**Spatial Distortion Approach to Traffic Congestion Visualization**

Capstone Research Report

Zbyněk Stara

New York University Abu Dhabi

zs633@nyu.edu

**Abstract.** The field of traffic congestion visualization relies heavily on color-coded topographical maps; little work has been done, however, to examine the merits and possible weaknesses of this approach, and to evaluate its performance in comparison to alternative visualizations. This paper presents a new approach to visualization of traffic congestion that uses spatial distortion instead of color to present traffic information in a fundamentally different manner. A user study was employed to empirically evaluate the ability of each of the two visualization approaches to influence the users’ routing decisions in a given traffic network.

**1 Introduction**

According to Arnott and Small, “time spent ensnarled in traffic is not simply time wasted; for most of us, it is time miserably wasted.”[[1]](#footnote--1) They offer a simple calculation to estimate the magnitude of the problem. About one third of all driving in metropolitan areas takes place in congested conditions, during which average speed decreases to half of the traffic segments’ free flow value.[[2]](#footnote-0) Even without considering the cost of additional fuel, accidents, air pollution, and other issues due to congestion, the economic finding that drivers are willing to spend “about 1.33 USD to save 10 minutes [of] travel time” puts the annual cost of driving delays in the United States at 48 billion USD.[[3]](#footnote-1)

Routing decisions of individual drivers depend on the current levels of congestion in their traffic network. However, if only limited information is available to drivers, either because of a lack of traffic data or the impossibility to interpret available information, the quality of drivers’ routing decisions necessarily suffers.[[4]](#footnote-2) The quality of a traffic visualization is as important as the accuracy of the reported traffic data; thus, it is necessary to evaluate the degree to which a given traffic visualization allows its users to make more optimal routing decisions in a variety of traffic scenarios.

The field of traffic congestion visualization currently relies heavily on topographical maps with color overlays – featured perhaps most prominently in Google Maps Traffic. Alternative traffic visualizations have been proposed, coming both from the field of traffic visualization,[[5]](#footnote-3) and from other fields;[[6]](#footnote-4) however, very little work has been done to try to empirically evaluate the efficiency of any of these traffic visualizations.

This paper proposes spatial distortion as a novel approach to visualization of traffic. A formal description of the algorithm is then followed by a description of a user study[[7]](#footnote-5) designed to compare this visualization to the technique of color-line overlays. Particular attention was paid to the ability of each visualization to influence the routing decisions of drivers. By examining the algorithms in this manner, this paper aims to identify which traffic situations are most efficiently presented in either visualization. By providing quantified information, this work aims to inform understanding of traffic visualization, and thus help reduce “the immense losses of time and money that could be spent more efficiently in most any other way.”[[8]](#footnote-6)

**2 Background**

Visualization is, at its core, the science of representing problems and their solutions according to a logical structure that may or may not be immediately apparent. Visualization techniques identify constituent elements of the problem, codify them as points in one or several data sets, and represent the data in a visual form that is intelligible for human readers.[[9]](#footnote-7)

The nature of this visual representation varies significantly depending on the identified relationships among data points. However, what matters is not the specific visual technique *per se*, but the nature of the information that may be gleaned by the observer from the presented data. The most successful visualizations are not those that merely illustrate the problem and/or its solution, but those that reveal information that might not have been accessible without the visual representation.

An important measure of the effectiveness of a visualization is whether it helps influence the user’s behavior. This is especially applicable in the field of traffic visualization and congestion prediction; better visualizations may allow users to better avoid highly congested areas of the traffic network, and thus reduce the amount of time and money lost due to traffic, as well as the amount of pollution generated by road vehicles. The state of the art in traffic congestion visualization is the technique of overlaying topographical maps of traffic networks with color-coded lines that represent congestion. More formally, this visualization approach allows the user to evaluate the level of congestion in the traffic network in terms of the level of congestion of constituent roads.

JamBayes, a traffic information and prediction system developed by Horvitz et al.,[[10]](#footnote-8) uses a visual interface that assigns colors to road segments in a simplified map of the Seattle road network according to current traffic congestion, and can be seen in Figure 1.

The color-line information is supplemented with clock illustrations that communicate expected time delay at crucial points of the traffic network more explicitly. Additionally, exclamation marks are employed to alert the user to potentially surprising traffic conditions, identified by comparing the current traffic conditions to those expected by the visualization’s underlying traffic model.

