APPLYING BOTTOM-UP VISUAL PERCEPTUAL PATTERN DESIGN PRACTICE

The early-stage visual system rapidly processes shared patterns and characteristics which when utilized in products can improve the experience design

Gupp, Erik
egupp@bentley.edu

Introduction

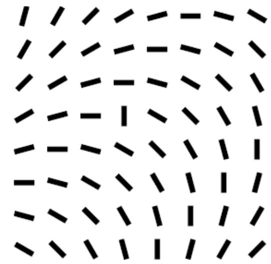

FIGURE 8.4 Local contrast produces bottom-up guidance. Note that there are five vertical lines in this display. Only one is salient.

When you looked at the image to the right, what was the first thing you immediately noticed? Most likely it was the vertical line in the middle even though you were not told what to look for. "The more an item differs from its neighbors, the more attention it will attract, all else being equal" (Wolfe & Gray 2007). The reason behind this is how the human visual system has evolved to first detect change and see the world in patterns. Within this design review, I will focus on the preattentive bottom-up process which is the automatic formation of image feature maps based on shared perceptual qualities and characteristics that make objects identifiable and pop out.

Movements of the Eye & Visual Searching

When you open your eyes, you begin moving visually from fixation to fixation. Foveal view is what is in your direct line of sight, while peripheral vision is outside of that. As we scan our environment, we focus our line of sight on areas of interest where task related information is seen. Well designed products are those that have easily visually identified features that help improve the interaction experience rather than confuse users.

When humans look around, they are employing different types of eye movements. If you are locked into a singular target that is moving that is a smooth-pursuit eye movement. In contrast, saccades are the very rapid movements of your eyes between different targets in an area of interest. "When eye movements are recorded in a visual search paradigm in which display size varies, it is found that the number of saccades prior to a target present/target absent decision is an increasing linear function of display size". Interpreting this result shows that during an unguided search between eight to ten items are being processed in parallel. (Liversedge & Findlay 2000).

These different types of eye movements speak to the visual mechanisms of serial and parallel search. Serial search is guided, and models behavior where a person searches for a particular target among distractors. This employs more top-down cognitive processing and is based on previous experience and working memory. (Wolfe & Gray 2007) Where our focus is on this design review is in the parallel stage of bottom-up cognitive processing.

Within Treisman's feature-integration theory of attention, "features come first in perception. In our model, features are registered early, automatically, and in parallel across the visual field, while objects are identified separately and only at a later stage, which requires (top-down) focused attention." (Treisman & Gelade 1980). In this bottom-up visual processing model, features of a visual scene are processed first in parallel with some features sticking out more than others. “Bottom-up mechanisms are thought to operate on raw sensory input, rapidly and involuntarily shifting attention to salient visual features of potential importance – the spot of red against a field of green that could be a piece of fruit, the sudden movement that could be a predator.” (Connor et al., 2004)

Neurological & Biological Systems

That red fruit against the green tree is a salient high contrast stimulus that is first processed in our retinas. That visual information is then passed up the optic nerve through a neural junction at the lateral geniculate nucleus (LGN). The LGN belongs to the “category of sensory projection nuclei of the thalamus” and has three distinctive cell types of magnocellular (M), parvocellular (P) and koniocellular (K). These cells are arranged in six different layers, with “M cells receiving input from large-field, motion-sensitive Y-type retinal ganglion cells, while P cells receive input from the small-field, color-sensitive X-type retinal ganglion cells”. (Covington & Al Khalili 2020)

The primary role of the LGN is to relay information to the visual cortex, wherein visual area 1 (V1) and visual area 2 (V2) are the first areas to receive those visual inputs. Each hemisphere of the brain has a visual cortex and receives visual information for the brain to ultimately recognize objects and patterns quickly without a “significant conscious effort”. (Huff et al., 2019) Specifically, what V1 and V2 are tuned to are pattern properties including orientation and size factoring in luminance, color, motion, contours, and texture.

“The great majority of neurons in the primary visual cortex of many primates are exquisitely sensitive to the orientation of a stimulus.” In the feed-forward model described by Hubel & Wiesel, orientation selectivity emerges from simple cells made up of elongated ON and OFF subfields. (Ferster & Miller 2000) V1 neurons are also highly tuned to spatial frequency which refers to bar-based patterns of alternating contrast light and dark patterns. These patterns are called gratings, and high spatial frequency refers to an increased number of gratings within a given distance. As we age, we become less sensitive to targets of low spatial frequency, however a study found both older and younger observers did not differ in their ability to see images with high spatial frequency. (Sekuler et al., 1980)

Color vision is a result of three classes of photoreceptor cells combined in the visual cortex. In a study using primates, it was found many V1 neurons combine cone signals nonlinearly. (Horwitz & Hass 2012) In the opponent process theory developed by Hering, there are three distinct opposing color channels of red-green, blue-yellow, and white-black. “That is, we may experience red-blues or green-blues but never yellow-blues, and we see yellow-greens or blue-greens, but never red-greens, and so on." This theory is in opposition to the Young-Helmholtz trichromatic theory which says we have separate blue, green, and red cones which translate different wavelengths of light into color. (Hurvich & Jameson 1957)

Visual perception of motion is served by three separate systems. The first order computation of the directional strength of visual motion is basic and responds to moving luminance patterns. The second order motion responds to modulations of texture contrast, and the third responds to the motion of marked locations in a feature map. (Lu & Sperling 2001) Important within motion or static pattern perception is the pop-out effect, where a target's location can be quickly detected among distractors.

