A Salience-based Quality Metric for Visualization

H. Jänicke\textsuperscript{1} and M. Chen\textsuperscript{2}

\textsuperscript{1}Ruprecht-Karls-Universität Heidelberg, Germany
\textsuperscript{2}Swansea University, Great Britain

Abstract

Salience detection is a principle mechanism to facilitate visual attention. A good visualization guides the observer’s attention to the relevant aspects of the representation. Hence, the distribution of salience over a visualization image is an essential measure of the quality of the visualization. We describe a method for computing such a metric for a visualization image in the context of a given dataset. We show how this technique can be used to analyze a visualization’s salience, improve an existing visualization, and choose the best representation from a set of alternatives. The usefulness of this proposed metric is illustrated using examples from information visualization, volume visualization and flow visualization.

Categories and Subject Descriptors (according to ACM CCS): I.3.3 [Computer Graphics]: Picture/Image Generation—Line and curve generation

1. Introduction

To which parts of a visualization do viewers pay attention, and how well does the expected attention match with the importance of the data depicted? These are questions that every visualization creator would be curious about as the answers would affect how a visualization is perceived and on what the observers learn from it. However, these questions are also rather intricate, and hence difficult to answer.

It is known that salience detection is a principle mechanism to facilitate visual attention \cite{Pal99,TCW95,DD95}, especially in object-based attentional selection \cite{Dun84,RLS98}. Visual salience measures how much an item stands out with respect to neighboring items. The higher this value, the more visual attention it attracts. This suggests that the effectiveness of a visualization can be improved by having a desired level and spatial location of salience in the visualization. It also suggests that a measure of “appropriateness” of the salience should be a quality metric for visualization.

Salience has been featured in previous works in visualization. For example, Ebert and Rheingans \cite{ER00} proposed to use non-photorealistic rendering to enhance salient regions of volume data. Hladuvka et al. \cite{HKG01} employed second derivatives to address the same problem. Kim and Varshney \cite{KV06} developed a tool that allows users to specify salience-based emphasis to guide the viewer’s visual attention. They also presented a user study where eye-tracking was used to show the effectiveness of salience-guided visualization. In \cite{LVJ05}, Lee et al. used salience to extract interesting regions of a surface mesh. Unlike these techniques, where visual salience was tailored to improve a specific visualization technique, we aim at a general metric to assess the quality of a visualization.

Dictionaries define quality in general as a grade of excellence or superiority. As for visualizations, a technique is commonly considered to be superior to another one, if questions concerning the underlying data can be answered more easily, faster or more accurately. Often this aspect of a visualization’s quality is assessed with user studies \cite{KHI03}, where subjects have to perform tasks that relate to the target group’s interaction with the visualization. User studies comparing different visualization techniques were carried out, for example, for uncertainty \cite{SZB09}, flow \cite{LKJ05,FCL09}, hurricane \cite{MSM08}, and volume \cite{HGL99} visualization.

While user studies are very useful to evaluate some fundamental characteristics of a technique, it is not possible to conduct a user study for each individual visualization every time it is created. The quality of visualization images depends on main factors, including the domain-specific requirements, the user’s needs and expectations, the source data set and the techniques used. Hence, it is desirable to pro-
Figure 1: Visualization-based importance: (a) Relevant structures in the (b) visualization as defined by the visualizer, where black areas are most relevant. (c) Visual salience of the image in (b).

2. A Salience-based Quality Metric

Visual salience is an important property in all visualizations. It defines which parts of the image stand out and will likely attract a lot of attention. Ideally these areas should coincide with the important parts of the data set. To measure the correspondence between relevant structures and the observer’s attention, these quantities have to be defined for each pixel in the visualization image, and can be computed by the visualizer to evaluate and optimize the visualization under construction. Here we purposely divide the users into two categories, namely visualizers who create visualization, and observers who are the target viewers of the visualization. Sometimes, a visualizer can also be the sole target observer.

2.1. Data Relevance

Data relevance is a measurement of importance from the visualizer’s perspective. In the image space, data relevance can be defined as a relevance mask, which tells for each pixel in the image, how important the corresponding data is in the context of the visualization. For example, for a sample visualization in Figure 1(b), a data relevance mask is shown in Figure 1(a). Ideally the data relevance would correlate to the salience of the visualization as depicted in Figure 1(c).

