

Compositionally-Descriptive Engines: Exploratory Notes

*Antony Smith, MA
Independent Scholar
Cleveland, England, U.K.*

Abstract

Elements of composition in a photographic image cause physical responses in the viewer; the rate of eye movement is directly determined by the composition, as can be the attitude of the viewer's head and shoulders. Such physical effects trigger specific memories in the viewer, and those memories affect the emotional state of the viewer. The act of looking at pictures is much more complex than this, but never-the-less there is a direct relationship between composition and emotive response. This article will explore those relationships and from the findings will suggest potential applications.

Introduction

These notes will briefly introduce the basic elements of pictorial composition as the means by which a set of predictable measures can be generated for any picture. The purpose of generating predictable measures being the definition of a none (or less) subjective categorisation of pictures; in order to improve their cataloguing and hence search-ability.

In the current climate of cheap personal computers (PCs), digital cameras and compact photo disks (CDs) writers the number of images requiring cataloguing is growing at a significant rate. Even with the organisational benefits of a hierarchical filing system and the convenience of a picture-browser, finding a suitable image can be time consuming – even for the home-user. This is compounded by the ability to archive images to CD, where browsing requires continual disk changes. On-line thumbnail (contact) picture-

indexes can help, but eventually they too need archiving to off-line store; whereby the search mechanism takes on an additional level of abstraction.

Furthermore, a browser-based search method relies on interpreting each browsed image on the fly against the search criteria, which can be an extremely high-order process. The selection of images is typically the remit of the artist or designer, as it is they who must make the subjective interpretation of the options. There is significant benefit then in automating the search process to a degree that at least renders a manageable sub-set of possibilities.

Typically a picture catalogue will contain simple EXIF data and subjective short textual tags. There is no data-oriented view of a picture's content or intent. Nor is there any consideration to the picture's inherent colour-mask. Such data-views provide the fundamental for a highly targeted search engine. But how can a data-view be defined for 'pictorial content and intent', and how may we make a meaningful selection from the data-view of a colour mask?

These notes develop a data-view of measurable compositional objects. The premise being that composition is inherently measurable and that it is also a statement of pictorial content and intent. The former is evident given that it deals with measurable things such as lines and shapes. The latter follows firstly from the fact that definition of compositional objects **is** the pictorial content-map, we simply lack any contextual word-tags. That composition is a statement of pictorial intent follows from the physical, and consequent emotional, effects upon the viewer caused by the compositional-form – if the eye tires, the brain follows quickly. Having elicited such measurable data-views from a consideration of the emotional impact of compositional observations, these notes conclude with a description of a possible application area: The pictorial-intent search-engine.

Elements of Composition

Mass and Mass-Balance

Definition: Mass-Objects

Mass refers to the objects of a composition that are considered to have sufficient self-definition to be classed as non-background. Each leaf on a branch in a tightly cropped close-up is an object of mass. The branch is not because:

- Its geometric shape in contrast to that of the leaves makes the branch a line. By definition lines do not have mass
- The colour of the branch will generally be closer to that of the background than that of the leaves in the colour-gamut. In which case the branch is considered to be visually subordinate to the leaves.

Of course, the photograph can be arranged so that the branch is placed against the sky, in which case a tight-crop may give the branch mass.

Measure: Mass

Digital image representations are not strictly two-dimensional since they exhibit not only horizontal and vertical extent but also colour depth. The “mass” of an image object can therefore be determined from:

$$\text{Area} * \text{Colour Depth}$$

However, depth should be measured as a relative separation from the background “average” depth, not from an absolute arbitrary scale.



Figure 1a. Here the yellow highlights indicate two mass-objects. The mass concentrates around the features; this impacts consideration of the relative placements of the objects.



Figure 1b. Here, by considering the enclosed areas of the highlights at the rear of the picture we discover three further mass-objects, as shown by the orange highlights.

Impacts: Mass

Mass per se does not impact emotive response. The relative placement of mass, however, does. Having identified mass-objects, it is then necessary to consider the balance of those objects.

Definition: Mass-Balance

The purpose of a mass-balance view of a scene is to concentrate on the relative “sizes” and positions of the objects of mass, in terms of balance. As if those objects would have a physical gravitational-effect upon the picture. A megalith along the left-hand cardinal points (see later) in a wide landscape shot would have enough compositional weight to encourage the viewer to look at the picture with a decided-tilt to their heads – which may be desired. If not a tree, or a cat, or a bulldozer, on the lower right-hand cardinal point would have some mass with which to offset the effect of the megalith.

