Illustration and Perception

Dirk Bartz[1], Zein Salah[2], and Douglas Cunningham[3]

[1] Universität Leipzig, ICCAS/VCM, Leipzig, Germany
[2] Universität Magdeburg, ISG, Magdeburg, Germany
[3] MPI for Biological Cybernetics, Psychophysics, Tübingen, Germany

Abstract Visual Computing is concerned with the processing of image-oriented information. It explores the combination of image analysis, computer graphics, and interaction. In a medical context, it provides methods and means for processing and presenting information to physicians to enable better patient treatment. Currently, the efficacy of visual computing relies on the individual abilities of engineers and physicians to communicate the relevant information. Illustrative visualization is a methodology that aims at improving the efficacy through reducing the visual complexity of irrelevant data, while keeping a focus on the relevant features of the data.
While the scientific concept appears sound and is based on basic perception research, it is not yet validated. Hence, psychophysical experiments must be conducted verify the illustrative visualization concept and evaluate if a better perception of the information is achieved. Here, we will discuss the concept and show early validation results.

1 Introduction

Illustrations are widely used in science, engineering, and medicine to visually represent an abstraction of an object. Typically, these illustrations are drawings, where the illustrator applies artistic techniques to emphasize relevant aspects of the objects. In contrast, visualization extracts a visual representation from simulations, scanned datasets, or modeled data sources. The current complexity of such datasets, however, renders the interpretation difficult.

Illustrative visualization aims at combining both approaches to create an abstracted representation of the dataset, where major visual features are highlighted. This is in particular important for large models that contain a large number of individual objects, like a mechanical part of an MCAD model. Illustrative visualization generally highlights selected, relevant information better than traditionally rendered images. However, rendering multiple proximate parts still represents a challenge, since they need to be clearly disambiguated.

In this paper, we present a guidelines-based approach for the combined illustrative visualization of models composed of multiple objects. Obviously, the visual differentiation between objects necessitates applying different styles for the better perception of the spatial relationship between structures. By additionally adjusting object attributes such as color, transparency, and silhouette thickness, the focus can be drawn to targeted structures. The guidelines-based approach presented here restricts the possible number of combinations through static heuristics derived from perception research (see Section 2) and it was validated in a perceptual experiment.

Note that our approach is not aimed at the visual exploration of unknown data objects, but at the presentation of well-segmented objects, e.g. for technical illustrations, based on the guidelines and a specified interest classification. Although we present several examples of rendering MCAD data, the approach presented in this paper is not limited to this class of datasets and can be applied as well to models from medical scenarios.

2 Guidelines for Rendering of Multi-Object Models

While the illustration of individual objects is relatively easy, the complexity of the illustration problem increases when rendering multiple neighboring objects (e.g., piston and cylinder head of a cylinder). In particular, the quality of the final rendering and the emphasis expressed in this rendering depends on the selection of the presentation styles.

Various illustrative rendering and shading techniques are used by our guideline-based visualization assistant. They provide a range of possible combinations on which the quality of rendering multi object models depends heavily. In our context, object silhouettes, saturated colors and transparencies, and different illumination effects are used for the different entities.

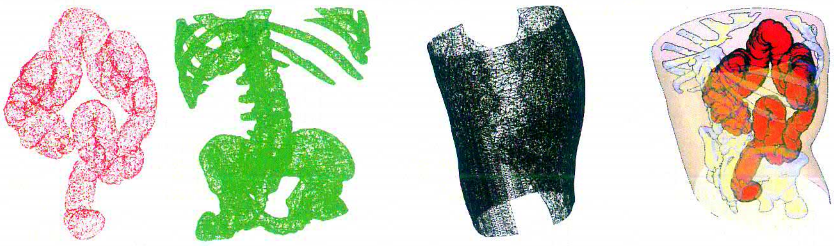

Figure 1. Extracted models of an abdomial dataset: colon, bones, and skin (left to right). The final rendering (right most) represents an example model object classification: Extracted focus (colon), context (bones), and container objects (skin).