The group of color-line visualizations also includes perhaps the most widely used system for visualization of traffic – Google Maps Traffic. Similar to JamBayes, the visualization employs the topographical map approach in tandem with a discrete colormap to overlay the current state of traffic at main thoroughfares of supported cities; an example is presented in Figure 2. The Google model presents several innovations over the JamBayes model. To eliminate the need for introducing additional graphical elements to a map interface that is already extremely busy, the Google model abandons the clock-icon features of Horvitz’s model, and partially replaces them with an extension to the conventional colormap. In addition to the conventional color palette of green, orange and red, the Google system includes the dark red color to indicate that traffic congestion in a segment is even worse than that signified by the color red.[[11]](#footnote-9) Primarily because of its ubiquity and familiarity to users, the Google Maps Traffic visualization was chosen as a fair comparison for the spatial-distortion approach presented in this paper.

The spatial-distortion visualization of traffic congestion was conceived as a traffic-specific application of the ideas behind the technique of the cartogram, introduced by Gastner and Newman,[[12]](#footnote-10) and presented in Figure 3.

Cartograms begin with an ordinary map; however, the visualization algorithm then distorts the map based on a variable of interest – e.g. population density. In so doing, the visualization sacrifices topographical accuracy for a clearer illustration of the visualized variable – the higher the value, the more prominent its area becomes in the visualization. Conversely, if the value of the variable is relatively small, its corresponding area in the cartogram is shrunk proportionately. As they encode information with position and area rather than color, cartograms prove more easily decipherable for their users than equivalent color-based visualizations.[[13]](#footnote-11) This comes with the qualification that the features of the original map must be discernible in the cartogram to supply visual clues to the reader.[[14]](#footnote-12)

**3 Model**

The spatial-distortion model uses traffic data to determine the degree to which a particular traffic segment should be emphasized. The more congestion a traffic segment experiences, the more it distorts geography around itself; individual drivers are thus able to quickly identify the worst routes,[[15]](#footnote-13) and are able to opt instead for a detour through the least emphasized areas.

While applying the technique of the cartogram to traffic is very appealing from a conceptual standpoint, its implementation proved challenging. The question became “How can information about the current traffic conditions in the network be applied to the input map in such a way that the output map presents traffic congestion in an internally consistent, intuitive, and helpful way?” It is important to remember that in order for them to be understandable by human users, cartograms still rely on visual cues on the boundaries of distorted areas; the distortion must not be too extreme. The design process required several iterations; the visualization algorithm underwent several fundamental paradigm shifts during the process.

**3.1 Image Distortion Approach**

The most obvious choice for a distortion algorithm is one that distorts the map at the level of the input image. Rather than iterating through every edge and every node of the traffic network and changing their positions, the algorithm would magnify and shrink areas of the image using image-editing algorithms. Despite its apparent simplicity, the option was abandoned after it proved too difficult to program a custom image-editing software to respond to changing traffic input.

It thus proved necessary to effect the spatial distortion by editing the nodes and edges of the input map directly. This can be done in two ways – either by changing the lengths of edges and adjusting node positions accordingly, or by changing the positions of nodes and adjusting the lengths of edges as necessary.

**3.2 Traffic Edge Distortion Approach**

Modifying the visualization by changing the lengths of edges is the approach used by Spring Network simulations. Each edge in the traffic network is interpreted as a spring and assigned a certain potential energy; the algorithm then simulates the network in an iterative way as it reshuffles itself in order to minimize the potential energy stored in each individual edge.[[16]](#footnote-14) Graphviz’s Neato program is designed to provide Spring Network functionality for graph visualization,[[17]](#footnote-15) so it was examined for potential use as a space distortion algorithm. Applied to the problem of traffic visualization, the desired length of each edge was made to correspond to the product of its original length and the observed traffic congestion – the heavier the traffic, the shorter the segment. Less congested roads were thus expected to dominate the output map; more congested maps would have been made less prominent by the shortening, and thus less attractive as possible paths.

Unfortunately, the spatial-equalization property of spring networks proved problematic; that is, areas of the network with a higher density of roads always experienced more compression than areas with fewer interconnected roads, regardless of local traffic situation. Because the visualization operated under the assumption that shorter roads were more congested, this inevitable side effect introduced an unacceptable bias in user perception of traffic congestion, illustrated in Figure 4. However, since the spatial-equalization property is inherent in all edge-based iterative distortion approaches, an alternative approach proved necessary.