For some quality metrics (e.g., abstraction [CWM’09]), data relevance is an integral part of computation. For other quality metrics (e.g., esthetics [FB09]), data relevance is often not included in the computation at all. A salience-based metric has the flexibility to accommodate both approaches.

In many situations, the visualizer knows which parts of the visualization are important and shall draw the attention of the observer. Commonly these parts depict important features or structures in the data set. In these situations, a data relevance mask can be created simply in the image space using a brush or space-fill tool, which are available in most image editing software. Such a tool can also be incorporated in a visualization software to support a quality metric.

In situations where the important features and structures are unknown to the visualizer, automated techniques such as feature detection algorithms (e.g., [PW94, Hen98]) or data statistics (e.g., [JWSK07]) can be used in the data space. The computed object-space data relevance is then projected onto the image space, resulting in a data relevance mask.

2.2. Visual Salience

The human visual system is very good at analyzing imagery data. Complex scenes can be interpreted in real time, despite the limited speed of the neuronal hardware. This high performance can only be achieved by selective analysis of the scene [TCW’95]. Visual salience provides the human vision system with stimuli that attract our attention, facilitating the selective analysis. Neural mechanisms and computational models of visual salience have been extensively studied in several disciplines, including neuroscience, physiology, psychology, and computer vision. There is a large volume of literature, ranging from models that mimic the human perception (e.g. [DD95, WRKP05, IKN98]), techniques
The first step towards a salience map is the computation of feature maps. A salience map then consists of three steps: compute the motion in salience computation, as we consider here only static visualizations. The algorithm to compute a salience map is based on the normalized color components. Normalization is performed with respect to intensity to decouple hue from intensity. The color opponency feature map for red/green \( F_{RG} \) is given by:

\[
F_{RG}(p, s_1, s_2) = \sum_{x \in \text{Nh}(p)} |(r(p, s_1) - g(p, s_1)) - (g(p, s_2) - r(p, s_2))| \tag{3}
\]

The third property, orientation, is measured using oriented Gabor pyramids for angles \( \theta \in \{0^\circ, 45^\circ, 90^\circ, 135^\circ\} \). The resulting orientation maps \( O \) tell how good each of the four filters represents the local image structure. The feature map, with angle \( \alpha \) of the Gabor filter, is given by:

\[
F_O(p, \alpha, s_1, s_2) = \sum_{x \in \text{Nh}(p)} |O(x, \alpha, s_1) - O(x, \alpha, s_2)|. \tag{4}
\]

2.2.2. Across-scale Analysis

The feature map computation results in 42 feature maps. The final goal is to compute a salience map that summarizes salience using a single scalar value for each pixel. Hence, the different values have to be combined in a single number. This is achieved in two steps. First, the feature maps for each property are combined, and in a second step, the salient regions are extracted from all three combined maps.

The three maps combining the different scales are called “conspicuity maps”. They are obtained through across-scale addition of the individual feature maps. Therefore, each feature map is reduced to scale 4 and normalized \( N(x) \), see [IKN98] for more detail) and afterwards, pixel values are added point-wise:

\[
F(p) = \sum_{s_1, s_2} N(F_i(p, s_1, s_2)), \tag{5}
\]

where \( s_1 \) and \( s_2 \) are defined as earlier. If applicable, different modalities are added as well and we receive conspicuity maps for intensity, color opponency, and orientation.

2.2.3. Salience Map

In our analysis we found that the contribution of the three conspicuity maps to total salience is not equal. Especially color is a very dominant visual clue and attracts a lot of visual attention. In Section 3.2, a contribution map is presented that depicts the contribution of the channels to visual salience. Based on this map, the formula of total salience is modified to account for different contributions:

\[
S = \lambda_1 N(T_I) + \lambda_2 N(T_C) + \lambda_3 N(T_O), \tag{6}
\]

where \( \lambda_1, \lambda_2 \) and \( \lambda_3 \), with \( \lambda_1 + \lambda_2 + \lambda_3 = 1 \), are weights to account for variable contributions.
2.3. Quality Metric

After the definition of data relevance (relevance mask $I$) and visual salience (salience map $S$), the quality of a given visualization can be determined. The goal is to create a visualization that directs visual attention to the areas that are relevant to the visualizer. Hence, a straightforward idea to measure the quality would be to sum squared differences over all pixels $p$. This measure, however, has only little informative value, as small variations in large insignificant regions often dominate the results. Hence, we propose a more detailed visual result that accounts for the different relevance classes and allows for an easy to analyze quality metric.