Measure: Mass-Balance

The main subjects align along a near *strong diagonal* (Figure 2a). The right-hand subject exhibits “dominance.” Looking at the photograph in Figure 2b, it can be seen that the right hand-subject covers a much greater area and so has more mass concentrated on the top-right cardinal point than the left-hand subject is able to exert at the bottom-left cardinal point.

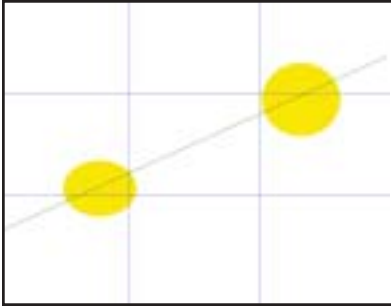


Figure 2a. Near strong diagonal.



Figure 2b. Dominant right.

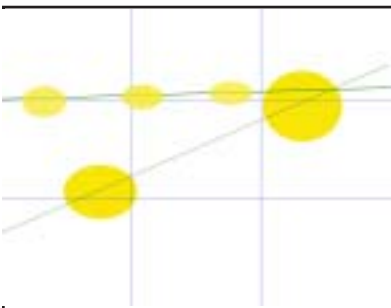


Figure 2c. Lesser objects.

The third image (Figure 2c) shows the positioning of lesser (rear) objects. The three objects balance each other perfectly around the top-left cardinal point. They exhibit harmony, unity. The line they occupy also extends directly across the eyes of the right-hand subject – and therefore is implicated in the unity of the back-line. The picture thus exhibits dominance, harmony, and unity.

Further it can be seen that left-hand subject is isolated from the other objects, whereas the right-hand subject abuts one of the rear objects. As well as exhibiting unity, the picture exhibits self. The mass-balance view thus describes the pictorial intent in terms of: Dominance, harmony, unity, self.

Impacts: Mass-Balance

Detecting mass objects is relatively straight forward. The relationships between those objects can be defined in terms of relative mass, position within the frame relative to other objects, position in the colour-gamut of the picture. Those relations can be described in terms of “emotive intent:” a group of similarly sized and coloured objects balanced around a cardinal point can be said to exhibit unity, and so on as stated above.

Cardinal Points, Lines, and Line-Objects

Definition: Cardinal Points

The four cardinal points are positioned where the horizontals and verticals intersect at the thirds of the picture area.

These four points do not have equal compositional strength. The viewer may come from a culture where printed materials are read left to right. The eye of that viewer will read the picture in a similar left-right manner. It's also common for a viewer to read a picture from bottom-left as opposed to the top-left, but this can be impacted by the relative heights of the picture and the viewer's eyes. The lower left cardinal point therefore has dominance.



Figure 3a. Notice in the original composition the main subjects are set in opposition across the strong diagonal, as discussed, providing a sense of dominance.



Figure 3b. Here, the main subjects are more equally balanced. The greater mass of the right-hand subject is centred around a weaker cardinal point, which reduces that subject's dominance.

Measure: Cardinal Points

The cardinal points may be ascribed a “power” factor, mass-objects falling within the circle of influence of a cardinal point effectively increase in mass in relation to that point's power factor. The following formula puts this in perspective:

Mass =

$$(Area * Relative Colour Depth) * (Nearest Cardinal Point Power/Distance to Nearest Cardinal Point)$$

Impacts: Cardinal Points

Since the point of creating a picture is to have it viewed, the purpose of composition is to draw-in the eye. In a well-composed scene the eye will be

drawn firstly towards one of the other cardinal points. Most typically this will be the diagonally opposing upper-right cardinal point. This diagonal opposition renders the longest path through the image. The lower-left to upper-right path is termed the “*Strong Diagonal.*”

The upper-left to lower-right (the “*Complementary Diagonal*”) path is also quite strong, its greatest strength being its ability to complement the strong diagonal.

Definition: Lines

A line is similar to a mass-object in terms of detection, but the length-width ratio is such that the object is considered to be without mass. Further subtlety is achieved by placing mass-objects in a manner that creates implied lines. Beauty arises out of subtlety – the attendant feeling of familiar comfort from an unrecognised source.

Measure: Lines

Lines, real or implied, also have measurable attributes, which help to define their impact. Line impacts are many and varied; the following pictures consider “radial patterns” of lines.

A measure of “vastness” can be taken from the distance from the convergence point to the image edge.

