Evaluating Sketchy Lines for the Visualization of Qualitative Uncertainty

Nadia Boukhelifa, Anastasia Bezerianos, Tobias Isenberg, and Jean-Daniel Fekete

Evaluating Sketchy Lines
for the Visualization of Qualitative Uncertainty

Nadia Boukhelifa[*], Anastasia Bezerianos[†], Tobias Isenberg[‡], and
Jean-Daniel Fekete[§]

Project-Teams AVIZ

Abstract: We report on results of a series of user studies on the perception of visual variables that are commonly used in the literature to depict uncertainty. To the best of our knowledge, we provide the first formal evaluation of the use of these variables to facilitate an easier reading of uncertainty in visualizations that rely on line graphical primitives. In addition to blur, dashing and grayscale, we investigate the use of 'sketchiness' as a visual variable because it conveys visual impreciseness that may be associated with data quality. Inspired by work in non-photorealistic rendering and by the features of hand-drawn lines, we generate line trajectories that resemble hand-drawn strokes of various levels of proficiency—ranging from child to adult strokes—where the amount of perturbations in the line corresponds to the level of uncertainty in the data. Our results show that sketchiness for the visualization of uncertainty in lines is as intuitive as blur; although people subjectively prefer dashing style over blur, grayscale and sketchiness. We discuss advantages and limitations of each technique and conclude with design considerations on how to deploy these visual variables to effectively depict various levels of uncertainty for line marks.

Key-words: Uncertainty visualization, qualitative evaluation, quantitative evaluation, perception

[*] Nadia Boukhelifa is with INRIA, France; email: nadia.boukhelifa@inria.fr
[†] Anastasia Bezerianos is with LRI (Univ Paris-Sud & CNRS) and INRIA, France; email: anastasia.bezerianos@lri.fr.
[‡] Tobias Isenberg is with the University of Groningen, the Netherlands, and with DIGITEO/CNRS/INRIA, France; e-mail: isenberg@cs.rug.nl.
[§] Jean-Daniel Fekete is with INRIA, France; email: Jean-Daniel.Fekete@inria.fr.
[¶] Minor editorial changes.

Evaluation de lignes dessinées á main levée
pour la visualisation d'incertitude qualitative

Résumé : Nous rapportons les résultats d'études utilisateurs sur la perception de variables visuelles qui sont couramment utilisées dans la littérature pour représenter l'incertitude. A notre connaissance, nous rapportons la première évaluation formelle de l'utilisation de ces variables pour faciliter la lecture de l'incertitude dans des visualisations qui reposent sur des lignes graphiques primitives. En plus du flouté, des pointillés et des niveaux de gris, nous avons étudié l'utilisation du dessin á main levée comme variable visuelle parce qu'il transmet l'imprécision qui peut être associée á la qualité des données. Inspirés par les travaux sur NPR (Non-Photorealistic Rendering) et par l'aspect des lignes tracées á la main, nous avons généré des trajectoires linéaires ressemblant á des traits dessinés á la main, et correspondant á différents niveaux de compétence—allant de l'enfant á l'adulte—pour lesquels la quantité de perturbations dans la ligne correspondait au niveau d'incertitude des données. Nos résultats montrent que l'utilisation du dessin á main levée pour la visualisation de l'incertitude dans les lignes est aussi intuitive que le flouté, bien qu'intuitivement les personnes préfèrent subjectivement les pointillés au flou, aux niveaux de gris ainsi qu'au dessin á main levée. Nous discutons les avantages et les limites de chaque technique et nous concluons sur des considérations de conception, concernant comment déployer ces variables visuelles pour représenter efficacement plusieurs niveaux d'incertitude pour des repères linéaires.

Mots-clés : Visualisation de l'incertitude, évaluation qualitative, évaluation quantitative, perception

1   Introduction

Information visualization can show not only what we know about the data but also the degree of our confidence in that data. This confidence could be considered as yet another data dimension. However, information on data quality in general—if at all available—rarely comes in numerical format. Qualitative measures of uncertainty are far more common and often come as ordinal meta-data. For instance, a utility company holds a positional confidence attribute for their assets where uncertainty is mapped to five categorical values from least certain to more certain: schematic, assumed, indicative, third party survey, and internal survey [7].

Ordinal data can be visualized using Bertin's [4] retinal variables texture, value, or size. When visualizing uncertainty, a number of visual variables are considered more 'intuitive' for this domain; examples include blur, sharpness of focus, and color saturation. These variables may bear direct perceptual resemblance to what the uncertainty indicates and, thus, may provide an easier reading of uncertainty [9, 29, 41]. However, to our knowledge, no formal studies are reported to back up these observations.

In this paper, we investigate *sketchiness* as a visual variable to depict uncertainty information in line marks such as for graphs, hierarchies and route maps. The design of these sketchy lines is inspired by the field of Non-Photorealistic Rendering (NPR) and by our own observations from analyzing child and adult hand drawings. Our analogy is that the 'cleanness' of the hand-drawn lines corresponds to the quality of the data. We hypothesize that *sketchiness* is a good metaphor for the qualitative visualization of uncertainty information.

Our contribution is two-fold: (1) a empirically-based method to generate line trajectories that resemble hand-drawn strokes of various levels of proficiency—ranging from child to adult strokes—where the amount of deviations in the line corresponds to the level of uncertainty in the data; (2) a qualitative and quantitative evaluation of four uncertainty visualization techniques using blur, dashing, grayscale, and sketchiness. In particular, we attempt to answer the following questions: (a) can people intuitively associate sketchiness to uncertainty, (b) is sketchiness as effective for depicting uncertain information as the other visualization techniques, and (c) which method do people subjectively prefer.

The remainder of the paper is organized as follows. After discussing previous work, we first present our model for generating sketchy lines, and then present a series of studies that answer our questions on intuitiveness, accuracy of reading uncertain visualizations, and subjective viewer preference in regards to the different uncertainty visualizations.

2   Related Work

The previous work that relates to our own can generally be thought of as threefold: (a) work in non-photorealistic rendering that studies the generation of primitives that appear sketchy or hand-drawn as well as their applications, (b) methods to depict uncertainty data, and (c) user studies that examine how people perceive and interpret uncertain data depictions. We discuss each of these three fields in turn.

2.1   Sketchy Lines in Non-Photorealistic Rendering

The field of non-photorealistic rendering (NPR) [18, 44] has been inspired by the many ways of traditional depiction that humans have employed over the last decades, centuries, and millennia. As such, computer-generated line drawings have been among the first goals for NPR. Early-on, for example, loose and sketchy line rendering and animation [11] have been simulated. Researchers have also developed ways to represent lines such that the line path can be separated from the (sketchy) line perturbation [13, 37]. These line models can, in turn, be applied to line renderers (e. g., [42, 48]) to generate images that resemble—to varying degrees—traditional hand-drawn ones. More recently, NPR line models have been extended to

be more example-based, e. g., by taking the human arm movement into account in their generation [3]—a goal that we share for generating our sketchy lines.

In addition to reproducing a generic hand-drawn look, sketchy non-photorealistic rendering has also specifically been employed to portray uncertainty. For example, Strothotte et al. [41] used a general level of sketchiness to indicate a general notion of uncertainly in the domain of archeology visualization. They also describe examples that continuously change the line thickness/line saturation or the degree of perturbation in order to visualize continuously changing degrees of uncertainty. In the same domain, Potter et al. [34] perturb the vertices of a line-based rendering style to control the degree of sketchiness of the depiction, which in turn is used to visually express the level of confidence in a reconstruction. A similar notion of general sketchiness to indicate uncertainty was used by Nienhaus and Döllner [31] for the visualization of 3D shape concepts in CAD—supported by earlier findings of Schumann et al. [38]. Rather than these implicit notions of uncertainty visualization, we are interested in a more precise analysis of how line sketchiness can be employed for qualitatively depicting a number of different levels of uncertainty, and thus in how sketchiness can be used intentionally as a dedicated visual variable.

2.2   Visual Variables to Depict Uncertainty

Many visual variables have been proposed for the depiction of uncertainty including Bertin's [4] retinal variables position, size, and value. However, only a subset of these variables can be considered to be 'intuitive' for visualization purposes. Most methods that are widely used to depict uncertainty and that are applicable to line primitives can be grouped into three main categories: (1) color-based techniques that manipulate hue, saturation, or brightness dimensions; (2) focus-based techniques that modify contour crispness, transparency, or resolution; and (3) geometry-based techniques that distort line marks by applying a rendering style as in sketchiness.

Color is repeatedly used to depict uncertainty in information visualization [44, 1, 12]. Since color scales are readily available in today's graphical packages and libraries, assigning a range of a color scale to a range in the data is straightforward. Amongst all color dimensions, saturation is often preferred because the reduction of color purity conveys more intuitively the notion of degrading data quality or confidence [29, 41]. Hue is also used [35, 49, 50] and people often use a rainbow scale for ordinal uncertainty data with up to seven levels. For value, darker lines suggest more certainty about an aspect of the underlying data. In this case, however, the line width may need to be adjusted to preserve perceptibility [9]. More advanced color mapping techniques are also deployed including whitening where white pixels are randomly placed or actual hue and white are blended [22, 23, 33].