Object Classification. We classify objects for a combined rendering into: *focus*, *near-focus*, *context*, and *container* objects (see Fig. 1). Focus objects are those of highest interest. Near-focus are objects in direct semantic connection with the focus object, and context objects are all other surrounding or neighboring objects of the model. A container object encloses a subset of focus, near-focus and context objects. In practice, this corresponds to the outer shell in a technical illustration (or to the skin in anatomical datasets). Note, however, that our perceptual validation in Section 3 differentiates only focus and non-focus objects.

Pop out Effect. Perception research has shown that certain visual attributes grab one's attention and "pop out" of a cluttered field during a visual search [5,3]. In other words, the time it takes to find "pop-out" features when they are embedded in a cluttered field is greatly reduced in comparison to other attributes. Furthermore, search times for "pop-out" features are nearly independent of the number of non-focus (or non-target) objects. This is in strong contrast to other attributes, whose search times generally increase linearly with increasing numbers of distractor items. One explanation for this effect claims that these attributes are processed prior to the allocation of attention, very early in the lower level of the human visual system (HVS), and thus refers to these attributes as "pre-attentive" [5].[1] In the following, we refer to these attribtutes as lower level attributes or features. It is important to note that different lower level features may interfere (or interact) with each other, thus possibly requiring linear time to identify the target object. In [1], it is shown that the human visual system prioritizes different visual cues; color is prioritized against shape, and brightness over hue (after normalization). Some cues, however, seem to be independent [3].

Here, we mainly use saturated hue, differences in transparency, and silhouettes (as boundary representations) as lower level attributes to highlight our focus objects, and to a smaller extent the near-focus and context objects.

Which rendering algorithm. For rendering, the shading styles and parameter tuning can be applied equally well. Hence the choice of the algorithm is mainly dependent on the type of the available representation of objects. If a segmentation of voxel objects is provided, the point-based algorithm can be applied, thus avoiding a triangulation step.

Color. Focus objects are emphasized with warm signal colors (e.g., red) that are fully saturated and with full luminance that provide a sufficient contrast to the surrounding objects. All other objects are given less saturated, possibly cooler colors than focus objects. In any case, their colors should be linearly separable to the focus object color to ensure pop out effects [2]. Significantly less saturated colors with low luminance of a different color category should be used for the container objects (see images of screwdriver in Fig. 2).

Which shading style. Container objects should be shaded with a less attention attracting style – e.g., standard Gouraud shading. Cool-to-warm shading provides a good differentiation, since the typically used colors are linearly separable and hence provide good target (focus) detection [4]. Hence, it does not suit container objects, unless it emphasizes a relevant feature (i.e., material property). On the other hand, for near-focus and context objects, cool-to-warm shading would help them to appear quite clear and distinctive. It is not advisable to apply the same style to both container and context objects, as they may become barely distinguishable. The same is true for near-focus and context ob-

[1] This explanation is still controversial in the perception literature, since some research indicates that the heightened saliency and efficiency of "pop-out" features is, in fact, due to a form of attention and not "pre-attentive".

jects, if they need to be differentiated. In this case, near-focus objects should use a stronger emphasis than context objects. For shading focus objects, halftoning is a good choice, especially for large, innermost objects.

Degree of transparency. Obviously, container objects must be very translucent to see the objects inside. For the other objects, we apply the *degree of interest* concept. In essence, objects that lie within the same nesting level from the viewer are given transparencies inversely proportional to their degree of interest. Focus objects are rendered with high opacity (or totally opaque, if only one exists), in contrast to near-focus objects, which may be rendered with a somewhat lower opacity.

Silhouette. For illustration purposes, rendering with silhouettes reveals better expressiveness and shape perception, since they serve as supportive cues for figure-to-ground segregation. Hence, we always render focus, and near-focus, and sometimes context objects with outlined silhouettes. Moreover, the thickness of the silhouette reflects the degree of object importance. Therefore, we render focus object with thicker silhouettes.

3 Perceptual Validation

Although the individual guidelines are based on general perceptual theory, it remains an empirical question as to whether their combinations have the intended effect in visualization. Here, we present a perceptual experiment designed to see whether it was easier to find the focus object of a complex, multi-object model in images generated according to the guidelines than in images that do not follow the guidelines.