**3.3 Traffic Node Distortion Approach**

The node distortion approach moves traffic nodes junctions to new positions as determined by a distortion function. Consequently, the length of traffic edges becomes dependent on the movement of their endpoint nodes. It is important to note that since traffic data records the traffic state of edges in a traffic network instead of nodes, a decision had to be made about the most appropriate way to determine traffic congestion at nodes from available data. In the end, the level of congestion at a node was calculated as the average of traffic congestion of the node’s incoming edges.

The following simple mathematical model of spatial distortion depending on traffic congestion was then assumed:



*a* is the distorting node, and *b* is the affected node, and *r* is the distortion radius. For the purposes of the model, the x- and y-coordinates of *a* and *b* may be considered separately; the above formula thus stands for a pair of formulas



The value of *r* is the radius of the effect of distortion around the distorting node *a*. Thus, when considered separately in x- and y-coordinates, the values of *rx* and *ry* are such that they describe a circle of radius *r* around *a*:



The radius of the distortion effect is found according to



where



*ta* is the *traffic slowdown* at node *a*; it encodes information about how much slower traffic at node *a* is in the observed traffic state when compared to the free-flow speed limit. Meanwhile, *R* is a constant maximum radius of distortion. The value of *R* has to be carefully chosen with respect to the density of the visualized map and variation in the level of traffic congestion to ensure that the final map is recognizable by the user. The size of the distortion effect around each node *a* thus depends on the level of traffic congestion at *a* – the more congested *a* is, the more it distorts the space around itself. Therefore, when *ta* = 1 (that is, traffic at node *a* moves 100% slower than it would in free-flow condition; it does not move at all), then *r* = *R*, and the effects of distortion due to *a* reach to the maximum distance. Conversely, if no congestion is observed at node *a* (*ta* = 0, since traffic moves 0% slower compared to the free-flow speed), *r* = 0, then node *a* does not have any distorting effect at all.

The outcome of the distortion function *f* is to produce a set of *b*’ coordinates –the new position of *b* after being affected by *a*. The distortion function *f* is iteratively applied to every combination of *a* and *b*; since *a* and *b* are drawn from the same set of nodes, the algorithm determines the distorting effect of every node on every other node in a given traffic map. The distortion is not dynamic, however. At every step, the original coordinates of *a* and *b* are considered, with the results of the *f* function adding up to the provisional coordinates of *b*’. It is only after all points *a* and *b* are considered by the distortion algorithm that the traffic nodes are drawn at their new *b*’ coordinates. This produces a distortion effect.

This still leaves open the question of how exactly the distortion should look. Thus, another metaphor was brought in, this time from physics. It was recognized that the desired result of the map looks very much like the curvature of spacetime due to gravity in physics visualizations, similar to the rubber-sheet visualization presented in Figure 5. After all, the more traffic congestion is observed at a node, the more distortion it effects – just like gravity does in reality. The more mass an object has, the more it distorts spacetime around itself.

While the metaphor in this case is that of antigravity – repulsion of the surrounding matter depending on mass of the distorting object – this could be achieved by simply switching the sign of the gravity constant. This is to evoke in the user the impression of having to climb a hill in order to traverse congested areas in the visualized maps. Just as it often requires more energy to follow a straight-line path over a hill than it takes to take a detour following the contours of a landscape, so, the thinking goes, it is advantageous to take a detour around areas of heavy traffic. In visual terms, instead of observing the gridded rubber sheet of the gravity metaphor from the top and seeing the constantly-spaced lines converge as they approach the center of mass, the effect would be observed from below, and seeing the constantly-spaced lines bulge around the distorting object. The distorting nodes are not black holes, drawing surrounding nodes in; they are charged particles that repel other particles according to their charge.

Having this theoretical framework in mind, it proved challenging to find the correct spatial distortion function – one, that is, that produces the illusion of *b* nodes around a given distorting node *a* being projected onto the surface of an imaginary solid protruding from the plane towards the observer. Multiple spatial distortion functions were examined, drawing from different fields of physics that involve spatial distortion, with the goal of finding one that would produce a hemispherical, bubble-like surface around *a* onto which nodes *b* would be projected. That being said, it is important to note that although the desired effect is the illusion of projecting nodes onto a surface of a three-dimensional object centered on *a*, this had to be achieved with displacement of *b* nodes to *b*’ in two dimensions; the metaphor employs an orthogonal projection, which does not take the third axis into account. Apart from the gravity function, the distortion functions examined included a simulation of refraction – the bending of light at the boundary of different transparent surfaces – as well as the logistic curve of population growth, and several geometric curves including the cosine curve, linear curve, and curves describing the cross-section of a cone, a sphere, and a convex object.

