Thinking Fast Thinking Slow

Abhishek Kulkarni, Ashish Aggarwal, Nishtha Shrivastav, Swarada Sathe

Abstract— Visualizations make it easier for the human brain to understand and comprehend complex datasets. It also helps detect patterns and draw useful insights from the data. Research on visualization of data using the conventional data encoding approaches is very vast and has been experimented with by many researchers. However, the relationship between the type of motion and data that is being represented is still not fully explored. We plan to study the effectiveness of motion in presenting quantitative data using moving visuals and compared them to traditional common static visual encodings by evaluating the responses of subjects exposed to both the visuals. We will further extend this work to understand different types of motion to evaluate whether it plays a role in perception of quantitative data. We fashioned a user study in which we showed various visuals with varying speeds to the study participants and asked them to identify the change in quantities such as area, length,angle and color gradients. We also asked the same questions with quantities represented using static visuals and compared both the results.

Index Terms—Motion Encoding, Information Visualization, Evaluation Study

1 Introduction

In today's world, large amounts of data is being generated and manual analysis of data is a daunting and practically impossible undertaking. Representation of the data in the form of effective visualizations is essential for understanding the gist of the data. Especially, quantitative data needs to be visualized for comprehension and to draw out important conclusions from it. For representation of a particular dataset, suitable data encodings are employed to visualize the data and show the underlying relationships in the dataset. The most common data encodings are color, position, area, length,etc. Utilizing motion as an encoding or in combination with other established encodings is a research area that has not been sufficiently explored yet. The main motivation of this work is therefore to examine motion as a quantitative data encoding parameter. The conclusions drawn from this work could potentially lay groundwork for various visualization techniques incorporating motion for illustration of a dataset. It also gives insights into effect of motion as an encoding technique on perception of data and whether the intended interpretation of the data is conveyed to the viewer.

To achieve the goal for this project, we designed a user study whose analysis and evaluation will enable us to understand whether motion aids perception of quantities. The user study was designed in such a way that it accounted for the interpretation of the static visualizations as well as the dynamic visualizations i.e. visualizations that contained motion as an encoding parameter. The study also had questions wherein it was tested if the speed of the motion affected the perception of the data. The questions were fashioned to evaluate the quantitative response from the participant. For example, the participants were given two samples of a combination of motion and position encoding and the participant was supposed to answer how different they are from each other in terms of quantitative value they hold.

After generation of data from a sufficient number of participants, we analysed the responses to evaluate how accurately was the intended information perceived. We attributed a range in which each answer could be considered as accurate, and if the answer was not in the determined range for a particular question for more than a certain percent of the participants, it was considered that the visualization could not represent the intended interpretation effectively.

2 Related Work

A number of studies have been conducted on this matter which helped us lay a foundation for our research study. [1] describes an experiment that was carried out to test how effectively graphs are perceived when it includes quantitative encoding of data. This paper gives us a possible plan of action as it tests a representation type for quantitative encoding, that has not yet been much explored. Motion encoding for animated edge textures for a node-link diagram is explored by [2]. The difference between our approach and the approach of the paper is that this paper exclusively studies motion only for node link diagrams and we will be studying motion as a generalized approach for quantitative encoding. [10] conducted a study based on 12 interactive graphic encodings to compare how adding interactivity directly to visualizations differs in results from having a control panel to manipulate them. This research paper is especially relevant to our project as it describes a user-study driven method to analyse the importance of various visualization encodings. Precedence to an online survey over a controlled experiment in this case is crucial, based on which we decided to opt for an online survey. In [11] the researchers designed a study to understand motion encoding, specifically flicker, direction and velocity in motion and its use in visualizations. The key takeaway from this paper is conduction of independent studies to study different properties of motion. Similarly, [13] conducts a study to understand three distinct types of periodic motion: circular, linear, and expansive/contraction and its effectiveness in visual search.