V1 and V2 are also responsible for detecting both symmetric and asymmetric contour patterns. End-stopping cells are sensitive to contrast at the end of a line and help us quickly understand direction. (Yazdanbakhsh & Livingstone 2006) Contour detection algorithms can be split into two distinct categories, local and global methods. In the former, the defining features are luminance, color, and texture while the later is identified on good continuation and closure. "Contour detection is a difficult task due to various practical reasons, such as possible low signal-to-noise ratio (SNR) or presence of textures in an input image. (Papari & Petkov 2011)

Objects with similar textures will be grouped together with segmentation occurring with measurable differences between the foreground and background in local mean contrast, local spatial frequency, orientation, and other local properties. (Landy & Bergen 1991) Within complex backgrounds, texture segmentation can be efficiently detected by passing an image through even-symmetric Gabor filters then computed over multiple spatial frequencies and orientations. (Jain et al., 1997)

Preattentive Processing

Preattentive processing is an initial, automatic, and subconscious organization of the visual field based on the factors mentioned above including orientation, color, size, motion, and textures. Individual eye movement saccades take at least 200 milliseconds to initiate, and visual tasks that can performed in that time or quicker are considered preattentive. (Healey 2007) This is the function of human vision that categorizes what we are looking at prior to when we pay attention to that target. "If the target differs from the distracters in number or kind of certain elementary features then human observers can effortlessly detect the presence of the target regardless of the number of distractors present." (Knierim & Van Essen 1992) This statement holds mostly true until the variety of distractors grows, at which point preattentive features do not pop out as efficiently. "This pattern of performance suggests parallel processing when the target has a unique distinguishing feature and serial self-terminating search when the target is distinguished only by the absence of a feature that is present in all the distractors. The results are consistent with feature-integration theory." (Treisman & Souther 1985)

Feature-integration theory proposed by Treisman suggests that features are coded in parallel and there are individual "maps" for each of those visually distinctive features with a master map of locations. (Treisman 1994) If you are looking at a product feature, the color of it is processed in parallel along with the orientation and size for example. In Wolfe's guided search theory, feature maps are still prevalent but differ in that instead of individual blue, green, red color maps there is just one "color" map and this feature categorization follows bottom-up processing.

One complication with feature maps is when a search contains a conjunction, where distractors include a combination of factors that are not unique. This type of search is not normally processed preattentively, but "search for conjunctions of highly discriminable features can be rapid or even parallel." (Treisman & Sato 1990) The degree of how much a conjunction search is parallel is debated among the various preattentive theories including feature integration, guided search, texton or similarity.

In a discussion on high-speed visual estimation using preattentive processing they investigated two preattentive features of hue and orientation in the context of numerical estimation tasks. While their focus was on data displaying software programs, watch faces can be substituted as they also have simple visual features including colors, shapes, and sizes. What they found was that rapid and accurate estimation was indeed possible using either hue or orientation preattentive values. (Healey et al., 1996)

Design Review

The wristwatch is a ubiquitous accessory found no matter where you go in the world. They are products designed to be worn by a wide range of demographics, from young users to old, in multiple use cases from dark environments to the bright outdoors. To appeal to all these users, watches come in multiple visual designs with available features that provide additional information to their wearers. Any visual feature that appears on a watch should have consideration if that can be processed preattentively, or if it includes a design that forces users to spend too much time searching instead of subconsciously pulling out patterns and identifying important information.

In their latest range of Aviation Pilot style watches, the Hamilton Converter Auto Chrono displayed to the right is a good example of a watch with a complex set of patterns and features. There are varying hues, luminance, lines, sizes, motions, contours that can complicate bottom-up processing. The number of feature maps you are visually processing is high in this design, which means many things are working in parallel.

When visually first landing on the watch face target, the most salient feature should be the hour and minute hands. It is a watch, so the quickest takeaway a user should have is knowing what time it is. The issue is that the white & gold outline of the hour/minute hands and hour markers are not immediately noticed due to the proximity and similarity of those hues used in other elements. My recommendation would be to increase the luminance of these features, add a textural component or change the hue to something not used in any other feature.

A noted feature of this watch is how it has been designed to help users calculate critical conversions and metrics. Around the bezel, there are various words written in red, such as GAL and MPH, indicating those various metrics. A positive about this design is that red is used sparingly, so they are preattentively perceived quickly. One issue user's may have is connecting the red tipped arrow on the face pointing to those words. There is a gap between the arrow and the bezel, so there is no connectedness visually helping the user. A design solution would be better connecting those points of information to reduce the cognitive load.