To visually represent the effects of different distortion functions, a distortion measure *d* was created:



where



*s* is the distance between *a* and *b* in the dimension considered (remember that *f* stands in for both *fx* and *fy*). Thus, the coordinates of *a* and *b* are related to the coordinates of *b*’ due to function *f*. Plotting *d* against *s*’, where *s*’ is the distance between *f* and *a*



then provides a visual representation of the cross-section of the imaginary solid onto which *b*’ nodes are displaced. Using this measure, the task was now to discover a distortion function that would produce the desired hemispherical distortion. In total, this process required eight iterations; the different functions investigated are detailed in Appendix A. The visual characteristics of the distortion functions are presented concisely in Figure 6.

It was determined that among the eight functions considered, the *spherical* function produces the most intelligible hill-like distortion. The function has the following piecewise definition:



The function’s desirable distortion properties are the reason why it was selected as the representative of the spatial-distortion approach to traffic visualization, to be compared with the color-line visualization in the user study. Several additional sample traffic networks were visualized with the *spherical* algorithm for illustration; they are presented in Figure 7.

**4 Evaluation**

To compare the spatial distortion visualization represented by the *spherical* model to the color-line visualization represented by Google Maps Traffic, a user study was performed.

In the absence of authoritative material on the visualization produced by Google, the color-line visualization used in Google Maps Traffic was emulated as closely as possible by the author of this study. Biersdorfer provides speed ranges that correspond to the four colors in the color-line visualization relative to New York state highway speed limits;[[18]](#footnote-16) these were translated into corresponding slowdown ratios for the purposes of the study. Meanwhile, the RGB values of the four colors were taken directly from the Traffic map overlay in the Google Maps web interface.

A set of nine different traffic situations was visualized with both algorithms and presented to users, who were asked to draw on the maps the path between three pairs of points that they think would take them the shortest time to traverse in a car. Their answers were then compared to the actual best path in a given situation, and the accuracy of their answer was recorded. In addition to accuracy, the time it took users to finish each map was noted as a secondary measure. The users were also asked to fill out a short questionnaire to quantify their sentiment about each visualization, and a set of long-answer questions asked for qualitative suggestions for improvement of both visualization approaches.

**4.1 Traffic Data**

The data for the user study was made to vary along four independent variables – Visualization, map Density, Traffic volume, and Path.

A set of three grid-based engineered traffic maps were created, with different levels of road Density (Dense, Medium, Sparse). Engineered maps were used instead of maps of real cities to ensure, as much as possible, that users base their routing decisions solely on the traffic visualizations alone, rather than on their acquired knowledge of usual traffic conditions in any given city, or individual preferences such as choosing a seafront road over a road in the city center. Additionally, it is easier to alter the density of engineered maps rather than real maps; it is enough to simply remove a pair of edges – there are no complications such as one-way roads, highway onramps, and terrain features. Additionally, the regular nature of the grid maps makes it possible to unambiguously name each node in a way that enables easy addressing.

Furthermore, each Density level required a different value of the Distortion visualization’s maximum radius constant *R*. These had to be discovered empirically, based on two competing criteria. The first was that there must be no *overlaps* in the generated Distortion visualization – this would make interpretation of the map impossible for the users – and that the edges of maps must not be too distorted. (The higher the value of *R*, the farther the edge junctions will be displaced due to the unopposed influence from their neighbors, who are concentrated in only one direction.) The second criterion was that as much traffic information should be visible as possible. Distortion becomes visible only if *r* is higher than the distance between junctions, thus a lower *R* results in less traffic congestion being shown overall; in the extreme case, when *R* is set to be lower than the distance between junctions, no distortion is shown at all, no matter the amount of traffic congestion. The final values of *R* were determined according to a pattern presented in Table 1.

In each of the three map densities, three different levels of Traffic were simulated (High, Medium, Low). SUMO, a free and open-source program for traffic network simulation developed by German Institute of Traffic Research, was used throughout the project to simulate traffic conditions. To enable a controlled evaluation of the effects of each visualization method, the same traffic data was used for spatial distortion visualizations as was for color-line visualizations. Since SUMO allows variation of traffic only in terms of the total number of vehicles, some trial-and error was required to determine the optimal quantity of vehicles for each traffic level. In the end, the following rule was adopted – the High traffic level in each map density has such a number of cars that there are no dark-red road segments in the corresponding color-line visualization. Dark red means that the traffic in a given segment is nearly stopped; this leads to an extreme snowball effect that quickly degenerates into a map-wide gridlock. The aim was to find the maximum number of cars for each map density that does not lead to gridlock. The number of cars in Medium and Low traffic levels were then calculated based on the High traffic level, according to a pattern presented in Table 1.