Similar to our proposed study, [8] conducted a series of user studies to verify if adding motion to the static visualization can help us compare the perception of quantitative values. Important takeaways from this paper were the provenance of the point that "that the trajectory of motion can provide a larger perception range than using color encoding" and that when the number of objects in motion at a time increases, attention span decreases and complexity increases. The latter takeaway helped us limit the number of moving objects in a visualization. [4] conducted a study on oscillatory motion to come up with a conclusion that multiple objects inphase can be understood better than the objects that are out of phase.

[6] talks about motion correlation as a principle which allows a user to follow and select an object based on the kind of movement it displays. This paper helps us understand how eye movements and gestures can be studied to determine the selection of objects for interaction. The authors of [5] investigate how different properties of motion relate to user's detection, focus and distraction. In our project

we plan to study the effectiveness of motion in presenting quantitative data as compared to traditional common static visual encodings. Thus the methodology used in this research is highly relevant to our goals.

[9] elaborates on how a heuristic evaluation should be carried out for any user interface. He lays down some guidelines that often lead to successful evaluation of a user interface. To carry out efficient evaluations, it is vital to understand what makes an evaluation successful. Thus, this paper helps gain clarity on how to approach an evaluation and what factors need to be considered when designing a study. [14] discusses the cognitive aspects of motion and visualization like ``mental maps" to better understand the internal representation of data in human mind. This helps is contextualizing our work and relate our experiment tasks with broader human perception ability discussed in the literature. [15] studies the effectiveness of animation using flickering points to alleviate overdraw in scatterplot matrix (SPLOMs). This paper is directly relevant to our research project as it studies the role of a type of motion in inferring scatterplot matrix which is similar to studying the quantitative inferring through motion variation. [16] studies the effect of different modalities of motion and its particular attributes on affective qualities of interpretation. It discusses the affective properties of motion which again gives us a better understanding of the potential of motion aesthetically so that when we analyze our results, we can have a strong and informed discussion to explain the interpretations of our results. [2] and [12] discuss the method of using micro-task markets like Amazon's Mechanical Turk. In our project, we need to create a dataset from scratch. A method like micro-task markets seems like a good fit to collect considerable data at a low cost. These papers give us a possible line of approach for expanding the range of participants and to ensure diversity.

3 Method

There are two major parts of this project, one is designing and conducting the survey and the other is analysis and evaluation of the results.

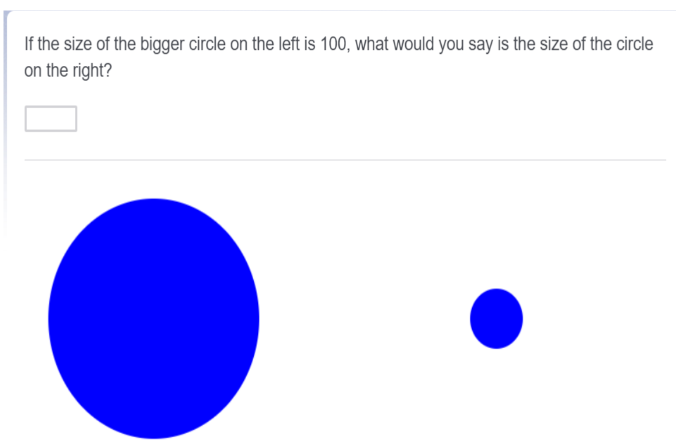

Fig 1. A static visualization from the survey

3.1 Survey Design

All the visualizations have been created in d3.js and embedded in qualtrics survey. Fig 1 is an example of the static visualization created for area encoding. We think that the variation of speeds in the motion will help us to infer the optimum speed for the nearest approximate encoding in addition to the understanding of variation of user encodings with different speeds.

We used qualtrics portal to create the survey and analysed the responses. We applied linear statistical regression to comprehend if varying speeds causes any significant differences.

We selected the following encodings which covered the entire spectrum of Cleveland and McGill's visual encoding rankings, thus providing a holistic view of the picture.

1) Position
2) Area
3) Length
4) Color
5) Angle

Each of these encodings was tested as a static visualization and as a visualization with motion. Further, three speeds were tested for all visualizations that included motion encoding. Each task asked the user to give their best estimate about the change in the visual encoding being tested. For instance, one of the questions asked the user the size of a circle after a circle with initial value 100 reduced to a smaller size x. Thus, we achieved our main goal of testing the difference in perception of quantitative values in static and moving visualizations.