Amongst focus-based techniques, blur, which is defined as the removal of high-frequency spatial detail from the information [8] has widely been used to indicate fuzziness and ambiguity in the data [6, 15, 29]. For example, Bisantz et al. [5] applied blur to a set of airplane symbols to provide decision makers with a fast way to understand the level and uncertainty of a given threat. Gershon [16] used blurred versions of images in an animation loop from focused to blurry (or the inverse) to draw users' attention to uncertain objects. Kosara et al. [24] described their 'semantic depth of field' technique that uses blur to de-emphasize objects of less importance in the scene. We use blur as one visual variable in our evaluation and compare it to others like sketchiness, dashing, and value.

Modifying the geometry of line marks can be a powerful way to convey uncertainty. For instance, Griethe and Schumann [19] argue that "wavy or dotted lines could convey less trusted relationships." Similarly, Strothotte et al. [43] show how sketch-like renditions (and the use of transparency) can express the uncertainty in archaeological reconstructions. The domain of NPR also proposed many geometry-based rendering techniques to convey uncertainty as described above. Using drawing primitives from NPR, Pang [32] used gaps in contour lines of geographic maps to encode uncertainty such that larger gaps in the contour line encode an increased uncertainty. In the related area of oceanography, Osorio et al. [2] also augmented contour visualizations using uncertainty bands that indicate different possible

locations of a contour line. Luboschik et al. [28] suggested the use of dashed and wavy lines to show uncertainty in parallel coordinates. In the context of maritime situational awareness, Matthews et al. [30] depict the timeliness and quality of sensor information using icon borders—solid or broken. More relevant to our work, however, is the distorted annotation technique by Cedilnik and Rheingans [9] who distort grid lines proportionally to the amount of uncertainty in the data. Similar to our approach for sketchiness, Cedilnik and Rheingans map the amplitude of line distortion to the amount of uncertainty in the data. Our method, however, does not separate the data from the uncertainty depiction because the affected lines themselves are data carriers in our case.

2.3 Perceptual Studies and User Evaluations

To our knowledge, there has not been a formal comparative evaluation of the use of the mentioned visual variables to facilitate an easier reading of uncertainty in visualizations that rely on line primitives. However, several perceptual studies and user evaluations have examined the general application of those visual variables in visualizations.

It is established that color is a powerful dimension to indicate data quality in general [25, 29]. Specifically for the context of multidimensional data visualization, Xie et al. [49] found that hue has stronger capacity to convey quality information for parallel coordinates than brightness or line width—even for large datasets. Color, however, may suffer from the lack of an intuitive order. MacEachren [29] showed, for example, that subjects cannot spontaneously order colors into a legend arrangement for bi-variate choropleth maps but that they can recognize order in that arrangement. The question of user preference is also pertinent to the problem of uncertainty visualization; in a user opinion survey conducted by Gerharz et al. [15] in the context of geographical information systems, people disliked whitening [23] to convey uncertainty. The authors argue that the principle of whitening is easy to understand but getting detail information from it is difficult. Li et al. [27] investigated the issue of scale for uncertainty visualization for astrophysical data and used a unified color scheme to represent log-scale distances and percentage errors. They found that participants were able to determine the amount of uncertainty using colors with 96.7% success rate. Similarly to the findings by Gerharz et al. [15], however, access to detail was difficult especially for neighboring color ranges.

Blur as an example of a focus-based uncertainty visualization technique is a well-studied visual variable in domains that go beyond information visualization (e. g., [20, 46]). Kosara et al.'s [24] evaluation of the previously mentioned 'semantic depth of field' specifically examined people's ability to read absolute blur levels. Their results show that participants can distinguish between different blur levels with good accuracy but cannot quantify this difference nor identify objects of the same blurriness. Kosara et al. thus concluded that blur can guide attention but is hard to quantify and thus may not be recommended for showing *quantitative* uncertainty. Moreover, Kosara et al. found that participants disliked looking at blurred objects. These results motivate our work in which we examine whether blur is effective in communicating *qualitative* uncertainty information.

Few quantitative evaluations exist for geometry-based methods. One exception is the work by Matthews et al. [30] who found that—in visual search tasks on maps where latency of information is indicated either by color hue (green or gold) or border style (solid or broken)—search time was faster with border style than with the color format.

More generally, the choice of which uncertainty visualization method to pick may be task-dependent [36]. Evaluating uncertainty depiction techniques for information visualization application in general is thus challenging. For example, MacEachren [29] states that "for exploratory applications, where there is no predetermined message to communicate, we can not judge uncertainty depictions using communication effectiveness standards. We can only evaluate these depictions in terms of how they alter decision making, pattern recognition, hypothesis generation, or policy decisions." Zuk and Carpendale [51] presented a set of heuristics for uncertainty visualization evaluations and stressed the need for more research
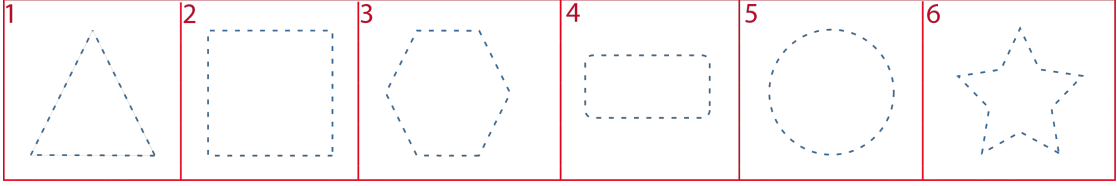
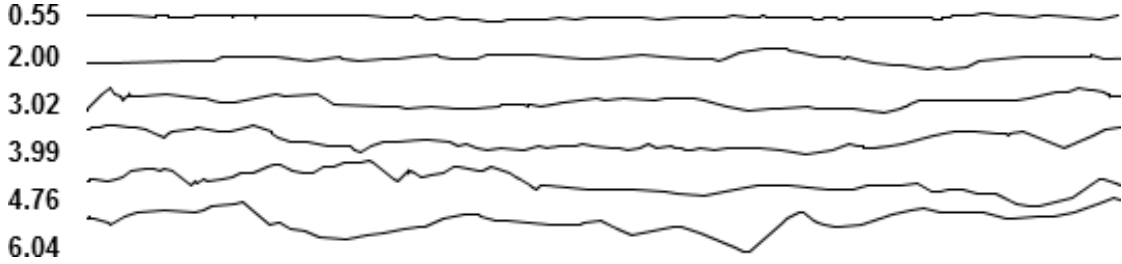Figure 1: Templates of figures for participants to trace.

Figure 2: Examples of hand-drawn strokes which we ranked by their average deviation (in mm) from straight lines.

in human factors and perception. We are inspired by this call for action and conduct a comparative study between four popular visual variables for uncertainty visualization including methods from the three described categories. For color-based methods we use the visual variable *value* (i. e., grayscale), for focus-based techniques we examine *blur*, and for geometry-based techniques we selected *dashing* and *sketchiness*. To enable this evaluation we first describe in the next section how to mimic hand-drawn trajectories to be able to create sketchy lines, before detailing our study.

3   Mimicking Hand-Drawn Trajectories

To to be able to synthesize lines that mimic hand-drawn ones we first studied characteristics of hand-drawn strokes collected using an AnotoPen. We asked 20 people whose ages ranged from 5 to 47 years (mean of 26) to trace six different shapes (Figure 1); each participant first drew with their dominant and later with their non-dominant hand. In total, we collected 204 hand-drawn shapes. With the exception of one person, none of our volunteers had a formal drawing training experience.

Our hypothesis was that drawing proficiency can be determined by examining the average deviations of the drawn path from the template path. A quick visual inspection showed that, indeed, the subjectively more proficient-looking drawings deviated less from the template shapes. Figure 2 shows examples of unfolded hand-drawn shapes, ranked by their average deviation from the target shapes. Within the set of hand-drawn strokes we collected, the average deviations per stroke ranged from 0.55 mm to 6.04 mm. An error distribution analysis of all stroke control point deviations (signed) showed a normal distribution with a mean of 0.36 mm and a standard deviation of 2.08 mm (min -13 mm; max 18 mm). We decided to use average range of deviations from a straight line to map levels of uncertainty in data.

Mimicking hand-drawn lines, of course, requires a generative model of hand movement. Flash and Hogan [14] described such a model for straight lines, based on minimizing jerk during a stroke from $p_0 = [x_0, y_0]$ to $p_1 = [x_1, y_1]$, jerk being the derivative of acceleration. At time t, the model (improved by AlMeraj et al. [3]) generates a point:

$$
\begin{aligned}
x(t) &= x_0 + (x_0 - x_1)(15\tau^4 - 6\tau^5 - 10\tau^3) + D \\
y(t) &= y_0 + (y_0 - y_1)(15\tau^4 - 6\tau^5 - 10\tau^3) + D
\end{aligned}
\tag{1}
$$

with $\tau = \frac{t}{t_f}$, t_f being the time of the end of the stroke, and D a random value in a specified (pixel) range.

AlMeraj et al. [3] empirically defined the time sampling parameter δt according to the length of the desired line: $\delta t = 0.5s$ for lines shorter than 200 pixels, $\delta t = 0.3s$ for lines of 200–400 pixels, and $\delta t = 0.2s$ for lines longer than 400 pixels. They also fixed $t_f = 2s$ based on empirical evidence. Using these parameters in Equation 1 yields multiple points that are connected using a smoothing spline. While AlMeraj et al. generated noise in a fixed range of $D \in [-5, 5]$ pixels, we vary the range of the noise to produce more or less 'sketchiness' using a normal random generator, based on our previous observations: we vary D, our sketchiness parameter to produce deviations within $[-20, 20]$ pixels values which correspond roughly to the average stroke deviation range we observed. More details on how we produce our sketchy lines are described in the next section.