There are 9 unique combinations of Density and Traffic – 18, if the spatial-distortion and color-line versions are considered separately. They are collectively presented in Figure 8. Each combination was presented to the users three times with three different pairs of start and goal nodes highlighted. The three pairs of path endpoints were the same across all the different map combinations. These endpoints were selected to increase variety in the data – to make sure that what is measured are the properties of each visualization in a given map and traffic level, rather than for one single path – and to force the users to consider different possible traffic scenarios. That is why one of the endpoint pairs connects the light-traffic map edge to the opposite map edge, and another asks the user to consider the best path into the center of the map – which was invariably the most congested area of each map. The third set of endpoints asked the user for a path between two edges of the map, but forced them to go through the center of the map. In total, therefore, there are 54 scenarios for users to consider.

**4.2 User Study Design**

A user study was performed with 32 non-colorblind participants, interviewed individually over the course of approximately one hour. Each user underwent a short introduction to the field of traffic visualization, a short training session, followed by drawing paths on 54 maps, and a short questionnaire. The drawing of paths occupied the majority of a research session time, and was the focus of the study.

During the initial part of the session, the subjects were broadly introduced to the terminology of traffic visualization, and were shown examples of both color-line and distortion visualizations. This was followed by a training session that had two main aims.

First was to ensure that the study subjects would be able to perform the main task correctly – draw a path following edges, and in the correct direction from A to B, and to recognize that there might be several possible valid paths in a grid road layout. Second, since the distortion visualization is not familiar to users as much as color-line visualization due to the ubiquity of Google Maps Traffic, it was explained to the study subjects that they should try to imagine the distortion visualization as a projection of a three-dimensional landscape, with higher hills around more congested nodes. Beyond this piece of information, however, they were not instructed in finding the best path in a distortion visualization; indeed, the primary investigator did not get to know the optimal path for the maps until all interviews were conducted.

After the introduction and the training session, the experiment proceeded to the main task. For each of the 54 maps provided, the users were asked “to draw on the map the path from A to B [they] think would take the shortest time to traverse in a car.” Using a provided marker, the users were asked to physically draw the path they thought would be the best in a given printout of a map according to the prompt. Paper-based approach was chosen despite being labor intensive to oversee and enter into digital form, due to worries that a digital solution – e.g. using an iPad to show paths and record users’ responses – might be confusing and distracting to study subjects. The maps were presented in a random order established before the first study session, with alternating distortion and color-line visualizations. The investigator noted the time it took the participant to provide a path in each map.

Finally, having marked all 54 visualization, the users were asked to fill out a short questionnaire. It had five question areas. The users were asked to evaluate the closeness of their paths to the ideal, the closeness of others’ paths to the ideal, the similarity of their paths to the paths provided by others, how useful they perceived the visualizations, and how useful others perceived the visualizations. Each question area had three questions – first asked to the user to provide a score for the color-line visualization on a 5-point scale, the second asked the user to provide the same for the distortion visualization, and the last question asked the user to directly compare the two visualizations. Two long-answer boxes gave the participants the opportunity to suggest improvements to each visualization.

**4.3 Output Data**

The primary output data from the experiment were the paths drawn by the participants onto each of the 54 traffic visualizations. These 54 visualizations represent 54 unique combinations of the four dimensions considered: 2 visualizations \* 3 map densities \* 3 traffic levels \* 3 path endpoint pairs. However, for the purposes of the subsequent analysis, the three different Paths were not used as an independent dimension in the subsequent analysis. The resulting tripling of the number of data points in the rest of the independent dimensions usefully increased the number of samples despite the constrained size of the participant pool.

To be able to process data about the paths provided by users, the coordinates of nodes traversed by the users’ paths were transcribed into a computer. Combining the path information with the detailed outputs of the traffic simulation provided by SUMO then determined the time it would take the user to go from point A to point B in a given visualization. The path providing the minimal traversal time, known as the ideal path, was identified for all 27 traffic situations. In each, the ideal path was unique. Both the list of nodes constituting this path and the time required to traverse it were returned. Each of the 54 paths provided by the user was then re-expressed in terms of Accuracy (*A*) calculated as a ratio between *Tuser*, the time it would take to traverse the user provided path in a given traffic visualization, and *Tideal*, the time it would take to traverse the ideal path in the same visualization:



Since, by definition, no user path can be traversed faster than the ideal path, the best possible value of *A* is 1. All other user paths will observe a higher *Tuser* than *Tideal* and *A* will rise accordingly. There is no theoretical upper limit to *A* – the highest observed value was 8.92, although the average value was 1.28 (SD = 0.14).