The survey itself took around 12 minutes to complete. We received 36 responses. We filtered these down to 26 based on completeness of the answers and their validity in required terms. All the responses we got were from undergraduate or graduate students with a background in computer science.

3.2 Evaluation

The user's responses for every individual visualization for different encodings was recorded in the survey. Since the base values were different in different visualizations for a particular encodings, error values were calculated by subtracting the users input with the correct encoding value. Then, median of this error of the user's input was used rather than mean as mean can be significantly influenced by a few outlier values. A one-way ANOVA was conducted on the error values to find out if the variation in motion was significantly different or not for individual encodings. Then in the post-hoc analysis, Bonferroni Correction was used to analyze the pairwise t-test.

4 Results

4.1 Survey

Link to the survey : https://ufl.qualtrics.com/jfe/form/SV_bHGa2Gn7YKLtw2h

4.2 Evaluation Results

Following are the results obtained by using the proposed evaluation methodology for individual encodings:

1) Position

The position encoding had four visualizations for static (A), slow (B), medium (C) and fast (D) motion. Table 1 presents the actual quantitative values of visualizations which users were asked to predict, the median of the responses and the median of the error.

Table 1: Median and median error of users' responses for position encoding

	Static (A)	Slow (B)	Medium (C)	Fast (D)
Actual Values	5	2	2.33	2.33
Median of Responses	4	2	3	3
Median Error	-1	0	0.67	0.67

A one-way ANOVA was conducted to determine if the mean error of various quantitative encodings for position differed with the motion speed.

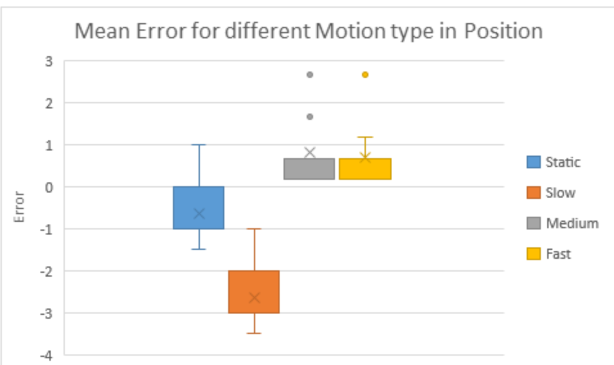

Figure 2: Mean Error for different motions in Position

We find that one-way ANOVA is **statistically significant** (F = 134.09, df = 3, 88, p = 1.1e-16), the effect size is rather large (η2 = .82; suggesting about 82% of the variance of error rate of various quantitative encodings for position is due to motion. The medians (Table 1) and standard deviations of the error rate for the various quantitative encodings of position were as follows: -1 (SD = 0.61) for the static motion, -0 (SD = 0.61) for the slow motion, 0.67 (SD = 0.74) for the medium motion, and 0.67 (SD = 0.68) for the fast motion. The medians and box-plot (Figure 2) suggest that with increase in motion, the median error first decreases. However there is no conclusive trend as it increases from slow to medium/fast and further evidence is required to make appropriate conclusions. The p-value corresponding to the F-statistic of one-way ANOVA is lower than 0.05, suggesting that the one or more treatments are significantly different.

A post-hoc analysis was conducted for pairwise t-test with Bonferroni Correction and **all the pairs (except one) were found to be statistically significant** (see table below). In pairwise comparison A-B (t=10.18, p = 1.33e-15), A-C (t=7.45, p = 3.6E-15), A-D (t=6.78, p = 7.68e-15), B-C (t=17.63, p = 0) and B-D (t=16.96, p =0) were statistically significant at alpha = 5%.

treatments pair	Bonferroni TT-statistic	Bonferroni p-value	Bonferroni inference
A vs B	10.1824	1.3323e-15	** p<0.01
A vs C	7.4508	3.6248e-10	** p<0.01
A vs D	6.7868	7.6890e-09	** p<0.01
B vs C	17.6332	0.0000e+00	** p<0.01
B vs D	16.9691	0.0000e+00	** p<0.01
C vs D	0.6641	3.0503022	insignificant

2) Length

The length encoding had three visualizations for static (A), slow (B), and fast (C) motion. Table 2 presents the actual quantitative values of visualizations which users were asked to predict, the median of the responses and the median of the error.