Given the number of possible models, component objects, viewpoints, and rendering parameters, it is impossible to explore the whole parameter space in a single experiment. The goal of this experiment is instead to provide an initial examination of some of the more extreme predictions of the guidelines. If no – or only small – effects are found for extreme violations, then detailed exploration of the guidelines would not be necessary. If, on the other hand, large effects are found where the guidelines predict them to be, then subsequent experiments can explore the more precise predictions and specific parameter dependencies in detail. Thus, the primary manipulation in this experiment is guideline conformity. To test this, 12 image pairs were generated. The two images in an image pair were nearly identical, differing solely in their conformity to the guidelines (see Fig. 2 for examples). Hence, the experiment only differentiates between focus and non-focus objects.

To produce some variation within the test set, and thus help to ensure that the results generalize beyond the specific images used, we used several different models, violations, viewpoints, and focus objects. More specifically, three different MCAD models (a car motor, a cross-section of a truck motor, and an electric screwdriver were used. Second, two potential focus objects were employed for each model (with only one per image). Third, two viewpoints were used per object per model. The combination of three models, two focus objects, and two viewpoints yielded 12 images, each of which was rendered twice (once conforming to the guidelines, one violating them) for a total of 24 trials. Obviously, three models are not enough to examine the effects of object type, complexity, or any other object or scene dimension in any detail – that is left to future work once it is determined if there is an effect of guideline conformity. Likewise, the manipulations of focus object and viewpoint are not intended to sample the parameter space or provide detailed insight into these dimensions, but instead to help improve generalizability. Finally, to maximize the test of extreme predictions, the specific violations depended on the model and focus object. It is important to emphasize, however, that the two images in a violation pair were identical except for the changes in guideline conformity.

Study Design. Ten experienced psychophysical observers participated in the experiment. Each was naïve to the purposes of the experiment. Additionally, each participant was reimbursed for their time at standard rates, was selected randomly from the participant database of a major perception research institute, had normal color vision, had normal or corrected to normal visual acuity, and had no known visual deficits. Each was seated, one at a time, in a closed, well-lit room. No one other than the experimenter and the participant was present in the room. The same experimenter was used for each participant. The participants were informed that we had generated a series of images following a set of guidelines (the specifics of which were not mentioned) regarding the visualization of models of complex objects. The participants were told that they would be shown the models, one at a time, and that they must examine each model to determine the focus object. The concept of focus object was explained, until it was clear that the participants understood the task. They were told that they would see each model several times, sometimes with different focus objects. They were told that both accuracy and speed were important. Each trial in the experiment started with the participant pressing the spacebar, which caused a single image to be displayed centered on a standard 53cm (21 inch) LCD screen set to 1280×1024 resolution. The images subtended approximately 20 degrees of visual angle along the largest axis. The participant was asked to examine the model until they decided which object was the focus object, and then press the spacebar again and verbally report their choice. After the experiment, they were given an informal debriefing. The images were shown in random order to the participants, with each participant receiving a different order.

Results. Overall, the participants were able to find the focus object rather well (total average percent correct was 71.7%). In both reaction times and accuracy rates, the images generated by following the guidelines produced superior performance (90.8% versus 52.5%; 2.3 versus 3.5 seconds). If one subtracted the scores for each violation image from the scores for their paired consistent image, the resulting difference score would provide some additional insight into the results. With one exception, the violation image in each pair produced lower accuracy scores and slower reaction times than the paired consistent images. Interestingly, both the smallest and the largest accuracy effects were seen for the same focus