The idea is to depict for each importance class, how well visual salience matches data relevance. Accordingly, the first step is the computation of the difference field between relevance mask and salience map. Negative values indicate that a structure with little relevance attracts a lot of visual attention and positive values that relevant structures obtain too little visual attention. In a second step, the differences are subdivided into three categories according to the desired relevance and sorted for each category individually. The resulting data can be represented using a bar chart that can be compared to the distribution in the relevance mask and salience map.

We found that a continuous representation, as induced by the continuous salience values, is rather difficult to interpret. Hence, we discretized the salience map according to the user defined relevance into three classes and obtain the color scheme as given in Figure 2(b). A sample quality chart with additional annotations is given in Figure 2(c). The x-axis encodes the number of pixels per image and the two lower bars indicate how many pixels were assigned to the three importance classes by the salience map and the user respectively. The upper bar depicts the matching between the two classes and can be thought of as a rearranged difference field as given in Figure 2(a).

The most important quantity is the matching of the high relevance structures (left section in the chart), as all relevant features shall gain visual attention. In the ideal case of a perfect visualization, the number of relevant pixels as defined by the user and the salience map would match exactly (same length of lower dark blue bars) and all pixels were classified correctly (“match” bar section entirely white). A secondary criterion is the matching of the irrelevant structures (rightmost section of the chart), which indicates how many irrelevant areas attract too much attention.

3. Analysis of Visual Attention

The quality metric as described in the previous section allows the user only to see whether a visualization is good or bad given the specified purpose. However, we want to allow for a more detailed analysis and an iterative improvement of the initial visualization. Therefore, the user needs information on where to improve the visual relevance. We propose two different visualizations for this purpose: The first representation is an overlay of the original visualization and the quantized salience map to indicate salient structures, and the second one depicts the contributions of the visual channels.

3.1. Highlighting of Salient Structures

To highlight salient structures in the original visualization, we have to merge two images, the visualization and the salience map. We found direct overlays combined with isoline representations difficult to read and therefore, chose to decrease the visual impact of the original visualization and use it solely as a reference frame. Hence, we transform the visualization to a gray scale image and manipulate the color distribution to make dark edges more pronounced and to attenuate large areas. The resulting gray scale visualization is then overlayed with the salience map, and isolines are added to distinguish between the three major salience classes as defined earlier.

3.2. Contribution of Individual Channels

To allow for the analysis of the contribution of the visual channels to the final salience map, we provide a contribution map. The contribution map in Figure 3(c) shows where each of the three channels contributes strongly. The thresholds can be interactively manipulated by the user. We commonly set the isovalue for all channels to 85 (range 0–255) to reflect the earlier partitioning into three relevance classes. Red areas indicate strong responses by the color opponency
4. Test Cases

We applied our quality measure to visualizations from three major fields in visualization: TagClouds from information visualization, Volume Ray Casting (VRC) from volume visualization, and methods from flow visualization. With each application, we concentrated on a different aspect of visual quality analysis: What attracts attention in a visualization? (TagCloud), Which visualization for which purpose? (VRC), and Which visualization is best? (Flow Visualization).

4.1. Information Visualization – Tag Clouds

A tag cloud is a visual representation of a set of words providing information about the relevance of the tags. The larger the size of a word, the more important it is within the given context, e.g., a web page or blog. Color is often used in an artistic sense to make the image visually more appealing.