3.1 Generating Levels for *sketchiness*

Our method to produce sketchy lines takes a number of control points generated by the Flash and Hogan model and a value for D, our sketchiness parameter, to produce deviations within $[-20, 20]$ pixels. Since we are essentially adding noise to straight lines to produce sketchy lines that are consistently perceived by viewers as belonging to the same level, we make sure that the mean deviation of the generated sketchy lines stays faithful to the input D value.

To achieve this, we sample values from a normal distribution where the mean is set to D and the standard deviation is set empirically to $D/6$. For our studies, we produced sketchy lines where the difference between D value and actual mean is less than 0.1 pixels. This approach can be described as being *mean-biased*.

An alternative strategy to generate levels consistently, described as *max-biased*, is to use the maximum deviation. We sample values for the D parameter from a normal distribution where the mean is equal to 0 and the standard deviation is equal to D. In addition, we add a constraint whereby D values outside $[-20, 20]$ pixels are not allowed.

Figure 3 shows 5 major levels as generated by the two methods. As D grows, lines under the mean-biased condition get more undulated in order to preserve the overall mean value, whereas sketchy lines overall have a more flat appearance with the exception of a few peaks. In a pilot study we compared the two variations and found that the *max-biased* variation resulted in smaller errors in the perception of the sketchiness level than the *mean-biased* variation, and thus we used the former in our studies.

3.1.1 Assessing Generated Sketchiness Quality

For verification, we conducted an online study to see if people can tell the generated lines (using the max-biased approach) apart from hand-drawn ones. The details for our study setup can be found in subsection 4.2.

We selected 45 samples from our library of hand-drawn strokes of three shapes (hexagons, rounded rectangles, and triangles) from a deviation range between 0.9 mm and 4.9 mm (within the range we observed in real strokes), and generated the same shapes using the generative model.

We varied D from 0.4 pixels to 10.0 pixels to obtain different sketchiness levels. Our hypothesis was that people cannot tell computer- from human-generated shapes apart for D values of up to 6 pixels (close to the average deviation we observed in hand-drawn shapes). This hypothesis is based on our own subjective observations that shapes start to look synthetic at $D \approx 6$, because for larger D values
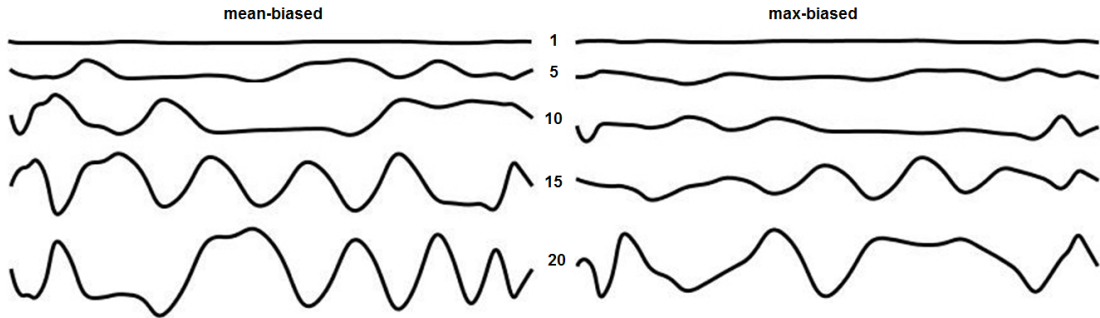
Figure 3: Ranges of *sketchiness* used in the study for the mean and max biased variants.

corners start to loose their shape and knots start to appear, something that is not associated with hand-drawn strokes. Thus the null hypothesis was that, for a given sketchiness *D*, participants will be unable to distinguish human-made from machine-generated strokes (i. e., selecting the right answer 50% of the time).

The results of the online study show that, overall, 60% of the computer-generated shapes were identified as hand-drawn ones, and 71% of the real hand-drawn shapes were correctly identified as such. A two-tailed non-parametric Bionomial test rejected the null hypothesis ($p < .0001$), showing that participants were significantly more likely to think the computer generated lines were human-made (thus the effect was not due to chance). Nevertheless, we did not find an effect of sketchiness level, indicating that regardless of *D* (i. e., even with extreme sketchiness) participants were still more likely to perceive the shapes generated by our algorithm as hand-drawn.

4   User Studies: Design Goals and Requirements

Using our results from mimicking sketchy lines based on empirically determined parameters, we can now pursue our main objective: investigate whether sketchiness can be used as a visual variable to encode qualitative uncertainty data in information visualization. For this purpose we compared sketchiness to popular uncertainty visualization techniques with respect to three criteria: *intuitiveness*, *accuracy*, and *subjective user preference*. For intuitiveness, we wanted to investigate if and to what extent do people associate sketchiness (and other techniques) to uncertainty. For accuracy, we wanted to investigate how accurately people can read values from uncertain line marks (characterized by the use of sketchiness, dashing, grayscale, and blur) and the number of distinctly perceivable levels for each technique. Finally, we wanted to study which technique would be preferred by participants to encode uncertainty.

4.1   Generating Levels for *blur*, *dash*, and *grayscale*

All our blurred, dashed and grayscale lines were generated using the GNU Image Manipulation Program (GIMP). We generated 20 levels per visual variable; each line was 400 pixels long, 3 pixels wide, and used black except for the grayscale case. Figure 4 shows the minimum and maximum levels for each visual variable. For blur, we used a Gaussian blur filter where we varied the blur radius horizontally and vertically in equal amounts ranging from 1 to 20 pixels.

For dashing, we selected a style where the dash/gap length grows proportionally with each consecutive level, so the minimum dash/gap size was set to 3 pixels and the maximum to 60 pixels.

For grayscale, we partitioned a linear grayscale from 4% up to value 90% into 20 discrete levels from black to white, since an increase in whiteness may be seen as an increase in uncertainty in the data
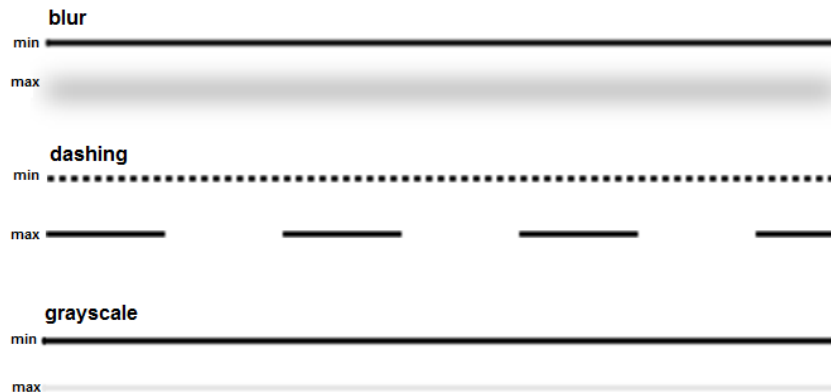
Figure 4: Ranges of three visual variables used in our study.

[23]. We set up the maximum grayscale level (i. e., white) in line with findings for the Just Noticeable Differences for grayscales by Levkowitz [26], but we set the minimum level (i. e., black) to 4% instead of the 60% as recommended by Levkowitz to get the same number of levels as for the other visual variables.

As mentioned in the previous section, the 20 levels for *sketchiness* were produced using *max-biased* generation, with D set to $[-20, 20]$. Figure 3 shows sample generated lines used in our studies.

4.2 Setup of the Studies

Our studies were conducted on Amazon Mechanical Turk (AMT), inspired by previous graphical perception experiments [21]. In total, we deployed six experiments and participants that participated in one our studies could not take part in any of the other ones. We had 1136 participants in total, and subjects were paid on average $ 0.25 per Human Intelligence Task (HIT).

5 On the Question of Intuitiveness

The question of whether people intuitively associate sketchiness with uncertainty is pertinent to our evaluation of sketchiness as a visual variable. We define *intuitiveness* as the spontaneous association between signifier and signified: signifier being sketchiness and signified being uncertainty. If we find that there exists an intuitive association between the two, we could hypothesize that people can spontaneously associate sketchiness with uncertainty and do not need to consult a legend to identify the meaning of the visual variable (note that at this stage we are interested in the meaning associated with the representation - that it is uncertain data-, not its perceived magnitude - i. e. how uncertain it is). We, thus, conduced a study to examine the intuitiveness of sketchiness as a visual variable for uncertainty.

This study was conducted in three parts. Each participant only completed one of the 3 parts, and each part was treated as a between-subjects experiment.

In Part I participants were shown visualization scenarios that included sketchiness (described in the next section), and were asked in an open ended question to explain the meaning of sketchiness. Their answers were then used to establish the main categories of interpretations that people spontaneously associate to sketchiness, and to examine where uncertainty is prominent among them. In Part II we ran a similar study, but participants were presented with a closed list of possible interpretations for sketchiness (multiple choice). The list was established from Part I. The goal of Part II was to check if the extent of this association changes with the introduction of limited alternatives. In Part III we compared sketchiness

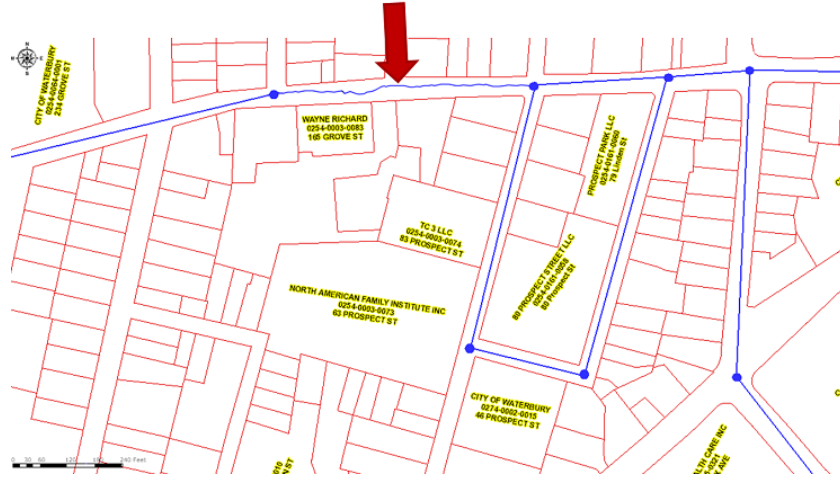**A utility map showing water pipes in blue**

Figure 5: An example scenario of the geometry context (a utility map).

to blur—a visual variable that is highly regarded in the literature as being congruent with uncertainty depiction—using the same closed list of categories as in Part II.

