Visual Complexity: Is That All There Is?

Alexandra Forsythe

Liverpool John Moores University, UK
{a.m.forsythe}@ljmu.ac.uk

Abstract. Visual complexity is conventionally defined as the level of detail or intricacy contained within an image. This paper evaluates different measures of complexity and the extent to which they may be compromised by a familiarity bias. It considers the implications with reference to measures of visual complexity based on users' subjective judgments and explores other metrics which may provide a better basis for evaluating visual complexity in icons and displays. The interaction between shading and complexity is considered as a future direction for the empirical study of visual complexity.

Keywords: Icons, Visual complexity, Familiarity, Metrics.

1 Complexity in Icon and Symbol Research

Visual complexity is a concept introduced by Snodgrass and Vanderwart [1] to refer to the amount of detail or intricacy in a picture. This concept has now been adopted in icon research and is frequently measured with reference to the number of lines within an icon or symbol [2, 3]. The amount of detail or intricacy within an icon influences the rate at which an icon is detected. Very simple or very abstract icons are detected faster than those in the mid-range. Some studies [4] report a negative effect on response latency have used abstract and concrete stimuli - a mixture of symbols and icons - but many have based their findings on arbitrary stimuli such as symbols [5, 6] or lattices of random black and white quadrangles [7].

Images that are more concrete or real world do not produce the same increase in response latency [8]. One explanation is that arbitrary stimuli are probably more semantically impoverished and are also less likely to have any predictable pattern. Only small pieces of information can be processed at any one time and the visual system is possibly unable to draw any semantic inferences to the same extent as can be achieved with pictures [7]. Thus, increases in response latency and response errors are perhaps more likely to occur.

Whilst we know that complexity is possibly related to response efficacy, there is still little consensus as to what complexity is or how it should be defined and measured. For example, Feldman's definition of an 'absence of pattern' reflects an emphasis on randomness, impoverishments and a degree of perceptual difficulty [10]. In contrast, Garcia et al. [11] relates complexity with increasing real worldness. Whilst the latter is perhaps an oversimplification, it does allude to principles of higher organization and the human propensity to search for pattern, an effort after meaning not lack of it [11]. Different approaches to measuring complexity will now be reviewed.

D. Harris (Ed.): Engin. Psychol. and Cog. Ergonomics, HCII 2009, LNAI 5639, pp. 158–166, 2009.
© Springer-Verlag Berlin Heidelberg 2009

2 Empirical Complexity

2.1 Early Approaches to Visual Complexity

The study of visual complexity emerged from the empiricist tradition. The tradition is based on the premis that people make poor intuitive judges in uncontrolled settings; understanding could only be advanced through quantification in controlled laboratory settings. When unusual, unexplainable results emerged, Gestalt psychology developed to explain them. The Gestaltists set out to understand the processes of perception, not through the meticulous analysis of patches of light, shape and colour, but through an analysis of the whole, configuration or form [12]. Their philosophy was that sensations are not elementary experiences; we "see" shape and form regardless of where the image falls on the retina or what neurons process the various image components. What was important was constancy.

One such law generated through the Gestalt movement was prägnanz. The Prägnanz principle contends that the forms that are actually experienced take on the most parsimonious or 'best' arrangement possible in given circumstances. In other words, of all the possible perceptual experiences to which a particular stimulus could give rise, the one most closely fitting to the concept of 'good' will be experienced. The term 'good' means symmetrical, simple, organised and regular [13]. This study of psychological organisation explained the tendency to create psychologically, simple order patterns from a wide range of perceptual stimuli.

This early study of 'simplicity' evolved into the study of 'complexity', with theorists attempting to re-write the Gestalt Law of simplicity within a more formal framework [14, 15, 16]. Both Hochberg and Attneave acknowledged that shape was a multidimensional variable that would vary with the complexity of an image.

Attneave [14] also developed a measure whereby an image could be measured in 'dots'. An outline image was presented to observers who were then requested to place dots in important image areas (bends and curves). These dots would be used to reproduce the image as accurately as possible. Simple images required fewer dots. 36 dots were used to recode the changes in contour for a picture of a sleeping cat. Attneave was able to produce an abstract of the image by connecting these points with a straight edge (Figure 1).

Hochberg and Brooks [14] developed a semi-automated measure of image complexity. They argued that relying solely on human judgments would mean that there would be no way of predicting how complex a novel image might be judged. Hochberg's calculations demonstrated that it was possible to predict how viewers would 'see' an image; the more interior angles, different angles and the more lines in an image the more likely it would be perceived in three-dimensions (Figure 2). The number of interior angles, the average number of different angles and the average number of continuous lines can be combined to provide a measure of complexity.