The first example in Figure 3(a) shows the tag cloud displayed on the InfoVis wiki homepage (www.infovis-wiki.net). The words in the tag cloud are colored in red, black, sand color and shades of gray. Judging from the size of the words, the most relevant terms are information, visualization, data, and research. However, when looking at the salience map in Figure 3(b), we see that the visual attention does not well agree with the size of the words and smaller words in intense colors appear very prominent. The contribution map in Figure 3(c) reveals why small words become visually so important. Each color indicates areas that are conspicuous according to one of the channels color opponency (red), intensity change (gray), and orientation (blue). The more channels are active, the higher the visual salience. In general the contribution map reveals a very inhomogeneous distribution of conspicuity. We see that the black words are conspicuous with respect to intensity change and orientation, i.e., on a fine scale we have strong contrast and more vertical orientation of parts of the letters and on a coarse scale lower contrast and horizontal orientation of the entire word. The high salience of the red words results from the strong color salience combined with changes in orientation again. In general we can say that the different channels compete strongly for visual attention and the choice of color does not support the intended representation of relevance of words.

A better choice of colors was made in the example in Figure 3(d). The contribution map (Figure 3(f)) shows a much more homogeneous pattern and the different channels support each other. Hence, largest words attract most attention. To improve a visualization with respect to the choice.
of colors, different color schemes have to be evaluated for their visual salience and depending on the application the color scheme that best fits the relevance has to be applied to achieve best salience patterns.

4.2. Volume Visualization – Volume Ray Casting

Volume rendering is a technique to compute a 2D projection of a three-dimensional data set that allows for a layered representation of three dimensional intricate structures. A common problem is to find a suitable transfer function that provides a good characterization of the different, partly overlapping structures in the 3D volume.

The image in Figure 4(a) shows a direct volume rendering (DVR) of a CT engine data set. We can observe that the outer surface is very easy to perceive, while the internal structures and tunnels through the object are barely visible. In [CWM*09], Chan et al. proposed perception-based transparency optimization to increase the perception of semi-transparent layers. The visualization of the same data set using their new algorithm is given in Figure 4(b). As proposed by the authors, the different layers of structures inside the data set are easy to perceive. However, it is rather difficult to relate the structures to each other and to observe depth clues.

Figures 4(c) and 4(d) show the respective salience maps of the two techniques. With the standard technique (Figure 4(c)), most attention is focused on the outline and surface of the engine and partially on the red structure in the center. In many areas multiple conspicuity channels are activated (Figure 4(e)). The attention to the outline of the structure is induced by all three attention channels. The central part, on the contrary, owes its visual attraction solely to the choice of red color. The perception-based algorithm (Figure 4(d)) features a more distributed salience map with smaller peaks that guide the observer’s attention to many different structures of the data set. This distribution is more easily visible in the contribution map in Figure 4(f). Like in the DVR contribution map, several structures exist that are conspicuous in multiple channels. The structures, however, are less homogeneous and more widely distributed over the image.

The salience analysis supports the first impression. With the standard DVR visualization, a lot of visual attention is directed to the outline of the engine representation and the overall impression is more homogeneous due to coherent conspicuous structures. The perception-based representation on the contrary guides the user to many different structures in the data set and makes the observer focus more on details.

Depending on the task, either of the two visualizations can be favorable. While the standard techniques provides a better impression of the overall structure of the engine, the perception-based algorithm disperses attention over the entire data set and makes internal structures easier to see.

4.3. Scientific Visualization – Flow Visualization

Most available visualization software tools provide a large variety of algorithms and in many application areas the same data set can be represented in several ways, even if this only means a change in the color scheme. The question that arises is which technique is suited best to represent the data. To find an answer the user commonly has to try different techniques and compare them or base the decision on experience. User studies are a useful tool to find typical characteristics of different techniques and compare them, but do not capture all the subtle details and differences that occur in daily life. Hence, we aim in this test case at a meaningful and simple way to compare the quality of different visualizations based on visual salience.
Figure 5: Flow visualizations: (top) The same field is depicted using different techniques (req courtesy of D. Laidlaw [LKJ*05]). (2nd row) Salient structures are colored in blue. (3rd row) Contribution maps, red for intensity-band, grey for intensity and blue for orientation. (bottom) Difference between the salience map and the relevance mask. Color coding as given in Figure 2(b).