For this series of studies, and based on our own experience and observations from related work, we hypothesized that:

H1 People are likely to spontaneously associate sketchiness to uncertainty in the data.

H2 People are more likely to associate sketchiness to uncertainty when in abstract contexts (e.g. hierarchy) than for non-abstract context (e.g. map) where geometry already has an inherent meaning.

H3 Participants are more likely to associate the meaning of blur to uncertainty, than they are to associate the meaning of sketchiness to uncertainty.

5.1 Part I: *Sketchiness*

Scenarios. We asked people to look at different visualizations with lines as a major visual feature and to interpret what sketchy lines mean to them. We designed six different scenarios, the first 4 being *abstract* contexts, and the last two *non-abstract* contexts : (S1) a bar chart where sketchy lines were applied to the contours of some of the bars, (S2) a social network graph where sketchiness was applied to some connections between nodes (i. e., to relationships between two connected people), (S3) a family tree where sketchiness was applied to some connections between parents and children (figure 6), (S4) a Venn diagram where sketchy outlines were applied to (parts of) the outlines of some ellipsoids that indicate set memberships, (S5) a rail network where sketchiness was applied to some of the links between two train stations , and (S6) a utility map where sketchiness was applied to line representations of some buried assets (figure 5).

For each scenario we generated 5 image variations, changing the percentage of lines in the visualization that were represented as sketchy (between 10% to 50% of all lines in the visualizations). The sketchy lines themselves were hand-drawn with an AnotoPen. Participants only saw one variation of each scenario.
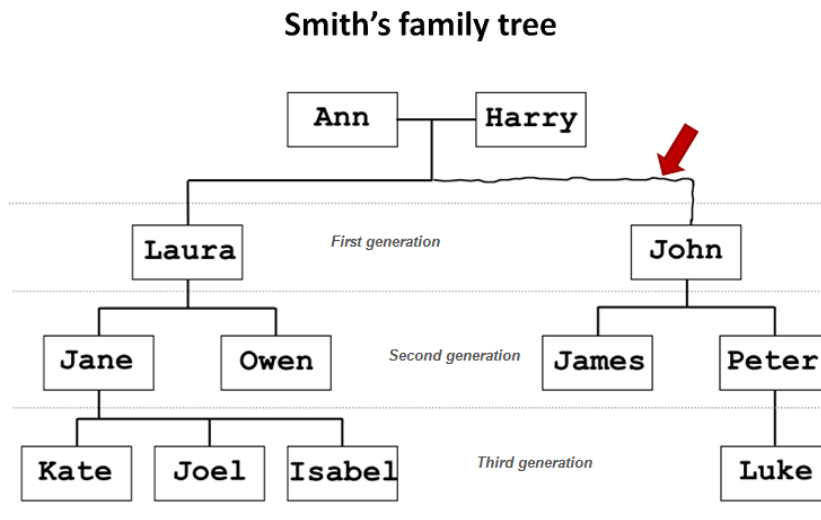
Smith's family tree

Figure 6: An example scenario of the geometry context (a family tree).

Participants, Study Design, and Procedure. 210 participants took part in this study. Participants were split into six groups of 35 people. Each group was exposed to one of the different scenarios described above and each participant in the group saw a single variation of the scenario. Participants were first introduced to the task, provided with a short scenario description, and then presented with an image as described above. We asked participants to type two different interpretations into a text box stating what these lines convey to them. We highlighted each of the sketchy lines in question using a red arrow. Overall, our experiment consisted of:

$$
\begin{array}{rl}
& 6 \quad \text{scenarios} \\
\times & 35 \quad \text{participants per scenario} \\
\hline
= & \mathbf{210} \quad \textbf{trials in total}
\end{array}
$$

After discarding entries that obviously did not represent an interpretation (e.g. jokes, unrelated questions) [47], we could analyze a total of 180 trials, 30 for each scenario. To ensure that our sample participant size was representative of the general participant pool, we conducted our study in 5 increments (of 42 trials). For each block we calculated the mean for the number of interpretations that associate sketchiness to uncertainty. After the forth block the means stabilized, and we ran a fifth block to verify the stabilization. This sampling procedure was followed throughout our experiments.

Results. We carried a qualitative evaluation of the results; one of the authors coded all interpretations; then a second author independently encoded 40% of the total number of interpretations. The concordance rate between the two encoders was around 74%. Coding conflicts were resolved and we were able to identify 6 major categories of interpretations for sketchiness as given below with their association rate (derived using the mean for both interpretations combined). In contrast to what we stated in our first hypothesis H1, we found that only **11.7%** of people associate sketchiness to uncertainty, less than we expected. The results for all categories are listed below.

- Alternative (36.7%): a different relation from what is conveyed by a straight line, which includes 'geometry' where sketchiness is attributed to the actual shape of the displayed feature.

- Qualitative (23.3%): the same relation as for a straight line but with emphasis on a particular quality other than uncertainty. Often, this was a negative quality.

- Ignore (13.9%): the exact same relation as for straight lines (no added information).

- Uncertainty (11.7%): all data quality descriptors that are related linguistically to the term uncertainty such as ambiguity, vagueness, impreciseness, doubt, or unreliability.

- Style (10.0%): intended drawing style; e. g., to draw attention to a particular part of the image.

- Glitch (4.4%): unintended style; i. e., human or computer error.

In the next two studies we use a reduced version of this list in which we included everything but the 'ignore' category.

5.2   Part II: Primed *Sketchiness*

Similar to Part I of the study in subsection 5.1, we asked people to look at a line drawing and interpret what the sketchy line might mean. We provided participants with the five-category list established above, further adapted for sketchiness *abstract* and non-*abstract* contexts.

Participants, Study Design, and Procedure. 168 participants took part in this study, and a between-subjects design was adopted. The design of the study was identical to the study in Part I, with the exceptions that we asked participants to pick an interpretation from the provided list of options and that we adapted to include the 'geometry' category for the non-*abstract* contexts S5 and S6Overall we had:

$$
\begin{aligned}
& 6 \ \text{scenarios} \\
\times \ & 28 \ \text{participants per scenario} \\
= \ & \textbf{168} \ \textbf{trials in total (1/3 for spatial scenarios)}
\end{aligned}
$$

Results. We conducted this study in increments of two blocks (discrepancy between the overall means per category was 6.5% at most).

Overall, people attributed sketchiness first to style (36.3%) and then to uncertainty (22.0%), showing an increase for uncertainty under the closed list condition. Next category scores, in increasing order, were qualitative (13.69%), alternative (12.50%), geometry (10.12%) and glitch (5.36%). Separating sketchiness abstract and non-abstract scenarios, we found an important difference: in agreement with our hypothesis H2, 28.6% of participants under the *abstract* condition attributed sketchiness to uncertainty in comparison to only 8.9% for the non-*abstract* condition. This may have implications on the type of visualization context for which sketchiness is more intuitive as a visual indicator of uncertainty.

5.3   Part III: *Blur*

We replicated the previous study (Part II) for the *blur* visual variable, therefore the tasks are the same as in subsection 5.2.

Participants, Study Design, and Procedure. 168 participants took part in this study, and between-subjects design was adopted. Participants were split into six groups with 28 people per group. Each group was exposed to one of the different scenarios S1–S6. Participants were first introduced to the task, were given a short description of the scenario, and then were presented with an image in which we had applied a Gaussian blur to one line (as outlined in section 4). We asked participants to pick a single interpretation from the five-category list in subsection 5.1. We highlighted the blurred line in question using a red arrow each time.

Overall, our experiment consisted of:

$$
\begin{aligned}
& 6 \ \text{scenarios} \\
\times \ & 28 \ \text{participants per scenario} \\
= \ & \textbf{168} \ \textbf{trials in total (1/3 for spatial scenarios)}
\end{aligned}
$$

Results. We conducted this study in 2 increments (the discrepancy between the overall means per category was 4.2% at most). Overall, people attributed blur mostly to style (26.2%), similar to the results for sketchiness. The second-most frequent attribution was a qualitative measure (23.8%) and only the third-most frequent one was uncertainty (22.6%). Contrary to our hypothesis H3, sketchiness is as intuitive as blur for the categories tested. The remaining attributions for blur were alternative (17.26 %) and glitch (10.12 %)

6   On the Question of Practical Considerations

After having established that sketchiness is just as associated with uncertainty as blur, we can now investigate the practical aspects of using *sketchiness* for this purpose, or as a visual variable in information visualizations in general. More specifically, we are interested in determining if we can express a scale using *sketchiness* as accurately as when using one of the other three visual variables traditionally used to represent uncertainty (blur, grayscale, and dashing). We thus compare the four techniques with respect to: (i) how close their perceived values are to the actual visual variable value; (ii) how many distinct levels can be identified by participants for each technique to use them in ordinal scales; and in (iii) whether the perception of these techniques follows previous models that express this perception mathematically.