However, knowing how many dimensions were needed to explain a shape was not sufficient to judge its complexity, since some dimensions (e.g. reference-axis or spaces) were more meaningful than others [15]. In other words, the calculation of a metric based on increasing tri-dimensionality tells us very little about either the complexity of unfamiliar images or the learning processes that can influence the perception of form. Attneave & Arnoult wanted to understand the degree to which

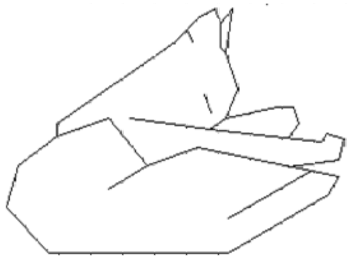


Fig. 1. Attneave's cat

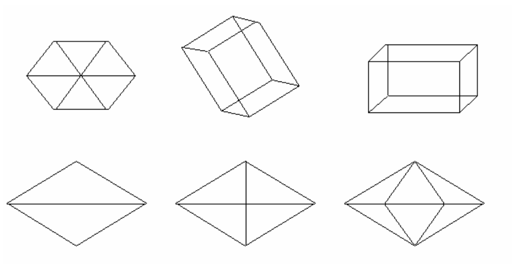
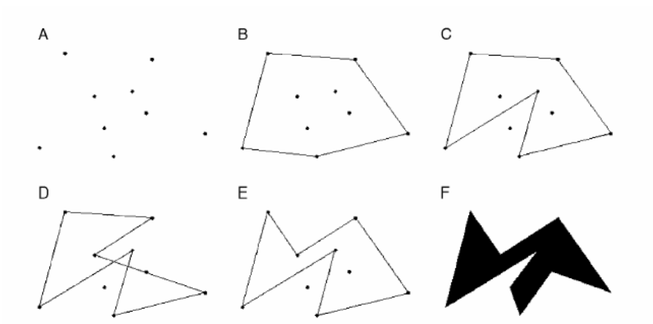


Fig. 2. Increasing tri-dimensionality



(A) Using a table or random numbers place a set of scatter points on a piece of 100x100 graph paper. The number of points will correspond to the number of sides the shape will have.
(B) Connect the peripheral points to form a convex polygon; some concavity will be tolerated.
(C) The smallest subset of points with convex angels are then connected by drawing the line towards a point on the opposite angle.
(D) Lines should not cross one another
(E) All points must be connected
(F) Nonsense shape

Fig. 3. Creating random polygons (Attneave & Arnoult, 1956)

size, contrast, method and familiarisation influenced the perception of form. They developed a system of calculations that could be used to generate nonsense shapes, the idea being that if testing using such a metric worked for images that had no meaningful relationship with 'real world' counterparts then it could be generalized to other stimuli. This system is outlined in Figure 3.

2.2 Later Approaches to Visual Complexity

Complexity has received less attention in recent years, in part because of the absence of a universally acceptable metric [17] and those measures that have been developed are not particularly well supported within a theoretical framework. For example, Geiselman et al [18] developed an index of discriminability between graphic symbols and identified nine 'primitive' attributes; e.g. numbers of straight lines, arcs, quasi angles and blackened-in elements. Symbols selected for high discriminability using this metric were responded to faster than those with lower discriminability. Similarly, Garcia, Badre & Stasko [10] developed a complexity metric based on a calculation of several icon features including the number of closed and open figures, and horizontal and vertical lines. This metric was developed primarily as a measure of the concreteness of icon. Garcia et al. reported that icons that are pictorially similar to their real world counterparts are more likely to be judged as complex. This has been found not to be the case icon: complexity is more closely related to search efficacy [1, 4].

A more valid and reliable measure of complexity would enable researchers to determine more accurately the effects of extra detail and intricacy on performance. Forsythe et al. [3] tested several automated measures of icon complexity based on measurements of the changes in image intensity. These measures were informed by arguments that course and fine lines are critical in providing information about a stimulus. The brain registers variations in an image as changes in intensity, and it is these coarse and fine changes that provide detail and local information about a stimulus [19, 20, 21, 22]. Coarse scales are thought to be treated by the brain as low-frequency components obtained from local information. This difference in processing speed would seem to be a function of image complexity: When an object is of a detailed nature, its global attributes are processed much faster than its local ones [23, 24].