The visualizations in Figure 5(top) depict different flow visualizations [LKJ*05] of the same data set. From left to right, we see a grid-based representation of the vector data, a depiction where the vectors have been jittered, a visualization using scaled triangles, a representation with image-guided flow fish and a line integral convolution (LIC) of the vector field.

The images in the second row depict the original image along with an overlay with the salience map. As all visualizations are gray-scale images, the color opponency channel in the salience computation was substituted by the intensity-band channel [IKN98]. We see that the first three visualizations attract visual attention in approximately the same areas. The flow fish and LIC representation feature a more irregular pattern, with only few areas of high salience and large unconnected areas of moderate salience.

The contribution maps of Figure 5(a,b) look quite similar with all three channels agreeing in most areas. In the scaled triangles visualization, the orientation contrast channel is activated in most areas where large triangles are clearly visible. The discontinuities in the representation, e.g., close to the critical points and separatrices, are picked up in the intensity-band and intensity channel. The contribution maps of the flow fish and LIC visualization show that the dense depiction of local directions induces only little visual attraction as the overall structure of the entire image is very homogeneous. Attention is only attracted by spurious changes in the intensity that are often artefacts of the algorithm.

Figure 6(a) shows the important features that the visualizer wants to show, and Figure 6(b) the corresponding relevance mask. The difference fields between the relevance mask and the salience maps are displayed in the bottom row.
Figure 6: Relevant structures in the flow visualization data set. (a) Depiction of the interesting features. (b) Relevance mask as designed by the visualizer.

of Figure 5. As these images are rather difficult to interpret, we reorganize the data and display it using the quality chart. The results are given in Figure 7. Good performance with respect to drawing attention to the features is achieved by the “scaled triangles” and “grid-based vectors” visualization, where a large subset of the highly important section (left part of the bars) is colored white, i.e., correct amount of visual attention. The “scaled triangles” visualization features much smaller and more focused high salience areas, whereas these areas are much larger in the “grid-based vectors” visualization. Worst quality according to visual salience have “LIC” and “jittered vectors” as all regions are about equally visually attractive. Hence, the user has to scan the entire image – which in this case is not too difficult, as the picture is very small – in order to detect relevant structures.

5. Discussions

In the previous section, we explored the application of our salience-based quality metric using test cases from three major fields in visualization. In general, the results gave a very good indication on which part of a visualization needs improving or which one to choose for a given purpose. In this section, we compare our results with earlier findings, discuss shortcomings, and outline the use of the metric in practice.

The first question that arises is, does the salience map reflect the distribution of human attention? The quality of the salience map model was evaluated in several studies. In [IK99], Itti shows that the model can be used to identify suspicious objects in complex scenes. In [BI05], Baldi and Itti found that the human gaze is attracted to surprising areas which is captured by the model as well and the study in [BI08] reveals that the saliency map can be used to predict eye location. Our results are consistent with these findings.

In the tag cloud example in Section 4.1, we measured the quality of different color selections. This analysis, however, is purely based on salience, and we do not take personal preferences about color or esthetics into account. Many people might find the first image (InfoVis wiki) visually more interesting or pleasing. From an information communication point of view, the second example features the better choice of color. A combined measure of visual salience and esthetics might further improve the analysis to create images that attract correct attention and are visually pleasing.

The flow visualizations investigated in Section 4.3 are part of a user study conducted by Laidlaw et al. [LKJ∗05]. In this study, the subjects of the study had to accomplish different feature location, definition and tracking tasks. They found that the methods have different strengths and weaknesses and none is optimal for all tasks. Best results were achieved with techniques that depict directed integral curves and critical points, e.g., the image-guided flow fish technique. Our analysis, however, rated the “scaled triangles” and “grid-based vectors” best, while LIC and “jittered vectors” feature a poor salience profile. These seemingly contradicting results have an easy explanation. While our quality metric picks up areas that stand out within the image, Laidlaw et al. measured performance according to time and error. The images are relatively easy and can be easily scanned as a whole by a human. Thus, the subjects could quickly spot the
relevant positions and took most of the time to perform the more challenging task, e.g., determining the exact location of a critical point. As features are found easily even if they do not stand out, visual salience has hardly any effect except a little influence on the timing. Local precision, the second performance measure in the user study, cannot be captured by our salience-based quality metric and is an interesting direction for future research.