Effectively, then, *A* is a normalization of *Tuser*. It was necessary to use a measure of normalized time-to-traverse instead of raw *T* values because of the use of the Path dimension to increase the number of samples – since the three different paths have different lengths and thus take different amounts of time to traverse, averaging their raw values would produce a value dominated by the longer paths. This problem disappears when Accuracy is used instead, however.

The Time it took each participant to draw each of the 54 paths was recorded as a secondary measure. Again, the values were normalized; instead of using raw time, each data point was expressed as a Z-score relative to a user’s mean – that is, the number of standard deviations from the mean a given time value was. For this measure, then, lower values are better than higher values.

The checkbox answers to questionnaire prompts were transcribed in terms of the corresponding numerical value on the 5-point scale, starting with 1 and ending with 5. Whenever possible, 1 represents a strongly negative answer to the question, and 5 represents a strongly positive answer. Whenever the question asks the users to compare the two color-line visualizations, 1 was arbitrarily chosen to indicate an answer in favor of the color-line visualization; 5 indicates an answer in favor of the distortion visualization.

The last measure recorded were the long-answer responses – the final part of the questionnaire.

**5 Hypotheses**

A significant difference in Accuracy is expected between Distortion and Color-Line visualizations. Since the Color-Line visualization is based on Google Maps Traffic – currently the dominant method of traffic congestion visualization – it is expected to benefit from the effects of users’ familiarity; this is expected to show in lower (that is, better) Accuracy for the Color-Line visualization. Nevertheless, it is expected that the Distortion approach will be more successful in several of the specific combinations of Density and Traffic, particularly in two extremes. First, the Distortion visualization is expected to be more successful in situations with fairly uniform distribution of traffic – where it may be able to communicate more detailed information than the Color-Line Visualization – and second, in situations with a single clearly defined local center of congestion – where it may more accurately capture the user’s attention and more readily present alternative routes.

A similar significant difference is expected between Distortion and Color-Line visualizations in the secondary measure of Times. The Color-Line visualization is expected to record lower (better) normalized Time values in the general case, while the Distortion visualization is expected to prevail in uniformly-congested and locally-congested maps.

The questionnaire results are expected to indicate that study participants are more familiar with the Color-Line visualization; the Color-Line visualization is expected to prevail in all questions.

**6 Results**

**6.1 Accuracy**

The Accuracy measure was analyzed using a 3 x 3 x 2 ANOVA with replication. An overview of the results is presented in Table 2.

Accuracy results are presented graphically in Figure 9. A significant main effect was found for Density, F(2,558) = 137.34, p = 0, indicating that participants were more accurate in Dense maps than in Sparse maps (M = 1.09 v 1.13 v 1.62). A significant main effect was found for Traffic, F(2,558) = 91.07, p = 0, indicating that participants were less accurate in High traffic than they were in Low traffic level (M = 1.56 v 1.15 v 1.13). However, no significant main effect was found for Visualization, F(1,558) = 3.49, p = 0.06. This indicates that the observed difference between the Distortion and Color-Line visualizations (M = 1.31 v 1.25) was not statistically significant.

The only significant two-way interaction was Density \* Traffic, F(4,558) = 72.29, p = 0. A 3 x 3 ANOVA with replication was run to analyze this interaction; full results are presented in Table 3. In the collapsed table, Density recoded F(2,279) = 97.81, p = 0; Traffic recorded F(2,279) = 64.86, p = 0; and the interaction had F(4,279) = 51.49, p = 0. From the graph in Figure 10, it can be seen that the interaction effect stems primarily from the extremely disproportionate Accuracy value of 2.40 in Sparse-High maps.

The three-way interaction Density \* Traffic \* Visualization, F(4,558) = 1.50, p = 0.20, did not qualify for significance.

Although neither Density \* Visualization nor Traffic \* Visualization qualified for significance, the expected trends in the Visualization dimension can be gauged from their data – with the important caveat that these results require further experimental validation. The data for the two interactions is presented in Table 4. It appears that the Color-Line visualization performs better than the Distortion visualization in the Medium and Sparse density maps, and in High and Medium traffic. The results for Dense maps and Low traffic are inconclusive.