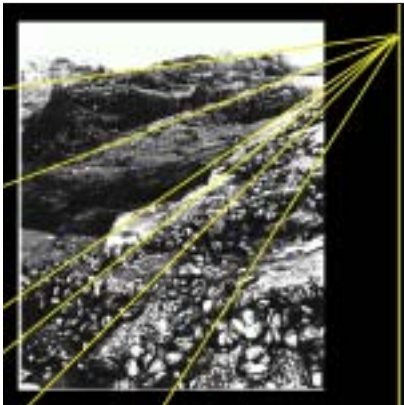


Figure 4a. Here folds and fissures in the landscape prescribe a series of lines that all converge on the same point. This point is out-of-frame which increases the sense that the picture presents a window on a greater whole; a sense of “vastness.”

A measure of expansion can be taken from the average angle of separation between the converging lines – for example, the greater that angle the “faster” the perceived expansion.



Figure 4b. Here many lines radiate from a cardinal point. This pattern gives the picture a sense of “expansion.”

Impacts: Lines

The above has defined measurable terms for the concepts of vastness and expansion. Lines can also express stability (in the horizontal), impregnability (in the vertical), universality, constancy (circles), beauty (S-curves), and many other interpretations.

Figure 5a. Here line objects create a cross shape at the lower-right cardinal point. Because crosses are highly recognizable the two lines forming the cross take on enough weight to balance the heavy treetop on the opposing cardinal point.



Definition: Line-Objects

Lines may occur in a pattern that implies a higher-order shape, such as a triangle, cross, pyramid, arc, or s-curve. Each such shape can also be ascribed descriptive words; for example a pyramid may be seen as strong, stable, permanent, etcetera. All compositional objects are carriers of pictorial intent, for they contribute to the overall pictorial “message” or “theme.” This should be particularly evident when considering the emotive impact of circle and cross shapes.



Figure 5b. Here the arrangement of the flower heads suggests a pentagon. The almost solid pentagon shape prescribed by the five-leaf arrangement below reinforces this suggestion.

Measure: Line-Objects

Each potential shape is prescribed an emotive property (pyramid = stability). The area of the line-object defines the “quantity” of that property within the picture.

To normalise results in comparing pictures “area” is always taken to be “percent of image area.”

Impacts: Line-Objects

Line-objects are inferred from the positioning of mass-objects or lines, they therefore tend to cover large areas. They do not impact other compositional objects, they add depth and resonance.

Colours

Since digital images are represented as colour maps detecting and measuring colours is a straightforward matter; there are some degrees of freedom in deciding exactly where in the spectrum one colour finishes and another starts.

Colours have a discernable emotive impact, both alone and in combination. Colours may be considered to exhibit strength, warmth versus weakness, coldness; colours may be advancing or retreating.

Figure 6. This market stall in Pashupatinath, Nepal, clearly indicates the cultural importance of colour; “standard” designations for the meaning and impact of colour have been derived over centuries.



Colours of a given strength (saturation) can be brightened (e.g. strengthened) when placed in contrast with white; all colours have a similar modifying effect on each other. For instance contrast can be modeled by increasing the strength and weakness values of colours placed in contrast. Colour harmony effects can also be modelled so; two harmonising colours (blue/orange say) can boost each other’s calculated strength.

Having defined the strength of a colour, it is possible to define the attribute that specific strength represents. The Chandamaharosana Tantra² states:

Black symbolizes killing and anger

White denotes rest and thinking

Yellow stands for restraining and nourishing

Red for subjugation and summoning and

Green means exorcism

Black: Signifies hate, ignorance – and inversely clarity, truth. Darkness represents the absolute, the threshold experience.

White: All colours are present in white; nothing is hidden, secret or undifferentiated. White stands for learning, knowledge, purity, holiness and cleanliness.

Yellow: Is WARM ADVANCING and signifies daylight. Humility and renunciation, desirelessness, the colour of earth, symbol of rootedness.

Red: Is WARM and signifies life and threat to life, passion, the sacred, longevity.

Green: Is in the middle of the spectrum - qualities of balance, harmony. Nature, pacifying, youthful vigour, action.

Blue: Signifies eternity, truth, devotion, faith, purity, chastity, peace, spiritual and intellectual life.



Figure 7a. No colour shift.



Figure 7b. Colour shift.

In comparing the pictures in Figures 7a and 7b, it is evident that the colour-shift in the second has had the effect of increasing the left-hand subject's mass. This is because the left-hand subject is further away from the picture's average colour temperature than the right-hand subject.