Where the design of the watch also falls short is within the three chronographs within the watch face. There is a lack of symmetry between all three, with varying measurement scales, line sizes and contour coloring. In an updated design, I would add more consistency to how those three are treated. Adding concavity or convexity would also help, as they are currently flat against the face and that switch would help them be more detectable.

The last feature I would consider updating is the month and day of the week feature on the watch face. With white text outlined in gold, it is again too similar despite the rectangular shape of the feature. I would inverse the colors and have black text against a white background to vary the element and have a contrasting effect that creates a unique identify for that feature.

Conclusion

When looking at a product, certain elements will naturally stick out due to our subconscious bottom-up pattern processing visual neurology. Poor design will have features that all seemingly bleed into each other, with the user misdirected, confused, or spending more time processing than necessary. Good design should segment its features into different groupings, and then design those groupings with contrasting elements that make them stick out in your visual feature maps. With watch faces, those potential groupings are current time, current date, chronograph timers, and conversion metrics. By having different hues, textures and orientations represent those different groupings users will have an improved experience that balances effectiveness and efficiency of visual information gathering.

Bibliography

Connor, C. E., Egeth, H. E., & Yantis, S. (2004). Visual attention: bottom-up versus top-down. Current biology, 14(19), R850-R852.

Covington, B. P., & Al Khalili, Y. (2020). Neuroanatomy, Nucleus Lateral Geniculate. In StatPearls [Internet]. StatPearls Publishing.

De Valois, R. L., & De Valois, K. K. (1980). Spatial vision. Annual review of psychology, 31(1), 309-341.

Ferster, D., & Miller, K. D. (2000). Neural mechanisms of orientation selectivity in the visual cortex. Annual review of neuroscience, 23(1), 441-471.

Healey, C. G. (2007). Perception in visualization. Retrieved February, 10, 2008.

Healey, C. G., Booth, K. S., & Enns, J. T. (1996). High-speed visual estimation using preattentive processing. ACM Transactions on Computer-Human Interaction (TOCHI), 3(2), 107-135.

Horwitz, G. D., & Hass, C. A. (2012). Nonlinear analysis of macaque V1 color tuning reveals cardinal directions for cortical color processing. Nature neuroscience, 15(6), 913-919.

Hurvich, L. M., & Jameson, D. (1957). An opponent-process theory of color vision. Psychological review, 64(6p1), 384.

Huberman, A. D., Feller, M. B., & Chapman, B. (2008). Mechanisms underlying development of visual maps and receptive fields. Annu. Rev. Neurosci., 31, 479-509.

Huff, T., Mahabadi, N., & Tadi, P. (2019). Neuroanatomy, visual cortex. In StatPearls [Internet]. StatPearls Publishing.

Jain, A. K., Ratha, N. K., & Lakshmanan, S. (1997). Object detection using Gabor filters. Pattern recognition, 30(2), 295-309.

Knierim, J. J., & Van Essen, D. C. (1992). Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. Journal of neurophysiology, 67(4), 961-980.

Landy, M. S., & Bergen, J. R. (1991). Texture segregation and orientation gradient. Vision research, 31(4), 679-691.

Liversedge, S. P., & Findlay, J. M. (2000). Saccadic eye movements and cognition. Trends in cognitive sciences, 4(1), 6-14.

Lu, Z. L., & Sperling, G. (2001). Three-systems theory of human visual motion perception: review and update. JOSA A, 18(9), 2331-2370.

Papari, G., & Petkov, N. (2011). Edge and line-oriented contour detection: State of the art. Image and Vision Computing, 29(2-3), 79-103.

Sekuler, R., Hutman, L. P., & Owsley, C. J. (1980). Human aging and spatial vision. Science, 209(4462), 1255-1256.

Treisman, A., & Souther, J. (1985). Search asymmetry: A diagnostic for preattentive processing of separable features. Journal of Experimental Psychology: General, 114(3), 285-310. http://dx.doi.org.ezp.bentley.edu/10.1037/0096-3445.114.3.285

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. Cognitive psychology, 12(1), 97-136.

Treisman, A. (1994). Visual attention and the perception of features and objects. Canadian Psychology/Psychologie Canadienne, 35(1), 107-108. http://dx.doi.org.ezp.bentley.edu/10.1037/h0084715

Treisman, A., & Sato, S. (1990). Conjunction search revisited. Journal of experimental psychology: human perception and performance, 16(3), 459.

Vision. (2016). In J. L. Longe (Ed.), The Gale Encyclopedia of Psychology (3rd ed., Vol. 2, pp. 1198-1200). Gale. https://link.gale.com/apps/doc/CX3631000784/SCIC?u=mlin_m_bent&sid=SCIC&xid=b6e9fc78

Wolfe, J. M., & Gray, W. (2007). Guided search 4.0. Integrated models of cognitive systems, 99-119.

Yazdanbakhsh, A., & Livingstone, M. S. (2006). End stopping in V1 is sensitive to contrast. Nature neuroscience, 9(5), 697-702.