Forsythe et al. [3] showed that these basic perceptual components [i.e. edges] are important in the measurement of complexity in so far as the extent to which an image is measured as having edges correlated highly with subjective judgments of image complexity. For example, the 'Perimeter' detection metric correlated (r_s=.64, p<.001) with a random set (n=68) of the McDougall et al. [2] icons and symbols and also correlated (r_s=.66, p<.001) with measures by Garcia et al [10]. This Perimeter metric has reasonably good predictive validity when applied to other pictorial images [25]. Perimeter measures are described by [26] as contour-based, global measures of shape. Perimeter measures do not divide the shape into parts; rather the whole shape contour is used to describe the shape. This makes this type of measure very straightforward for users to implement and as such it tends to be a popular method of image measurement.

An alternative automated measure of complexity that is also very straightforward to implement is based on the size of the compressed image file [27, 28] Image compression techniques take advantage of the fact that many images contain a lot of repetition. This information is removed or reduced to enable storage of the image in a compact form (it takes up less disc space). A more complex picture will have more image elements and these elements will be less predictable (there will be less repetition). The file string (an ordered sequence of storing variables) will be longer and contain an increasing number of different sequences.

Donderi [29] revisited information-theory as a possible framework that could explain the success of image compression techniques (such as Jpeg) as a determinant of complexity. Information-theory treats a message as a series of components to be communicated and the components in a complex image are its primitives. Donderi argues that when a picture is compressed the string of numbers that represent the organisation of that picture is a measure of its information content. When the image contains few elements or is more homogenous in design, there are few message alternatives and as such the file string contains mostly numbers to be repeated. This is consistent with a basic premise of information theory: the information content in a message is inversely proportional to its probability of occurrence.

Forsythe et al [30] applied this theory in the evaluation of several compression measurement techniques. Compression scores were collected for the image sizes of several published image sets. Gif compression provided a good approximation of human judgments of visual complexity across three image sets[1].

This study also demonstrated that many complexity metrics – based only on human judgments- were biased by a familiarity effect; unfamiliar images were rated as more complex than they actually were using the automated metrics (see Forsythe et al. for a list of studies). Forsythe et al [30] demonstrated this effect by training participants on a group of nonsense shapes. When subsequent ratings for complexity were collected from this group they rated the shapes used in training as simpler than a group of naïve participants. These results suggest that humans are not best placed to make judgments relating to the complexity or simplicity of an image. Compression techniques are a fast and user friendly option for the measurement of visual complexity and they are not so affected by judgments of familiarity. These metrics have some underlying theoretical bases, such as information theory, and produce good approximations of human judgments.

3 Nativist Complexity

The following section examines how our perception of visual complexity is overlaid by other factors. On this basis, I suggest that visual complexity should be considered in relation to other factors rather than alone.

3.1 Familiarity

Recent work has moved some way closer to developing a theoretically informed measure of visual complexity, *but is that all there is?* The finding that familiarity is

---

[1] Jpeg is a technique that reduces the size of the image file by removing redundant information, but generally assumes that some loss of information is acceptable. Gif works on a similar principle except that when the image is to be recovered no image loss occurs.

related to complexity resurrects and an old argument that complexity is meaningless; it is the way in which a stimulus is perceived that is important, not the number of elements [31]. If complexity correlates negatively with familiarity is it intrinsically bad? If familiarity is a part of the construct of complexity, then this is what researchers and designers may need to take into consideration rather than simply arriving at the best, context free measure of visual complexity. So, removing familiarity effects is not an advantage in its own right, it depends on what one wants.

3.2 Novelty and Interest

In addition to familiarity, one important concept which appears to be linked to complexity is the degree to which stimuli are able to capture attention and interest. There are many interface environments where capturing an individual's attention or relieving boredom by introducing interesting stimuli are more important than simply ensuring fast processing of simple stimuli. Berlyne [34] argued that interest is a monotonic function of collative variables such as novelty, complexity, surprise and ambiguity, suggesting that icon detail and intricacy is likely to be closely related to how interesting an icon is. At present however technology cannot measure an 'interest factor'.

A separation of the *symbol* property 'complexity' from 'detail and intricacy' in *icon* research is warranted. 'Detail' is perhaps a more neutral description of the structural components within an icon. Complexity implies difficulty; it suggests that an image will be more difficult to understand than a 'simple' image. This perhaps explains why observers rate familiar shapes (even nonsense shapes) as less-complex than they actually are [30]. What this means is that when *detailed icons* are used, the most important property for the observer is that they are meaningful. Interest and meaningfulness helps us focus our attention, and retrieve salient information about the message and reduce the interpretational burden.