**6.2 Times**

Similarly to Accuracy, the normalized Times measure was analyzed using a 3 x 3 x 2 ANOVA with replication. An overview of results is presented in Table 5.

Times results are presented graphically in Figure 11. A significant main effect was found for Density, F(2,558) = 104.11, p = 0, indicating that participants were able to spend less time with Dense maps than with Sparse maps (M = -0.35 v 0.01 v 0.35). A significant main effect was found for Traffic, F(2,558) = 30.91, p = 0, indicating that participants were slower to solve High traffic level maps than they were Low traffic level maps (M = 0.22 v -0.10 v -0.12). A significant main effect was found for Visualization, F(1,558) = 5.39, p < 0.02. This indicates that users require more time to use the Distortion visualization than to use the Color-Line visualization (M = 0.05 v -0.05).

The interaction between Density and Traffic proved significant, with F(4,558) = 7.24, p < 0.0001. A 3 x 3 ANOVA with replication was run to analyze this interaction; full results are presented in Table 6, while its graph is presented in Figure 12. In the collapsed table, Density recoded F(2,279) = 93.37, p = 0; Traffic recorded F(2,279) = 29.93, p = 0; and the interaction had F(4,279) = 6.91, p < 0.0001. It can be seen that the interaction effect stems from the fact that each Density level exhibits a different pattern among Traffic levels. In Dense maps, High traffic was the slowest, followed at a distance by Low and then Medium. In Medium maps, all traffic levels recorded similar times. Finally, in Sparse maps, High traffic was followed by Medium and Low.

The three-way interaction Density \* Traffic \* Visualization, F(4,558) = 2.54, p = 0.04, proved significant. Comparing Figure 15 with Figure 17, one can see that the Visualization dimension adds nuance at Medium map density, inverting the relationship between High and Medium traffic level – which also indicates that the Distortion visualization is better than Color-Line for Medium-High maps. Additionally, Color-Line visualization has more prominent time spike at Sparse-High maps, while Distortion has more pronounced effect at Dense-High maps

Although the other interactions did not qualify for significance, the trends in the Visualization dimension can be gauged from their collapsed data – again, with the qualification that these results are tentative. This data is presented in Table 7. It appears that the Color-Line visualization performs better than the Distortion visualization in Dense maps and Medium traffic; other results are inconclusive.

**6.3 Questionnaire**

The participants’ questionnaire responses are presented in Table 8. It can be seen that the color-line visualization earned better scores on average than the spatial-distortion visualization, with a few notable exceptions.

When it comes to scoring the accuracy of one’s own paths, the color-line visualization earned better marks than the distortion visualization with 3.25 v 3.16, but the direct comparison question was exactly tied at 3.00.

When it comes to guessing the accuracy of others’ paths, the color line visualization earned better marks than the distortion visualization with 3.44 v 3.06; in direct comparison, the color-line visualization prevailed with 2.81.

When it comes to guessing the agreement of one’s own paths with others’ paths, the color line visualization earned better marks than the distortion visualization with 3.25 v 3.06; in direct comparison, the color-line visualization narrowly prevailed with 2.97.

When it comes to scoring the usefulness of visualizations to oneself, the color line visualization earned better marks than the distortion visualization with 3.53 v 3.50. However, in direct comparison, the distortion visualization narrowly prevailed with 3.03.

Finally, when it comes to guessing the usefulness of visualizations to others, the color line visualization earned better marks than the distortion visualization with 3.66 v 3.31. In direct comparison, the color-line visualization prevailed with 2.75.

**6.4 Long Answer Questions**

When it comes to suggested improvements to the color-line visualization, participants’ answers fell into several groups.

The majority of study participants (21) complained that it was very hard to see the difference between the color states in the visualization (which distinguish different levels of traffic slowdown, *ta*); Eight participants specifically mentioned the orange and red traffic states, and three mentioned the red and dark red states. One participant remarked that, compared to the other colors, the green stood out too much. Suggested improvements to allay this problem included switching the dark red color for blue, adopting a blue/green/red color palette, and using light and dark colors as a redundant way to encode traffic level. Additionally, the use of different types of lines was proposed, to aid in differentiating the traffic states.

Some people declared that the abundance of red was not friendly to people’s eyes, although some have also added that once they got used to visualization, it got easier. One participant suggested using grayscale.