The angle between the subjects seems less acute; the eye stays longer on the bottom-left cardinal point. The compositional sense has shifted from "self in community" to "isolated within community."

The histograms in Figures 7c (representations from Adobe Photoshop™) give an indication as to how colour data can be used to compare picture content; they indicate how much of the picture (or object) area falls into the designated "blue" band of the spectrum.

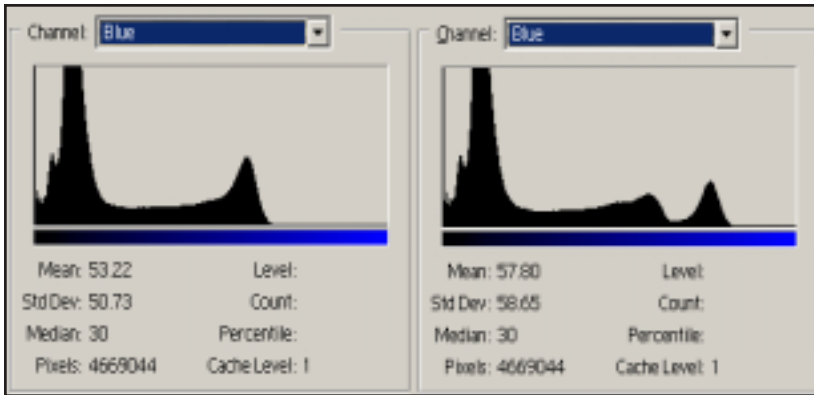


Figure 7c. Set of histograms comparing how much of an image falls into the blue band-width area.

Summary of Composition Elements

The foregoing is a brief introduction to the components of an image which are measurable and which are carriers of meaning or intent. All these findings are presented here in summary and it can be seen that the identified components are capable of differentiating around forty-five “expressions of intent.”

These expressions are somewhat arbitrary, however they illustrate the point well significant volumes of research are available to aid the definition of an accurate vocabulary. The following points are worthy of note:

- It is necessary that similar expressions repeat throughout, as this ensures that a single factor does not unreasonably bias any selection method (i.e., it is undesirable to always select largely red images given a criteria of “longevity”).
- Expressions relating to colours are much fuller than those relating to compositional objects – the colour definitions reveal the extent of the scheme and it is only a matter of research and/or experimentation to fill out all expressions similarly. (Consider the FIORES II¹ project where the actual daily vocabulary was measured against “design intent” so that mathematical models of that intent could be constructed).

Compositional Objects/Functions	Express Intent
Equal Mass-Objects, Line-Balanced	Unity
Equal Mass-Objects, Point-Balanced	Community
Equal Mass-Objects, Unbalanced	Dominance
Unequal Mass-Objects, Line-Balanced	Leadership
Unequal Mass-Objects, Point-Balanced	Gathering
Unequal Mass-Objects, Unbalanced	Crowding
Radial Convergence (convergence point is in frame)	Expansion
Radial Divergence (convergence point is outside of frame)	Vastness
Line, Horizontal	Stability
Line, Vertical	Impregnability
Line, Diagonal	Acceleration / Deceleration
Curves of Varying Orders	Beauty
Circles	Wholeness
Line-Objects, Cross	Spirituality / Prohibition (dependant upon orientation)
Line-Objects, Pyramid	Stability
Line-Objects, Pentagonal	Homeliness
Colour Temperature	Hot, Warm, Cold, etc.
Colour Contrast	Contrast
Colour Harmony	Harmony
Blacks	Hate, Ignorance
Whites	Learning, Knowledge, Purity, Holiness and Cleanliness.
Yellows	Humility and Renunciation, Desirelessness, Rootedness
Reds	Life and Threat to Life, Passion, the sacred, longevity
Greens	Balance, Harmony. Nature, Pacifying, Youthful Vigour, Action
Blues	Eternity, Truth, Devotion, Faith, purity, chastity, peace, spiritual and intellectual life

The vocabulary presented above is strictly limited; many more “expressions of intent” are required in order to effectively select a given image from a large set. It will be necessary to prescribe a mathematical model which is capable of determining how many unique measurable objects must be supported in order to effectively make a selection from a given picture-store size. Also in the foregoing there has been no consideration of “texture effects,” which need to be incorporated into the scheme as a whole.

Further, many of the intent descriptions in the foregoing relate to single objects or objects in relation to the whole image. Much more consideration is required of phenomenon surrounding:

- Inter-object effects (as per colour harmony and contrast discussions in fact)
- Object-grouping effects – pictorial object repetitions give rise to pictorial-intent resonance.