3.3 Spatial Frequency Information

Forsythe [3] found that observers are unlikely to judge a detailed icon as simpler if it contained a large amount of low spatial frequency information, relative to high (i.e. ratings of complexity are reduced when shading is reduced). Queen [33] also found that responses were faster to icons and symbols that were of low spatial frequency and that this frequency was unique from the frequency of other icons in the set. Little attention, however, has been given to the interplay between visual complexity and shading.

The relationship between icon complexity and shading can be explained as follows. The visual system "knows" that an object will reflect different amounts of light, but this reflection does not depend on the properties of the stimulus [34]. Almost all the variation in light-levels is due to the illuminant, the physical properties of the object account for only a small fluctuation in the waveform. Depending on the reflective surface, the time of day, a dark night or a snowy day the light variation changes considerably. Encoding absolute light-levels would be an inefficient strategy for the neural system, thus the brain has adapted ways in which the importance of light levels can be minimised. It does this by attenuating to zero and near-zero spatial frequencies.

(a) Original **(b) Low spatial filter** **(c) High spatial filter**

Fig. 4. High and low spatial filtering

Low spatial frequency information is the most consistent of information that the visual system receives. By attenuating to it the visual system is able to maximise the perception of the object and 'know' where edges occur [34]. Figure 4 illustrates the differences between high and low spatial information.

This attenuation comes at a cost; small details are over looked. Ginsburg et al. [35] reported that pilots with high contrast sensitivity were able to detect a blocked runway at a greater distance than those with poor contrast sensitivity. Likewise, Harvey, Roberts & Gervais [36] reported that letters with common features [e.g. O, Q] were not confused when their spatial frequencies differed. However letters with non-common features [e.g. "A", "S"], but similar spatial frequencies were confused. These findings contradict feature integration theory [37, 38] and suggest that spatial frequency information is much more important to visual processing than the integration of image parts.

4 Conclusions

Compression techniques such as Gif and Jpeg offer researchers the most reliable and user friendly option for the quantification of visual complexity, they are also unbiased - they are not affected by familiarity with an image set. These metrics have a strong theoretical basis (information theory) and produce good approximations of human judgments. However, it is a reality that visual complexity is related to familiarity and researchers should consider what it is that they want from a measure of visual complexity and if removing familiarity from the equation is warranted. Further work may want to explore the usefulness of familiarity in our understanding of what is perceived as complex or simple. Finally, spatial frequency offers a new direction for complexity research. It seems likely that understanding the ratio of high to low spatial frequency information in icon design can improve reaction times and performance.

References

1. Snodgrass, J.G., Vanderwart, M.: A standardized set of 260 pictures. Norms for name agreement, image agreement, familiarity and visual complexity. Journal of Experimental psychology. Human Learning & Memory 6, 174–215 (1980)
2. McDougall, S.J.P., Bruijn, D.O., Curry, M.B.: Measuring symbol and icon characteristics: Norms for concreteness, complexity, meaningfulness, familiarity and semantic distance for 239 symbols. Behavior Research Methods 31(3), 487–519 (1999)
3. Forsythe, A., Sheehy, N., Sawey, M.: Measuring icon complexity: an automated analysis. Behavior Research Methods, Instruments, and Computers 35, 334–342 (2003a)
4. McDougall, S.J.P., Bruijn de, O., Curry, M.B.: Exploring the affects of picture characteristics on user performance: The role of picture concreteness, complexity and distinctiveness. Journal of Experimental Psychology: Applied 6, 291–306 (2000)
5. Arend, U., Muthig, K.P., Wandmacher, J.: Evidence for global feature superiority in menu selection by pictures. Behavior and Information Technology 6, 411–426 (1987)
6. Bryne, M.D.: Using pictures to find documents: Simplicity is critical. In: Proceedings of the conference on Human Factors in Computing systems, INTERCHI 1993. Addison-Wesley, Reading (1993)
7. Brunel, N., Ninio, J.: Time to detect the difference between two images presented side by side. Cognitive Brain Research 5, 273–282 (1997)
8. Rossion, B., Pourtois, G.: Revisiting Snodgrass and Vanderwart's object set: The role of surface detail in basic-level object recognition. Perception 33, 217–236 (2004)
9. Feldman, J.: How surprising is a simple pattern? Quantifying "Eureka!". Cognition 93, 199–224 (2004)
10. Garcia, M., Badre, A.N., Stasko, J.T.: Development and validation of icons varying in their abstractness. Interacting with Computers 6(2), 191–211 (1994)
11. Bruner, J.S.: Beyond the Information Given: Studies in the Psychology of Knowing. Norton, London (1973)
12. Hochberg, J.E.: Perception, 2nd edn. Prentice-Hall, Englewood Cliffs (1986)
13. Koffka, K.: Principles of Gestalt Psychology. Lund Humphries, London (1935)
14. Attneave, F.: Some informational aspects of visual perception. Psychological Review 61, 183–193 (1954)
15. Attneave, F., Arnoult, M.D.: The quantitative study of shape and pattern perception. Psychological Bulletin 53, 452–471 (1956)
16. Hochberg, J.E., Brooks, V.: The psychophysics of form: Reversible perspective drawings of spatial objects. American Journal of Psychology 73, 337–354 (1960)
17. Johnson, C.J., Paivio, A., Clark, J.A.: Cognitive components of picture naming. Psychological Bulletin 120(1), 113–139 (1996)
18. Geiselman, R.E., Landee, B.M., Christen, F.G.: Perceptual discriminability as a basis for selecting graphic symbols. Human Factors 24, 329–337 (1982)
19. Beck, H., Graham, N., Sutter, A.: Lightness differences and the perceived segregation of regions and population. Perception and Psychophysics. 49(3), 257–269 (1991)
20. Harwerth, R.S., Levi, D.M.: Reaction time as a measure of suprathreshold grating detection. Vision Research 18, 1579–1586 (1978)
21. Sutter, A., Beck, J., Graham, N.: Contrast and spatial variables in texture segregation: Testing a simple spatial-frequency channels model. Perception and Psychophysics 46(4), 312–332 (1989)
22. Vassilev, A., Mitov, D.: Perceptual time and spatial frequency. Vision Research 16, 89–92 (1976)