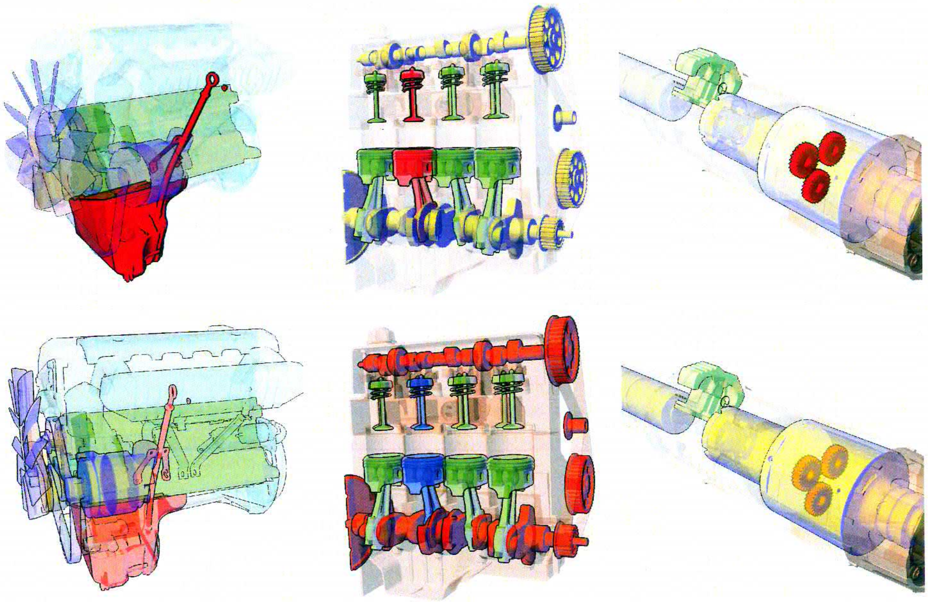

Figure 2. Examples of stylized multi-object object visualization. Objects are parametrized according to the guidelines (top row) and in violation of the guidelines (bottom row). Left to right: Car Motor - Oil pan is select as focus object. Truck Engine - 2nd cylinder is selected as focus objects. Electric Screwdriver - Gears are selected as focus object. Note the specific images on the right were chosen to be clear, extreme violations of the guidelines.

object: The car motor with the fan as the focus object. For one viewpoint, performance was 100% for both consistent and violation images despite large violations of color and transparency. For the second viewpoint, performance in the violation condition dropped to 20% despite the fact that the violations are nearly identical. A closer look at the images suggests that the location of the focus object may be in large part responsible for the difference. In the “easy” viewpoint, the focus object is in the center of the image, in front of and occluding all other objects. In the “difficult” case, the focus object is off-center and not occluding anything, allowing the violations of color to be more prominent. Note that the reaction time difference for the easy case was relatively high (1.45 second), while it was low in the hard case (0.98 second), demonstrating a speed/accuracy trade-off. This highlights the importance of measuring both accuracy and reaction time, and suggests that the effect may not be as pronounced as it at first seems. Nonetheless, there seems to be interesting interaction between transparency and color modulated by viewpoint.

Finally, it is important to re-emphasize that the goal of stylization is to highlight one form of information and de-emphasize others. Here, our goal was to highlight which object was of specific interest while maintaining information about its relation to the rest of the model. The results clearly show that the first part of the goal was met – guideline conformity made it easier to find the focus object. A brief glance at the images also shows that the second aim was at least qualitatively met, the general relation of the object to the model was retained. It is unclear, however, if the stylization had any quantitative effect on other critical image properties. For example, while the relative location of the objects within a model may be clear, the metric spatial relations may be altered by the changes in transparency.

4 Conclusion

In this paper, we have presented an approach to select illustrative presentation styles for multi-object models and validated the conformity of the guidelines of that approach through a perceptual experiment.

The major goal of our work is to obtain a clearly depictable display of the spatial relationship of the various objects, like different organs for anatomical data. To achieve that goal, we combined an illustrative rendering method using stylized shading with different object attributes (transparencies, color, and silhouettes), which were designed to draw the focus of a user to or away from specific objects. In particular, we showed how conventional illustrative techniques can be integrated to produce high quality, expressive illustration of proximate objects.

In order to validate the conformity of our guidelines, we conducted a perceptual experiment, which showed that presentation styles according to the guidelines generated significantly better emphasis results than violating them. Consequently, the experiment overall confirmed the encoded static heuristics.

References

1. T. Callaghan. Interference and Domination in Texture Segregation: Hue, Geometric Form, and Line Orientation. *Perception & Psychophysics*, 46(4):299–311, 1989.
2. M. D'Zmura. Color in Visual Search. *Vision Research*, 31(6):951–966, 1991.
3. C. Healey, K. Booth, and J. Enns. Harnessing Preattentive Processes for Multivariate Data Visualization. In *Proc. of Graphics Interface*, pages 177–184, 1993.
4. I. Spence and A. Efendov. Target Detection in Scientifc Visualization. *Journal of Experimental Psychology*, 7(1):13–26, 2001.
5. A. Treisman. Preattentive Processing in Vision. *Computer Vision, Graphics, and Image Processing*, 31(2):156–177, 1985.