And it therefore follows that it is necessary to consider the impacts of relationships between different objects and groups of objects.

Applications of Composition

Pictorial Composition As A Search Engine

For a picture search, the search-criteria may be expressed in a variety of abstraction-levels, for instance consider the following:

- The picture is a known frame number from a known film roll – such information is commonly held with digital pictures in formats such as EXIF.
- The picture should have a given tonal range – which may be required in a pure graphic design application such as brochure or web site layout. In theory the embedded ICC-colour profile could serve as a search mask.
- The picture should contain a recognisable component – a boat, a dog, a space station. Such descriptive terms can become limited in their ability to describe images, particularly if an image is of an “abstract” nature. The prescribed terms can also be subjective. The searcher may not make the same subjective decisions as the cataloguer. The search results will omit images that are entirely suitable as well as include a selection of those that are not.
- The picture should complement a given emotional state, we wish to search upon the pictorial intent – this would be the most “natural” way to search for an image from the perspective of an artistic application.

Searching by descriptive words is subjective and renders incomplete results. The problems lie in differences between the cataloguer and the searcher's:

- Word selection, and
- Way in which those words are associated with aspects of a given image. For example to be catalogued as an image containing a dog, should that imply one and only one dog? Or does it include pictures with dogs in the background and so on?

Differences in word selection can be controlled by only allowing search words that it is known appear in the index. When the index becomes large it can still be difficult for the searcher and cataloguer's word selections to match. This is because any word-list will be alphabetically sorted and so words of similar meaning do not necessarily occur in close proximity. An answer to that is to categorize the word-list, by emotional state or "ambience," and to provide alphabetised word-lists within each category. Categories could be anything, such as: family, danger, refuge, love, fear, spiritual, etc.

Control of the word-lists provides a structuring framework to a subjective search engine. The key to targeted searching is the minimisation of the effects of subjective categorisation. The best way to minimise such effects in any search engine is to normalise the information upon which the search is based. Normalisation here means to align the searcher and cataloguer's subjective interpretations. Normalisation is achieved by prescribing a predictive method to the application of a relationship. As an example, the normalisation method could be:

- Apply a measurement technique,
- Perform a mathematical transformation to yield an ordinal set,
- Traverse an agreed ordinal-place to descriptive-word map.

In this scenario search results are predictable. Which will allow the searcher to "tune-in" to the system's operation. This is never possible when cataloguing is performed by many individuals over long periods of time.

The earlier part of this article has discussed the physical and emotional impacts of compositional factors on the viewer. It is partly through these impacts that the pictorial intent is communicated. A prescribed set of compositional factors therefore, is an ideal starting point for a normalised

search engine measurement technique, particularly since compositional factors deal explicitly with measurable features as follow:

- The mass-balance can be profiled thus:
 - o Break the image into cells.
 - o Assign each cell a mass-depth (gravitational strength) based upon percentage occupation by a mass-object multiplied by the object's mass-factor (specific density). The mass-factor is determined by the occupying mass object's strength – determined by the relative distance in the colour gamut to the background's average position.
 - o Transform the two-dimensional cell grid into a one-dimensional sequence.
 - o Plot the mass depth against the cell sequence
 - o Take the resulting curve's differential
- The mass-balance differential is assigned an ordinal position within the set of all mass-balance differences within the search set.
- Each position within the ordinal set of mass-balance differentials is ascribed a single descriptive word. Possibly from Calm through Turbulent, since the differential describes the 'rate of change of balance'.

A similar process can be defined for all compositional factors described earlier. Furthermore, the identification of compositional objects (mass-objects, line-objects) lends itself to automatic detection. Edge-detection is commonplace in image editing software. Colour information can be easily transformed mathematically. Angles and cardinal point offsets can be equally easily handled.

Because the above search strategy is predicated on the control of subjective effects, its efficacy can only be known by a process of live-trials. Appropriate trial dataset and prototype software will be required. However, note that such a cataloguing/searching technique avoids the dangers apparent in "metadata abuse" – the inappropriate assignment of text tags to catalogued objects in order to bias the "hit rate." The prescribed Compositionally-Descriptive Search Engine avoids this through automatic analysis.

References

- 1 <http://rkk.mv.uni-kl.de/FIORES/FIORES2/publication/CADToolsBasedOnAestheticProperties.pdf>
- 2 <http://www.ethangka.com/article/colors>