Table 2: Median and median error of users' responses for length encoding

	Static (A)	Slow (B)	Fast (C)
Actual Values	40	30	30
Median of Responses	40	20	25
Median Error	0	-10	-5

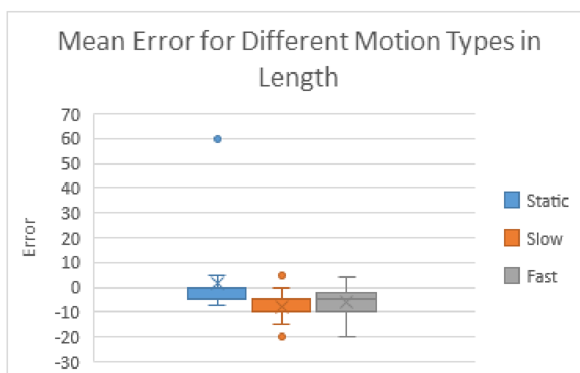

Figure-2: Mean error for different motions in Length

A one-way ANOVA was conducted to determine if the mean error of various quantitative encodings for length differed with the motion speed.

We find that one-way ANOVA is **statistically significant** (F = 7.22, df = 2, 66, p = 0.0015), the effect size is rather large (η2 = .17; suggesting about 17% of the variance of error rate of various quantitative encodings for position is due to motion. The median (Table 2) and standard deviations of the error rate for the various quantitative encodings of position were as follows: 0 (SD = 13.14) for the static motion, -10 (SD = 5.75) for the slow motion, and -5 (SD = 5.65) for the fast motion. The medians and box-plot (Figure 2) suggest that with increase in motion, the median error rate increases. The p-value corresponding to the F-statistic of one-way ANOVA is lower than 0.05, suggesting that the one or more treatments are significantly different.

A post-hoc analysis was conducted for pairwise t-test with Bonferroni Correction and **only one out of the three pairs were found to be statistically insignificant** (see table below). In pairwise comparison A-B (t=3.64, p = 0.001) and A-C (t=2.75, p = 0.022) were statistically significant at alpha = 5%.

treatments pair	Bonferroni TT-statistic	Bonferroni p-value	Bonferroni inference
A vs B	3.6471	0.0015724	** p<0.01
A vs C	2.7533	0.0228401	* p<0.05
B vs C	0.8939	1.1239224	insignificant

3) Angle

The angle encoding had four visualizations for static (A), slow (B), medium (C) and fast (D) motion. Table 3 presents the actual quantitative values of visualizations which users were asked to predict, the median of the responses and the median of the error.

Table 3: Median and median error of users' responses for angle encoding

	Static (A)	Slow (B)	Medium (C)	Fast (D)
Actual Values	30	45	60	60
Median of Responses	40	47	60	60
Median Error	10	2	0	0

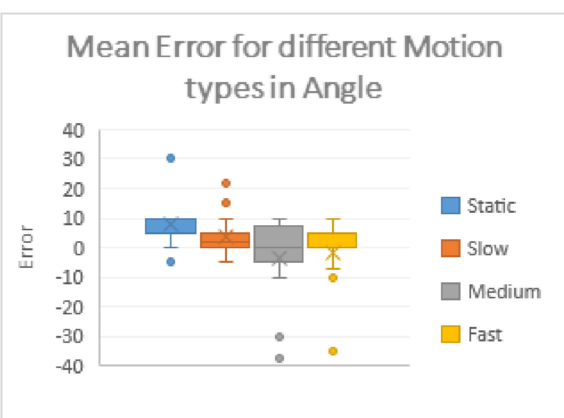

Figure-3: Mean error for different motions in Angle

A one-way ANOVA was conducted to determine if the mean error rate of various quantitative encodings for angle differed with motion speed.

