: 3(2)*. AHNCA Press.
39. A.H. Munsell: 1905, *A Color Notation*, Boston.
40. Hugo Liu & Pattie Maes (2005). The Aesthetiscope: Visualizing Aesthetic Readings of Text in Color Space. *Proceedings of FLAIRS2005*, pp. 74-79. AAAI Press.