6.1   Hypotheses

Based on our experience and observations from related work, we hypothesized that a few discrete levels of line *sketchiness* can be perceived but that people cannot accurately quantify the amount of *sketchiness* applied to a line. This led to the following hypotheses:

H4 Overall, people cannot accurately estimate the exact level for all techniques (sketchiness, blur, grayscale, and dashing), but this perception error will be different between techniques.

H5 People can discriminate between at least three levels of *sketchiness* but more levels for the remaining techniques.

H6 We can express the relationship between reported and actual levels for all visual encoding techniques (including *sketchiness*) using a mathematical model.

6.2   Task

To assess the visual perception of the techniques we conducted a study with tasks from psychophysics, the domain that focuses on measuring relationships between perceived and actual properties of visual objects [17, 45]. Of the methods that help assess a viewer's visual perception of an object compared to its subjectively experienced magnitude, numeric estimation methods are most relevant to our research and have been frequently used in in the past (e. g., [17, 45]). Participants are shown a ï¿½standardï¿½ modulus object with an assigned value (e. g., 100% *sketchiness*) and are then asked to assess a second object (the stimulus) and assign it a value based on the modulus (e. g., a percentage).

In our study participants were shown one line representing the maximum magnitude for the specific visual variable technique (100%, our modulus) as well as a stimulus line that they had to express as a percentage of the modulus. We used the previously generated 20 magnitude *levels* for stimulus lines for all our visual variables. These 20 stimulus levels correspond to 5–100% of the modulus object. Participants were also shown a line representing the minimum (0%) for the technique which was always a straight black line of 3 pixels for all techniques.

6.3   Participants, Study Design and Procedure

160 participants took part in this study. A mixed factorial design was used: the *visual variable* was treated as a between-subjects factor and the *level* was treated as a within-subjects factor. Participants were split into four groups of equal sizes. Each group was exposed to a different technique/visual variable and each participant conducted magnitude estimation tasks for all levels using that technique.

Participants were first introduced to the task. They were then presented with lines of different levels of their respective visual variable and were asked to mark on a scale from 0% to 100% how sketchy each line was with respect to the maximum modulus line (always visible). The order of presentation of the different levels was randomized across participants and techniques. We tested a total of 21 levels per technique (including level 0). Therefore, our experiment consisted of:

	4	techniques (dash, blur, grayscale, and sketchiness)
×	21	magnitude levels (including level 0)
×	40	participants per technique
=	**3360**	**trials in total**

6.4   Results

The metrics used in our analysis were the absolute perception error *AbsErr* and the perceived magnitude level *PerMag* of a line. As done by similar magnitude estimation studies (e. g., [10]) we define *AbsErr* as the absolute difference between the true percent of the stimulus compared to the modulus and the reported estimation percent $|reported\ level - true\ level|$. Absolute magnitude errors have a skewed distribution and, as suggested by Cleveland [10], we normalize it by using the log variation of this metric for our analysis $log_2(\frac{1}{8} + AbsErr)$. The means reported here are before normalization.

Trials were marked as outliers when metrics were beyond two standard deviations from the mean for a given technique and level. 161 trials (5% of all trials) were identified as outliers and removed from further analysis. We performed an ANOVA and post-hoc pair-wise mean comparison p-values are adjusted using the Bonferroni criterion.

6.4.1   Perception Error

We first examined how closely participants came to predicting the real level value for each technique, conducting an analysis on *AbsErr*. The overall *AbsErr* was higher for *sketchiness* (16.6%), followed by *dash* (13.2%), *blur* (12.7%), and *grayscale* (12.3%).

AbsErr increased steadily from 5% to 20% overall as we increased the level up to 11 (half of maximum magnitude). Then we observed a steady decrease in error for all techniques from 20% to 15% for level 19, and almost 5% in level 20. The ANOVA showed that the main effect of technique on *AbsErr* was statistically significant ($F_{3,156} = 4.890$, $p < .01$). A significant interaction effect between technique and level was also present ($F_{60,3120} = 6.364$, $p < .001$). Post-hoc comparisons showed that level predictions with *sketchiness* were significantly further from the true level values than for the other techniques, but this was true only after level 10, with no difference between techniques for levels up to 10 (all $p < .05$). This evidence supports H4.

6.4.2   Level Groupings per Technique

We then tried to determine for each technique how many levels participants can distinguish comfortably, and which level values are good candidates for an ordinal scale, assuming that our data are an accurate representation of real perception effects for the different visual variables. We thus conducted the following exploratory analysis:

We first plotted the mean and standard deviation of the *PerMag* each level per technique and identified possible level values that are distinguishable between them. We chose levels that have similar standard
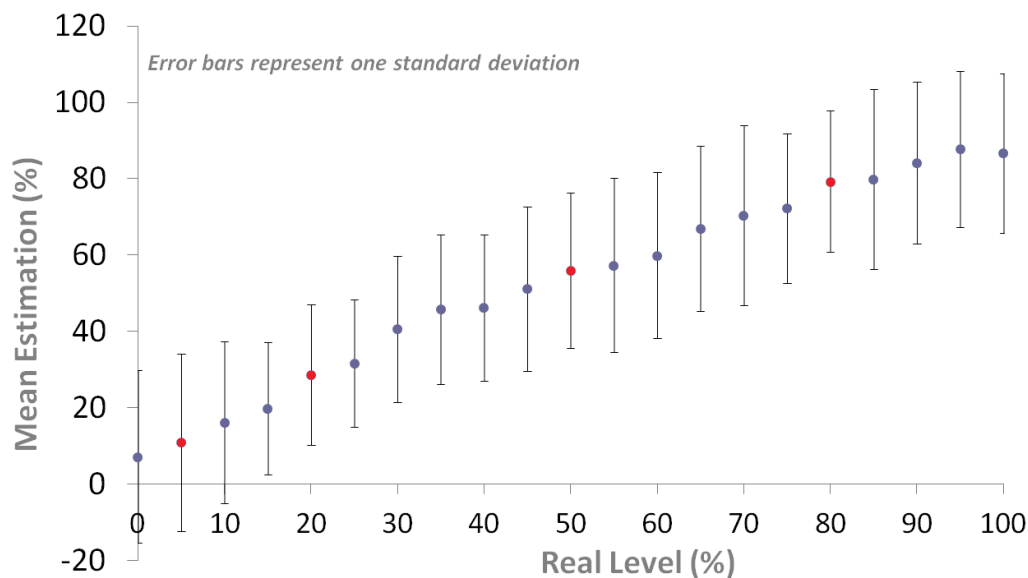
Figure 7: Real vs. estimated blur with red dots for distinct levels.

deviations to their neighbors (to ensure we select representative levels for their neighborhood), and that have no overlap between their mean and the standard deviations of the next selected levels. This process gave us initial estimates for the number levels that can be distinguished per visual variable and the level value that corresponds to them (see Figure 7, 8, 9 and 10 with levels highlighted).

We then ran an ANOVA to see what levels were significantly different with respect to their perceived magnitude *PerMag*. We plotted the levels that are *not* significantly different on the matrices in Figure 11 where gray squares indicate levels that were not significantly different and risk being perceived as similar. By inspecting these images of statistical results, we can see visual clusters of levels whose mean values are similar and thus should not all be used to represent two distinct levels. We ensured that our chosen levels (in black in Figure 11) fall under different such visual clusters, to further supporting our level choices.

Our ANOVA was conducted with *visual variable* treated as a between-subjects variable and *level* as repeated measures. There was a main effect of technique on *AbsErr* ($F_{3,156} = 24.27$, $p < .0001$) and, more importantly for us, its interaction effect with level was also significant ($F_{60,3120} = 6.047$, $p < .0001$), indicating that level perception was different between variables. A post-hoc comparison showed levels that were clearly different in their perception from each visual variable technique (all $p < .05$) and are seen in white in the matrices.