We find that one-way ANOVA is **statistically significant** (F = 5.06, df = 3, 88, p = 0.0028), the effect size is rather large (η2 = .14; suggesting about 14% of the variance of error rate of various quantitative encodings for angle is due to motion. The medians (Table 3) and standard deviations of the error rate for the various quantitative encodings of angle were as follows: 10 (SD = 6.64) for the static motion, -2 (SD = 6.31) for the slow motion, 0 (SD = 16.23) for the medium motion, and 0 (SD = 11.63) for the fast motion. The medians and box-plot (Figure 3) suggest that with increase in motion, the median error rate decreases. The p-value corresponding to the F-statistic of one-way ANOVA is lower than 0.05, suggesting that the one or more treatments are significantly different.

A post-hoc analysis was conducted for pairwise t-test with Bonferroni Correction and **4 out of 6 pairs were found to be statistically insignificant** (see table below). In pairwise comparison A-C (t=3.48, p = 0.004) and A-D (t=2.91, p = 0.03) were statistically significant at alpha = 5%.

treatments pair	Bonferroni TT-statistic	Bonferroni p-value	Bonferroni inference
A vs B	1.2211	1.3519422	insignificant
A vs C	3.4820	0.0046623	** p<0.01
A vs D	2.9118	0.0273185	* p<0.05
B vs C	2.2610	0.1573580	insignificant
B vs D	1.6907	0.5665972	insignificant
C vs D	0.5703	3.4196599	insignificant

4) Area

The area encoding had four visualizations for static (A), slow (B), medium (C) and fast (D) motion. Table 4 presents the actual quantitative values of visualizations which users were asked to predict, the median of the responses and the median of the error.

Table 4: Median and median error of users' responses for area encoding

	Static (A)	Slow (B)	Medium (C)	Fast (D)

Actual Values	25	20	20	40
Median of Responses	10	10	16	30
Median Error	-15	-10	-4	-10

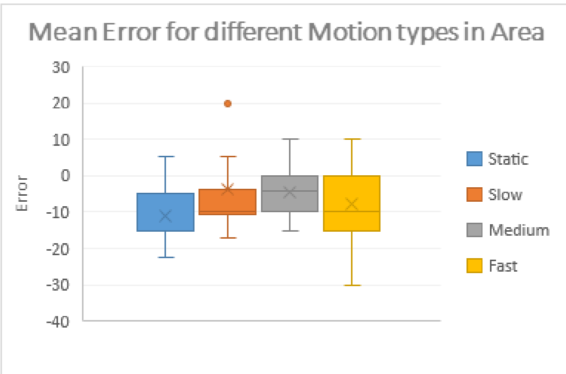

Figure-4: Mean error for different motions in Area

A one-way ANOVA was conducted to determine if the mean error rate of various quantitative encodings for area differed on the level of motion.

We find that one-way ANOVA is **not statistically significant** (F = 1.62, df = 3, 84, p = 0.19), the effect size is rather large ($\eta 2$ = .054; suggesting about 5% of the variance of error rate of various quantitative encodings for area is due to motion. The median (Table 4) and standard deviations of the error rate for the various quantitative encodings of area were as follows: -15 (SD = 7.6) for the static motion, -10 (SD = 18.4) for the slow motion, -4 (SD = 6.3) for the medium motion, and -10 (SD = 11.3) for the fast motion. The medians and box-plot (Figure 4) suggest that with increase in motion, the median error rate decreases. However there is no conclusive trend as it increases from medium to fast and further evidence is required to make appropriate conclusions. The p-value corresponding to the F-statistic of one-way ANOVA is higher than 0.05, suggesting that the treatments are not significantly different for that level of significance. Even though the data does not suggest the presence of significantly different treatment pairs in one-way ANOVA, we proceed with the multiple comparison tests. In some instances, a Bonferroni test of a small set of pairs might show significance, even though 1-way ANOVA suggests that there is too much noise and randomness in the data.