The second question that arises is, how can one use this quality metric in practice? In the previous sections, we mentioned several different maps, namely salience map, contribution map, and relevance map. The first two maps are generated directly from a visualization by using the method discussed in Section 3. The relevance map can be created automatically from the data or manually by the visualizer as discussed in Section 2.1. Based on these, we can envisage different levels of support from a visualization system for a salience-based quality metric. At the basic level (level 1), a system provides a utility for computing salience and contribution maps, allowing users to evaluate their created visualizations visually in a manner similar to Figures 4(b,c,e,f), 5(e,f) and 6(rows 2 and 3). At level 2, a system provides users with a utility to create a relevance map by painting over a visualization. Alternatively, the system allows the users to import an image of a relevance map painted by a third-party image editing tool. With a relevance map, a visualizer can evaluate a visualization using a difference field as exemplified by Figure 3 and Figure 6(row 4). At level 3, a system provides a utility to generate a relevance map using a data-space algorithm (e.g., feature detection, uncertainty estimation, etc.). Such a utility is normally data-type dependent, and often application dependent.

With these three levels of support, we can also envisage different scenarios, where the quality metric is used. Two typical scenarios are given below:

**Scenario A: Evaluating Techniques.** For exploratory visualization, a user is a visualizer as well as an observer, and may face a collection of optional technique and a large parameter space for each technique (e.g., opacity and color transfer functions, visual mapping, etc.). It would take a huge effort to explore all such options on a daily basis, while the priority should be given to exploring a large data space through visualization. Hence, the user can determine a default setting which would be most effective for a visualization process as follows. (1) Take a typical example data set, create a visualization using a method, obtain a relevance map with the abovementioned level 2 or 3 support. (2) Experiment with different methods and parameters, and apply the salience map, contribution map, and difference field to each of the created visualizations. Determine the most effective method and parameter set. (3) Using the selected method and parameter set as the default setting to explore other data sets.

**Scenario B: Evaluating Visualizations.** In this case, a user is a visualizer and may need to determine how effective the visualization created will be in conveying information to the observers, and whether they would mislead observers. The visualizer can apply the salience and contribution maps to evaluate each created visualization in an objective and visual manner. In this case, the creation of the relevance map is not essential, as the visualizer usually has a mental image of such a map, and can mentally overlay this mental image with a salience or contribution map.

In both Scenarios, the use of this quality metric also provides users with opportunities to accumulate understanding and experience about different methods and associated parameter space from an information theory and perception perspective. From the perspective of quality assurance, Scenario A focuses on the tools used in a visualization process, while Scenario B focuses on the “product” of the process. Like most quality assurance processes, it is unlikely that a single quality metric can meet all the requirements. Hence using a combination of quality metrics is highly desirable.

### 6. Conclusions

In this paper, we presented a novel quality metric for visualization based on visual salience. This metric favors visualizations that guide the observers visual attention to relevant parts of the image and is generically applicable to all areas of visualization. Using different analysis tools, i.e., salience overlays, contribution maps and quality charts, we could show how our quality metric can be used to choose the best visualization from a set of alternatives, and to establish an effective default setting for a visualization process.

In most cases visualizers have a mental image of what they want to show to the observer with their visualization. So far the common technique to assess the match between expectation and actual results were user studies that are very time consuming. With our technique the visualizer is provided a fast and easy to apply tool that tells how well the visualization meets the expectations.

The most important aspect of our quality-metric is that the technique can be easily integrated into existing visualization software. Within these frameworks it is a valuable addition not only for visualization experts but also for novices that obtain easy to understand feedback on the quality of the visualization they created.

As we have seen in the discussion, there are several directions for an improved quality metric for visualization. Not only correctness of the visualization has to be ensured (abstraction measures), but also task performance (user studies) and esthetics (esthetics measure). Although we recognize that it is unrealistic to expect a single quality metric to adequately serve all needs, combining different quality measures might result in a more comprehensive feedback on the quality of a visualization that helps create correct, meaningful, easy to understand and esthetic visualizations.
References