23. Hoeger, R.: Speed of processing and stimulus complexity in low-frequency and high-frequency channels. Perception 26, 1039–1045 (1997)
24. Parker, D.M., Lishman, J.R., Hughes, J.: Integration of spatial information in human vision is temporally anisotropic: evidence from a spatiotemporal discrimination task. Perception 26, 1169–1180 (1997)
25. Forsythe, A., Sheehy, N., Sawey, M.: The automated measurement of pictorial image complexity: a feasibility study. In: Harris, D., Duffy, V., Smith, M., Shephanisdis, C. (eds.) Human-Centred Computing: Cognitive, Social and Ergonomic Aspects, vol. 3, pp. 205–209. Lawrence Erlbaum, Hillsdale (2003b)
26. Zhang, D., Lu, G.: Review of shape representation and description techniques. Pattern Recognition 37, 1–19 (2004)
27. Vitevitch, M.S., Armbrüster, J., Chu, S.: Sublexical and Lexical Representations in Speech Production: Effects of Phonotactic Probability and Onset Density. Journal of Experimental Psychology: Learning, Memory, and Cognition 30(2), 514–529 (2004)
28. Donderi, D.: Visual Complexity: A review. Psychological Bulletin 132, 73–97 (2006)
29. Shannon, C.E., Weaver, W.: The mathematical theory of communication. University of Illinois Press, Urbana (1949)
30. Forsythe, A., Mulhern, G., Sawey, M.: Confounds in pictorial sets: the role of complexity and familiarity in basic-level picture processing. Behavior Research Methods 40(1), 116–129 (2008)
31. Rump, E.E.: Is there a general factor of preference for complexity? Perception& Psychophysics 3, 346–348 (1968)
32. Berlyne, D.E.: Novelty, complexity, and interestingness. In: Berlyne, D.E. (ed.) Studies in the new experimental aesthetics: Steps toward an objective psychology of aesthetic appreciation, pp. 175–180. Hemisphere Publishing Corporation, Washington (1974)
33. Queen, M.: Icon Analysis; Evaluating Low Spatial Frequency Compositions, Boxes and Arrows (2006), `http://www.boxesandarrows.com/view/icon_analysis`
34. De Valios, R., De Valios, K.: Spatial Vision. Oxford Series, vol. 14. Oxford University Press, Oxford (1990)
35. Ginsburg, A.P., Evans, D.W.: Contrast sensitivity predicts pilots' performance in aircraft simulators. American Journal of Optometry and Physiological Optics 59, 105–109 (1982)
36. Harvey, L.O., Roberts Jr., J.O., Gervais, M.J.: The spatial frequency basis of internal representations. In: Geissler, H.G., Buffart, H.F.J.M., Leeuvenberg, E.L.J., Sarris, V. (eds.) Modern (1983)
37. Treisman, A.: Features and objects in visual processing. Scientific America, 106–115 (November 1986)
38. Treisman, A., Gelade, G.: A feature integration theory of attention. Cognitive Psychology 12, 97–136 (1980)