A post-hoc analysis was conducted for pairwise t-test with Bonferroni Correction and **each pair was found to be statistically insignificant** (see table below).

treatments pair	Bonferroni T-Statistic	Bonferroni p-value	Bonferroni inference
A vs B	1.9602	0.3196989	insignificant
A vs C	1.7826	0.4695771	insignificant
A vs D	0.9199	2.1617052	insignificant
B vs C	0.1776	5.1566633	insignificant
B vs D	1.0404	1.8068804	insignificant
C vs D	0.8628	2.3443637	insignificant

5) Color

The color encoding had four visualizations for static (A), slow (B), medium (C) and fast (D) motion. Table 5 presents the actual quantitative values of visualizations which users were asked to predict, the median of the responses and the median of the error.

Table 5: Median and median error of users' responses for color encoding

	Static (A)	Slow (B)	Medium (C)	Fast (D)
Actual Values	60	30	50	50
Median of Responses	10	10	16	30
Median Error	-15	-10	-4	-10

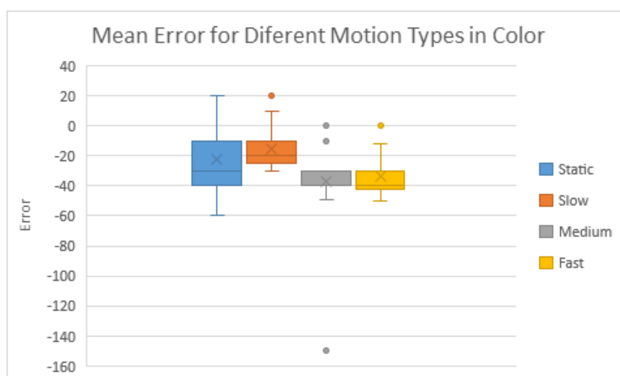

Figure-5: Mean error for different motions in Color

A one-way ANOVA was conducted to determine if the mean error rate of various quantitative encodings for color differed on the level of motion.

We find that one-way ANOVA is **statistically significant** (F = 5.4, df = 3, 88, p = 0.0019), the effect size is rather large ($\eta 2$ = .15; suggesting about 15% of the variance of error rate of various quantitative encodings for color is due to motion. The median (Table 5) and standard deviations of the error rate for the various quantitative encodings of color were as follows: -15 (SD = 22.70) for the static motion, -10 (SD = 15.03) for the slow motion, -4 (SD = 28.84) for the medium motion, and -10 (SD = 14.39) for the fast motion. The medians and box-plot (Figure 5) suggest that with increase in motion, the median error rate decreases. However there is no conclusive trend as it increases from medium to fast and further evidence is required to make appropriate conclusions. The p-value corresponding to the F-statistic of one-way ANOVA is lower than 0.05, suggesting that the one or more treatments are significantly different. The Bonferroni comparison tests would likely identify which of the pairs of treatments are significantly different from each other.

A post-hoc analysis was conducted for pairwise t-test with Bonferroni Correction and 4 out of 6 pairs were found to be statistically insignificant (see table below). **In pairwise comparison, B-C (t=3.56, p = 0.003) and B-D (t=2.94, p = 0.02) were statistically significant at alpha = 5%.**

treatments pair	Bonferroni TT-statistic	Bonferroni p-value	Bonferroni inference
A vs B	1.1112	1.6169679	insignificant
A vs C	2.4531	0.0968114	insignificant
A vs D	1.8381	0.4165490	insignificant
B vs C	3.5643	0.0035526	** p<0.01
B vs D	2.9493	0.0244742	* p<0.05
C vs D	0.6150	3.2407623	insignificant

5 Conclusion And Future Work

5.1 Conclusion

- We observed significant difference in perception when we add motion to the visuals for encodings angle, position, color and length.
- Our tests did not show any significant difference in perception of area for static and moving visuals.
- There is no conclusive trend to differentiate between the accuracy for the static or moving visuals. We also could not find a trend suggesting that a certain speed is better than the others.

5.2 Challenges

- Recording response time is a crucial element of studies such as this one. Although given the varying times that our visualizations take to load, we cannot have a fair response time estimate.
- Our limited knowledge of evaluation tools and methodologies was a challenge we needed to overcome.
- There is a possibility of results being different if the survey was conducted in a controlled setting.