Finally, we ran a hierarchical clustering algorithm across level means for each visual variable to ensure that the levels we chose fall under different clusters found by the algorithm (Table 1). This was indeed the case, further supporting our level choices. Based on our generation process we propose in Table 1 levels that are clearly distinguishable by technique. Indeed at least 3 (4 counting the maximum sketchiness level) levels were identified for *sketchiness* with 4 levels for blur and grayscale (supporting H5).
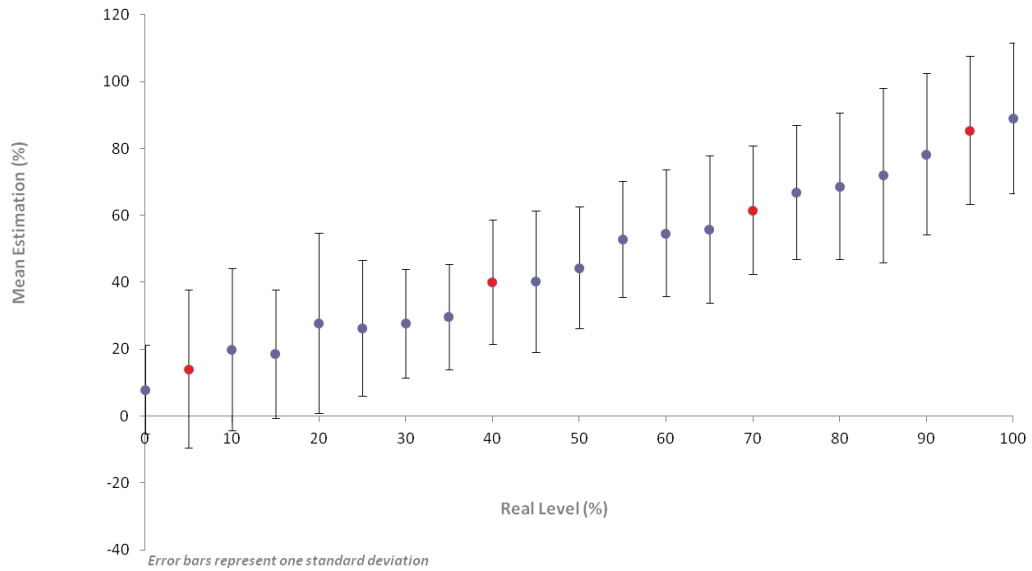
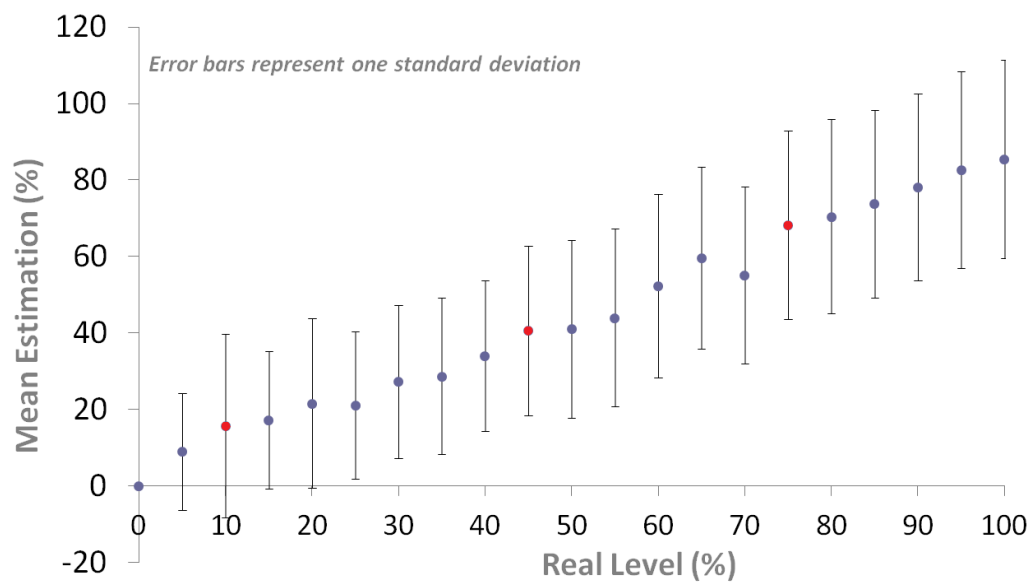Figure 8: Real vs. estimated grayscale with red dots for distinct levels.

Figure 9: Real vs. estimated dashing with red dots for distinct levels.
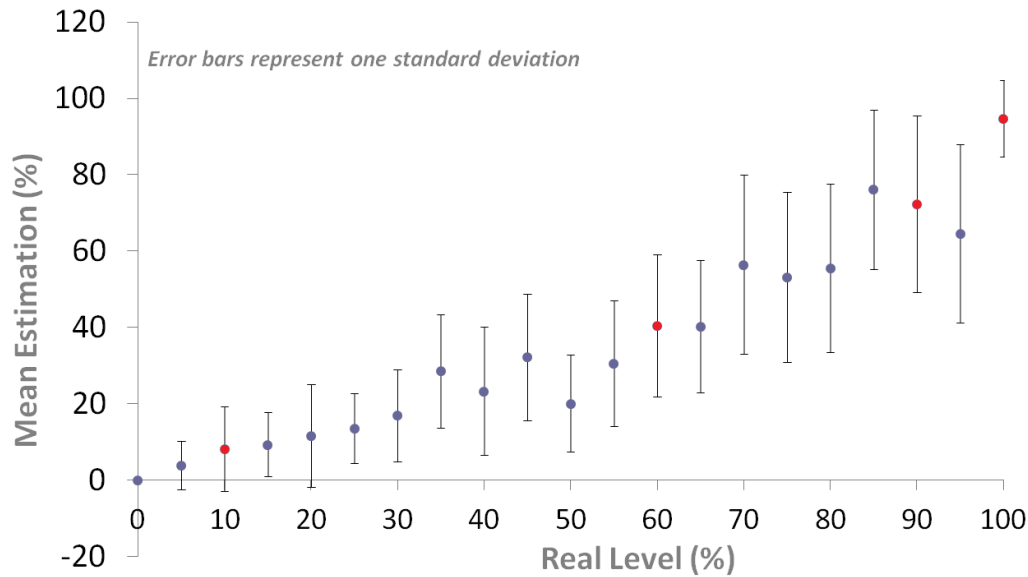
Figure 10: Real vs. estimated sketchiness with red dots for distinct levels.

	Levels	Level Values (and their clusters)
blur	4	[0, **1**, 2, 3] [**4**,5,6,,7,8,9] [**10**,11,12,13,14,15] [**16**,17,18,19,20]
dash	3	[0, 1, **2**, 3, 4, 5, 6, 7, 8] [**9**,10,11,12, 13,14] [**15**,16,17,18,19,20]
grayscale	4	[0, **1**, 2, 3] [4,5,6,7,**8**,9,10] [11,12,13,**14**,16,17] [18,**19**,20]
sketchiness	3-4	[0, 1, **2**, 3,4,5,6,8,10] [7,9,11,**12**,13] [14,15,16,17,**18**,19] [**20**]

Table 1: Distinct levels, representative values and their clusters.

6.4.3 Mathematical Description of Perceived Techniques

Previous work has attempted to mathematically describe the differences between physical and perceived magnitude of objects as collected from user studies. One popular function describing this difference is Stevens' [39] power law: $J = \lambda D^{\alpha}$, with $J = judged\ magnitude$, $D = actual\ magnitude$, $\alpha = exponent$, $\lambda = scaling\ constant$. Wagner [45] provides a meta-analysis of articles reporting values for α collected under different conditions. To best of our knowlege, no conditions matched *dash* or *blur* (and *sketchiness*), nevertheless there are values for *grayscale*. Stevens and Galanter [40] found an exponent of 1.2 for black-gray-white series. Given the previous discrepancies between the α varying across experiment setups [45], we decided to mathematically describe the perceived magnitude of all four different visual variables.

An initial curve fitting for our visual variables techniques indicated that, indeed, a power model best fits our data (for all visual variables their respective fits had $R^2 > .9$, $p < .0001$). Using the parameter estimates from the fit we conducted a detailed regression analysis on our data to verify the mathematical relationship. Linear regression analysis for each of the techniques showed a very good fit (all $R^2 > .9$ and all adjusted $R^2 > .9$), as hypothesized (H6). Our results and the used coefficients are summarized in Table 2 (all results significant at the 99% level).
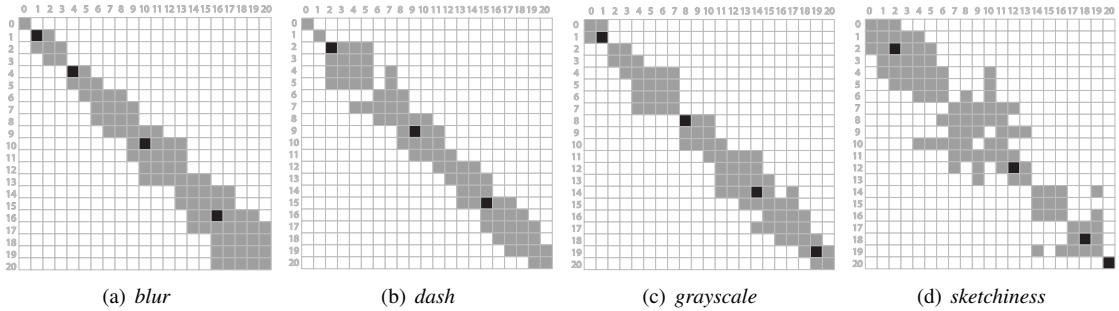
(a) *blur* (b) *dash* (c) *grayscale* (d) *sketchiness*

Figure 11: Black squares show our chosen representative levels, and gray ones indicate levels that are *not* significantly different.

	$J = \lambda D^{\alpha}$		Regression	
	λ	α	R^2	Adjusted R^2
blur	3.149	0.729	0.993	0.993
dash	2.188	0.774	0.974	0.973
grayscale	3.819	0.649	0.958	0.955
sketchiness	0.601	1.031	0.916	0.912

All results significant at the 99% level.

Table 2: Relationships and regression results.

6.4.4 Discussion on Levels

We found that the perceived *sketchiness* is farther from the real value of a level than for the other techniques, but not by much (4%). This prediction error varied across techniques and was always $> 10\%$. This value is somewhat large and can be too prohibiting if participants attempt to accurately retrieve values from the visual representations of these techniques in real life applications (H4). We thus believe all techniques should be used for scales but not for value retrieval.

As in previous studies, our perceived variables can be mathematically modeled using a power law (H3). The fact that our coefficients do not match others on *grayscale* as reported by Wagner [45] can be attributed to differences in study setups (as is often the case for perception experiments). An interesting observation is that the perception of sketchiness is fairly close to linear. This could be explained by our generation model that defines a sketchiness level based on the maximum distortion, as opposed to an average distortion across a line.

Using exploratory methods we were able to identify distinct levels within each of the visual variables (at least 3 for *sketchiness*—H5). We note here that our selection is somewhat conservative. We base our selection on data that assume that each of the selected levels will be compared independently with the min/max values (e. g. when seeing a line using *sketchiness* in a visualization the user will immediately be able to determine if it indicates low, medium, or high uncertainty). Nevertheless, if viewers are presented with visual representation of other levels (e. g. comparing the *sketchiness* between two lines in a visualization, or between a line and a legend), it is possible they will be able to distinguish even more levels. This requires further investigation.

7 On the Question of Preference

One issue that still needs to be investigated is which techniques are preferred by people for visualizing uncertainty. Based on our own results from the intuitiveness study in section 5 and also based on related

work (notably on blur [24]) we had the following hypothesis:

H7 People prefer sketchiness to blur; but dashing and grayscale may be preferred overall due to people's familiarity with these styles.

7.1 Task

To judge people's preference about uncertainty encoding we asked participants to select one of four visual styles—blur, dashing, grayscale, or sketchiness—in a visualization scenario. To motivate participants into thinking about their choice and to avoid random answers we asked participants to justify their choice.

7.2 Participants, Study Design and Procedure

129 participants took part in this study and a between-subjects design was used. Participants were split into 6 groups of equal size. Each group was exposed to a different scenario chosen from our first study (section 5). Participants were first introduced to the task, and a short description of the scenario was given. They were then presented with four side-by-sideversions of the same image which only differed in the rendering of a single line (depicted blur, dashing, grayscale, or sketchiness). All values of the uncertainty visual variables were selected from the middle ranges for consistency. Participants were informed that the study is about comparing different styles to show uncertainty using the aforementioned visual variables and they were asked to choose which style they prefer and give a reason for their choice. In each image we used a red arrow to highlight the uncertain line in question.

Overall our experiment consisted of:

$$
\begin{array}{rl}
 & 6 \ \text{scenarios} \\
\times & 32 \ \text{participants per scenario} \\
= & \textbf{192} \ \textbf{trials in total}
\end{array}
$$

7.3 Results

We ran our study in batches of 48 trials. For each batch we calculated the mean for each visual variable. After the third batch the means stabilized. Our results show that dashing is the preferred style for the participants (chosen by 68.3%), in agreement with our hypothesis (H7). However, blur (chosen by 15.10%) did better than both grayscale (chosen by 12.5%) and sketchiness (chosen by 3.12%).

We were able to process comments from 155 participants (80.7% of the responses). A closer look at their comments shows that the primary reason for preferring dashing over the other visual variables was '*noticeability*' as participants valued the ability to easily distinguish between data and uncertainty (26.5% of our participants), whilst blur was chosen because it was regarded as congruent to what uncertainty conveys in terms of vagueness and reduced precision (4.5% of our participants). Those who preferred grayscale did so because they deemed it easy to understand (2% of our participants).

We only had 4 comments on the use of sketchiness, 2 of which argued that it is an intuitive way to represent instability in the data. More interesting, and which may explain the low preference score for sketchiness, are comments by participants who argued against the use of sketchiness. The primary reason for this seems to be that sketchiness tends to imply a disliked notion of 'unprofessionalism' or informality. Some participants thus disregarded it as a standard visual variable (4.5%). This highlights the need for paying attention to the context in which sketchiness is applied as a visual variable to denote uncertainty.

8 Discussion and Conclusion

In this paper, we have reported on six user studies to investigate the appropriateness of using *sketchiness* as a dedicated visual variable for depicting uncertainty. We started by studying characteristics of

collected hand-drawn strokes. Findings from this evaluation and existing work on a generative model of hand movement fed into our method for synthesizing sketchy lines that mimic stroke characteristics of hand-drawn lines. Our mapping for uncertainty depiction consists of varying the amount of pixel deviations from a straight line in accordance to the amount of uncertainty in the data. An online study using Amazon MechanicalTurk provided evidence that lines generated by our model were significantly more often thought to be handwritten then not.

We then investigated whether sketchiness is an appropriate visualization technique for representing uncertainty. The three questions we attempted to answer were:

(a) Can people intuitively understand sketchiness as an indicator of uncertainty? In a series of studies we found that if users are asked what a sketchy line represents, their most common reaction is that it is associated with a different semantic relation from what is conveyed by a straight line (other than uncertainty), with only 11.7% attributing it to uncertainty. When given a multiple choice of possible explanations for sketchiness, it was associated with uncertainty 22% of the time. This was lower than expected, nevertheless a similar study showed a very similar effect for blur (22.6%) which is traditionally associated with uncertainty in visualizations. Thus, although sketchiness was not spontaneously associated with uncertainty as much as we would have liked, we feel it is still a viable alternative to other visualization techniques, that can be now used to represent other information in the data.

(b) Is sketchiness as effective for depicting uncertain information as the other visualization techniques? In terms of practically applying sketchiness as a visual variable to information visualizations our results were encouraging. We compared sketchiness with other visual variables in a magnitude estimation task (which is very common in information visualization, e. g. comparing a given object to a legend). Our studies showed that sketchiness was only slightly more error-prone than other techniques, the difference in error not being more than 4%. In fact, none of the techniques was very accurately perceived (all had very similar error rates of more than 10%). This indicates that none of them is particularly well suited for true value retrieval, and the common practice of using them in small ordinal scales of uncertainty is indeed a sound one. Our mathematical modeling for predicting all visual variables indicates that this error is predictable for all techniques (including sketchiness).

To help uncertainty visualization designers we proposed, based on our findings, a number of levels for ordinal scales, for each visual variable (3–4 excluding maxima). The number of levels we identified and proposed are close in number between sketchiness (3) and other techniques. Our process for selecting levels is based on the assumption that estimates are made when comparing to a maximal value object. Nevertheless, very likely in the presence of other visual levels (e. g., in a legend) more levels can be distinguished. Further investigation is need to see how sketchiness fairs under these situations.

(c) Which method do people subjectively prefer? Both our preference study and our exploration of visualization scenarios where sketchiness can be attributed to the geometrical form of data (non-abstract contexts) indicate that the use of sketchiness as a visual variable to indicate uncertainty should be done with caution. Sketchiness is not spontaneously associated with uncertainty in non-abstract contexts, even when participants are told that sketchiness is not a geometric feature. It is thus preferable to be avoided in these situations. Moreover, some participants commented on the 'informal' and 'unprofessional' look of sketchy lines in a negative way.

We conclude that sketchiness as a visual variable is a viable additional choice for depicting uncertainty in ordinal data, but further work is needed to investigate the appropriate visualization contexts where it can be deployed.

References

[1] J. C. J. H. Aerts, K. C. Clarke, and A. D. Keuper. Testing Popular Visualization Techniques for Representing Model Uncertainty. *Cartography and Geographic Information Science*, 30(3):249–261, July 2003. doi> 10.1559/152304003100011180

[2] R. S. Allendes Osorio and K. W. Brodlie. Contouring with Uncertainty. In *Theory and Practice of Computer Graphics*, pp. 59–65, Aire-la-Ville, Switzerland, 2008. Eurographics Association. doi> 10.2312/LocalChapterEvents/TPCG/TPCG08/059-065

[3] Z. AlMeraj, B. Wyvill, T. Isenberg, A. A. Gooch, and R. Guy. Automatically Mimicking Unique Hand-Drawn Pencil Lines. *Computers & Graphics*, 33(4):496–508, Aug. 2009. doi> 10.1016/j.cag.2009.04.004

[4] J. Bertin. *Semiology of Graphics: Diagrams, Networks, Maps*. ESRI Press, Redlands, California, 2010.

[5] A. M. Bisantz, T. Kesevadas, P. Scott, D. Lee, S. Basapur, P. Bhide, P. Bhide, and P. Bhide. Holistic Battlespace Visualization: Advanced Concepts in Information Visualization and Cognitive Studies. Technical report, UB Engineering, University of Buffalo, USA, June 2002.

[6] R. P. Botchen, D. Weiskop, and T. Ertl. Texture-Based Visualization of Uncertainty in Flow Fields. In *Proc. IEEE Visualization*, pp. 647–654, Los Alamitos, 2005. IEEE. doi> 10.1109/VISUAL.2005.1532853

[7] N. Boukhelifa and D. J. Duke. Uncertainty Visualization – Why Might it Fail? In *CHI Extended Abstracts*, pp. 4051–4056, New York, 2009. ACM. doi> 10.1145/1520340.1520616

[8] R. Brown. Animated Visual Vibrations as an Uncertainty Visualisation Technique. In *Proc. GRAPHITE*, pp. 84–89, New York, 2004. ACM. doi> 10.1145/988834.988849

[9] A. Cedilnik and P. Rheingans. Procedural Annotation of Uncertainty Information. In *Proc. IEEE Visualization*, pp. 77–84, Los Alamitos, 2000. IEEE Computer Society. doi> 10.1109/VISUAL.2000.885679

[10] W. S. Cleveland and R. McGill. Graphical Perception and Graphical Methods for Analyzing Scientific Data. *Science*, 229:828–833, 1985.