5.3 Future Work

The scope of research in this field is extremely vast. Through our research we came across a lot of research about visualizations but very few focused-on motion let alone adding motion to encodings. There is no dearth of opinions in this domain, but few can be formulated and confirmed through extensive research and reliable surveys.

In [7] Ma et al. present the analysis of a sports visualization system called LucentVision which uses real time analysis to create trajectories and provide visualization options to the viewer. Since sports is a huge medium that is consumed by millions worldwide, we can further explore the possibility of motion in sports related visualizations such as in LucentVision. Through our project we aim to determine which types of data can be visualized with motion encoding, and apreal time sports analytics would be a good arena to look into.

Acknowledgements

We wish to thank Dr. Eric Ragan for his valuable guidance and unwavering belief in us and our project. We also appreciate the anonymous study participants for their feedback and timely submissions.

References

[1] W. S. Cleveland and R. McGill. Graphical perception: Theory, experimentation, and application to the development of graphical methods. Journal of the American statistical association 79(387):531-554, 1984

[2] J. Heer and M. Bostock. Crowdsourcing graphical perception: using mechanical turk to assess visualization design. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,203-212, ACM, 2010

[3] H. Romat, C. Appert, B. Bach, N. Henry-Riche, and E. Pietriga. Animated edge textures in node- link diagrams: A design space and initial evaluation. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems pages 187:1–187:13, New York, NY, USA, 2018. ACM.

[4] J. Driver, P. McLeod, and Z. Dienes. Motion coherence and conjunction search: Implications for guided search theory. Perception & Psychophysics, 51(1):79–85, 1992.

[5] L. Bartram, C. Ware, and T. Calvert. Moticons: detection, distraction and task. International Journal of Human-Computer Studies, 58(5):515–545, 2003.

[6] E. Velloso, M. Carter, J. Newn, A. Esteves, C. Clarke, and H. Gellersen. Motion correlation: Selecting objects by matching their movement.

[7] A study of using motion for comparative visualization, Chien-Hsin Hsueh ; Jia-Kai Chou ; Kwan-Liu Ma

[8] G. Pingali, A. Opalach, Y. Jean and I. Carlbom, "Visualization of sports using motion trajectories: providing insights into performance, style, and strategy," Proceedings Visualization, 2001. VIS '01., San Diego, CA, USA, 2001

[9] J Nielsen, Nielsen Norman Group (1995) – How to conduct a heuristic evaluation

[10] Bahador Saket, Arjun Srinivasan, Eric D. Ragan, Alex Endert, Journal of Latex Class Files (2015) – Evaluating Interactive Graphical Encodings for Data Visualization

[11] Daniel E. Huber Christopher G., IEEE (2005) – Healey Visualizing Data with Motion

[12] Aniket Kittur, Ed H. Chi, Bongwon Suh, CHI proceedings (2008) – Crowdsourcing User Studies with Mechanical Turk

[13] Bartram, L., & Nakatani, A. (2010, November). What makes motion meaningful? Affective properties of abstract motion. In 2010 Fourth Pacific-Rim Symposium on Image and Video Technology (pp. 468-474). IEEE.

[14] Helen Chen, Sophie Engle, Alark Joshi, Eric D. Ragan, Beste F. Yuksel, and Lane Harrison. 2018. Using Animation to Alleviate Overdraw in Multiclass Scatterplot Matrices. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18). ACM, New York, NY, USA, Paper 417, 12 pages. DOI: https://doi.org/10.1145/3173574.3173991

[15] Bartram, L., & Ware, C. (2002). Filtering and Brushing with Motion. Information Visualization, 1(1), 66–79. https://doi.org/10.1057/palgrave.ivs.9500005

[16] Daniel Archambault and Helen C. Purchase. 2016a. Can Animation Support the Visualization of Dynamic Graphs? Information Sciences 330, Supplement C, SI Visual Info Communication (February 2016), 495–509. DOI: http://dx.doi.org/10.1016/j.ins.2015.04.017