[11] C. Curtis. Loose and Sketchy Animation. In *SIGGRAPH Technical Sketches*, p. 317, New York, 1998. ACM. doi> 10.1145/281388.281913

[12] S. Diepenbrock, J.-S. Praßni, F. Lindemann, H.-W. Bothe, and T. Ropinski. Interactive Visualization Techniques for Neurosurgery Planning. In *Eurographics Short Papers / Dirk Bartz Prize for Visual Computing in Medicine*, pp. 13–16, Goslar, Germany, 2011. Eurographics Association.

[13] A. Finkelstein and D. H. Salesin. Multiresolution Curves. In *Proc. SIGGRAPH*, pp. 261–268, New York, 1994. ACM. doi> 10.1145/192161.192223

[14] T. Flash and N. Hogans. The coordination of arm movements: An experimentally confirmed mathematical model. *Journal of neuroscience*, 5:1688–1703, 1985.

[15] L. E. Gerharz and E. J. Pebesma. Usability of Interactive and Non-Interactive Visualisation of Uncertain Geospatial Information. In *Proc. Geoinformatik*, pp. 223–230, Münster, Germany, 2009. ifgiPrints.

[16] N. D. Gershon. Visualization of Fuzzy Data Using Generalized Animation. In *Proc. IEEE Visualization*, pp. 268–273, Los Alamitos, 1992. IEEE. doi> 10.1109/VISUAL.1992.235199

[17] E. B. Goldstein. *Sensation and Perception*. Brooks/Cole Publishing, Pacific Grove, USA, 5th edition, 1999.

[18] B. Gooch and A. A. Gooch. *Non-Photorealistic Rendering*. A K Peters, Ltd., Natick, 2001.

[19] H. Griethe and H. Schumann. The Visualization of Uncertain Data: Methods and Problems. In *Proc. SimVis*, pp. 143–156, Erlangen, Germany, 2006. SCS Publishing House e.V.

[20] J. R. Hamerly and C. A. Dvorak. Detection and Discrimination of Blur in Edges and Lines. *Journal of the Optical Society of America*, 71(4):448–452, Apr. 1981. doi> 10.1364/JOSA.71.000448

[21] J. Heer and M. Bostock. Crowdsourcing Graphical Perception: Using Mechanical Turk to Assess Visualization Design. In *Proc. CHI*, pp. 203–212, New York, 2010. ACM. doi> 10.1145/1753326.1753357

[22] T. Hengl and D. J. J. W. A. Brown. Pixel and Colour Mixture: GIS Techniques for Visualisation of Fuzziness and Uncertainty of Natural Resource Inventories. In *Proc. Accuracy*, pp. 300–308, Delft, the Netherlands, 2002. Delft University Press.

[23] T. Hengl and N. Toomanian. Maps Are Not What They Seem: Representing Uncertainty in Soil-Property Maps. In *Proc. Accuracy*, pp. 805–813, Lisboa, Portugal, 2006. Instituto Geográfico Português.

[24] R. Kosara, S. Miksch, H. Hauser, J. Schrammel, V. Giller, and M. Tscheligi. Useful Properties of Semantic Depth of Field for Better F+C Visualization. In *Proc. VisSym*, pp. 205–210, Aire-la-Ville, Switzerland, 2002. Eurographics Association.

[25] M. Leitner and B. P. Buttenfield. Guidelines for the Display of Attribute Certainty. *Cartography and Geographic Information Science*, 27(1):3–14, Jan. 2000. doi> 10.1559/152304000783548037

[26] H. Levkowitz. *Color Theory and Modeling for Computer Graphics, Visualization, and Multimedia Applications*. Kluwer Academic Publishers, Norwell, MA, USA, 1997.

[27] H. Li, C.-W. Fu, Y. Li, and A. J. Hanson. Visualizing Large-Scale Uncertainty in Astrophysical Data. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1640–1647, Nov./Dec. 2007. doi> 10.1109/TVCG.2007.70620

[28] M. Luboschik, A. Radloff, and H. Schumann. Using NPR-Rendering Techniques for the Visualization of Uncertainty. In *Posters of IEEE InfoVis*, Los Alamitos, 2010. IEEE Computer Society.

[29] A. M. MacEachren. Visualizing Uncertain Information. *Cartographic Perspectives*, 13(Fall):12–19, 1992.

[30] M. Matthews, L. Rehak, A.-L. Lapinski, and S. McFadden. Improving the Maritime Surface Picture with a Visualization Aid to Provide Rapid Situation Awareness of Information Uncertainty. In *Proc. IEEE TIC-STH*, pp. 533–538, Los Alamitos, 2009. IEEE. doi> 10.1109/TIC-STH.2009.5444441

[31] M. Nienhaus, F. Kirsch, and J. Döllner. Sketchy Illustrations for Presenting the Design of Interactive CSG. In *Proc. IV*, pp. 772–777, Los Alamitos, 2006. IEEE Computer Society. doi> 10.1109/IV.2006.97

[32] A. Pang. Visualizing Uncertainty in Geo-spatial Data. In *Proc. Workshop on the Intersections between Geospatial Information and Information Technology*, 2001.

[33] J. J. Pfeiffer, Jr. Using Brightness and Saturation to Visualize Belief and Uncertainty. In *Proc. Diagrams*, pp. 279–289, Berlin, 2002. Springer-Verlag. doi> 10.1007/3-540-46037-3_27

[34] K. Potter, A. Gooch, B. Gooch, P. Willemsen, J. Kniss, R. Riesenfeld, and P. Shirley. Resolution Independent NPR-Style 3D Line Textures. *Computer Graphics Forum*, 28(1):52–62, Mar. 2009. doi> 10.1111/j.1467-8659.2008.01297.x

[35] P. J. Rhodes, R. S. Laramee, R. D. Bergeron, and T. M. Sparr. Uncertainty Visualization Methods in Isosurface Rendering. In *EG Short Papers*, pp. 83–88, Aire-la-Ville, Switzerland, 2003. Eurographics Association.

[36] J. Sanyal, S. Zhang, G. Bhattacharya, P. Amburn, and R. J. Moorhead. A User Study to Compare Four Uncertainty Visualization Methods for 1D and 2D Datasets. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):1209–1218, Nov./Dec. 2009. doi> 10.1109/TVCG.2009.114

[37] S. Schlechtweg, B. Schönwälder, L. Schumann, and T. Strothotte. Surfaces to Lines: Rendering Rich Line Drawings. In *Proc. WSCG*, volume 2, pp. 354–361, 1998.

[38] J. Schumann, T. Strothotte, A. Raab, and S. Laser. Assessing the Effect of Non-photorealistic Rendered Images in CAD. In *Proc. CHI*, pp. 35–42, New York, 1996. ACM. doi> 10.1145/238386.238398

[39] S. S. Stevens. *Psychophysics*. Transaction Publishers, New Brunswick, USA, 2nd edition, 1975.

[40] S. S. Stevens and E. H. Galanter. Ratio scales and category scales for a dozen perceptual continua. *Journal of Experimental Psychology*, 54:377–411, 1957.

[41] T. Strothotte, M. Masuch, and T. Isenberg. Visualizing Knowledge about Virtual Reconstructions of Ancient Architecture. In *Proc. CGI*, pp. 36–43, Los Alamitos, 1999. IEEE. doi> 10.1109/CGI.1999.777901

[42] T. Strothotte, B. Preim, A. Raab, J. Schumann, and D. R. Forsey. How to Render Frames and Influence People. *Computer Graphics Forum*, 13(3):455–466, Aug. 1994. doi> 10.1111/1467-8659.1330455

[43] T. Strothotte, M. Puhle, M. Masuch, B. Freudenberg, S. Kreiker, and B. Ludowici. Visualizing Uncertainty in Virtual Reconstructions. In *Proc. EVA Europe*, p. 16, Berlin, 1999. EVA Conferences International/GFaI.

[44] T. Strothotte and S. Schlechtweg. *Non-Photorealistic Computer Graphics. Modeling, Animation, and Rendering*. Morgan Kaufmann Publishers, San Francisco, 2002.

[45] M. Wagner. *The Geometries of Visual Space*. Lawrence Erlbaum Associates, Mahwah, NJ, USA, 2006.

[46] R. Watt and M. Morgan. The Recognition and Representation of Edge Blur: Evidence for Spatial Primitives in Human Vision. *Vision Research*, 23(12):1465–1477, 1983. doi> 10.1016/0042-6989 (83)90158-X

[47] W. Willett, J. Heer, and M. Agrawala. Scented widgets: Improving navigation cues with embedded visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1129–1136, Nov. 2007.

[48] G. A. Winkenbach and D. H. Salesin. Computer-Generated Pen-and-Ink Illustration. In *Proc. SIGGRAPH*, pp. 91–100, New York, 1994. ACM. doi> 10.1145/192161.192184

[49] Z. Xie, S. Huang, M. O. Ward, and E. A. Rundensteiner. Exploratory Visualization of Multivariate Data with Variable Quality. In *Proc. VAST*, pp. 183–190, Los Alamitos, 2006. IEEE. doi> 10.1109/ VAST.2006.261424

[50] B. Zehner, N. Watanabe, and O. Kolditz. Visualization of Gridded Scalar Data with Uncertainty in Geosciences. *Computers & Geosciences*, 36(10):1268–1275, Oct. 2010. doi> 10.1016/j.cageo. 2010.02.010

[51] T. Zuk and S. Carpendale. Theoretical Analysis of Uncertainty Visualizations. In *Visualization and Data Analysis, Proc. SPIE-IS&T Electronic Imaging*, pp. 606007/1–14. SPIE, 2006. doi> 10.1117/ 12.643631

Contents