Four participants requested a more explicit way to communicate the intensity of traffic – either with a smooth color gradient, or with line thickness as a redundant encoding. One participant wanted a legend that would explain the traffic levels. One participant requested an indication of the actual time it would take in traffic to reach one’s destination.

When it comes to the distortion visualization, there was less consensus on what could be improved – also, more people (13) chose not to answer this question than in the case of the color-line visualization (3).

Most fundamentally, five users did not know how to imagine the 3D effect due to distortion. Three users found the distortion confusing. One user complained that undistorted maps are already hard to read. One participant wanted a true 3D simulation.

Six participants proposed adding colors as if the map was a topographical map of a landscape. Four users explained that this is because when roads are missing from the map – as is the case in Medium and Sparse maps – it is impossible to know high ground from low, and thus hard to know if one’s route takes one away or into traffic. Additionally, this would help users who are concerned about not being able to distinguish the features of maps from the distortions due to visualization.

Two users wanted the colors to be added to the edges – proposing an outright compromise solution between the two visualizations, with distorted colored lines.

Instead of using colors, one participant suggested shading to be employed, akin to tourist maps. Another participant proposed using line thickness as a redundant measure to show the level of traffic.

Finally, three users requested the distortion to be purely due to traffic, effectively suggesting the distortion to be based on *tobserved* rather than the traffic slowdown *ta*. This approach was considered in the original design, but it has its problems, as demonstrated in Figure 13.

Despite their criticism in the first part of their response, one participant noted that they felt more confident navigating areas of high traffic when using the Distortion visualization.

**7 Discussion**

The main hypothesis for Accuracy expected there to be a significant difference, on average, between the two visualizations. Although a difference was observed between the Distortion and Color-Line visualizations – with the Color-Line visualization being on average more accurate – the main effect fell just short of significance (p = 0.06). This suggests that the user study could have benefitted from a higher number of research participants. Nevertheless, it is useful to examine the data further to try to determine the validity of the secondary claims of the Accuracy hypothesis. The first-order interactions between Visualization and Density/Traffic suggest that the Color-Line visualizations are more accurate in most maps – except for Dense maps and Low traffic level, the results for which are inconclusive.

The main hypothesis for Times expected there to be a significant difference, on average, between the two visualizations. A significant main effect was indeed observed between the Distortion and Color-Line visualizations – with the Color-Line visualization being on average faster. The first-order interactions between Visualization and Density/Traffic seem to favor the Color-Line visualization in Dense maps and Medium traffic, but do not offer conclusions for the other combinations.

Taken together, the Accuracy and Times observations seem to be in line with the specific claims of the two hypotheses. Dense maps were the most congested maps overall, and the color-banding effect of the Color-Line visualization – due to the segmentation of the traffic color spectrum into just four discrete levels – made it very difficult for users to gauge the level of traffic at a specific road segment as opposed to any of its neighbors. The case of Dense-Low maps – reproduced in Figure 14 – was especially difficult because most of the map ended up covered by the orange color. Unsurprisingly, the users’ top suggestion for improvement of the Color-Line visualization concerned this lack of information. Despite recording higher Accuracy than the Color-Line visualization, the Distortion visualization was slower to use. This indicates that while the Distortion visualization was able to communicate traffic information more granularly than the Color-Line visualization, the users did not know how to interpret an only slightly distorted grid of streets and ended up making spending a lot of time trying to weave through homogenous traffic. Although at least one user reported that they felt more confident when navigating areas of traffic, the question remains as to whether this confidence was misplaced. The first claim expecting the Distortion visualization to be more suitable than the Color-Line visualization when visualizing uniformly-congested traffic situations may yet be proven with additional data; although the results are promising, they are inconclusive.

The other specific claim of the Accuracy and Times hypotheses – which expected the Distortion visualization to be better in maps with highly localized traffic congestion – seems to only hold true for Time, and not for Accuracy. While the Medium-High and Sparse-High maps – reproduced in Figure 15 – required less time for users to fill in when using the Distortion visualization, the provided paths proved to be less accurate on average than when participants used the Color-Line visualization. This suggests that the Distortion visualization’s prominent bulging effect around high traffic made it easier for users to spot traffic congestion and give an answer; however, the routing choices the users made to avoid such areas of high traffic seem to be suboptimal compared to the choices made when using the Color-Line visualization. This may be due to the maximum traffic distortion radius *R* being set too high for these visualizations – making it seem like the congested areas extended farther than they should in the Distortion visualization. More work would be needed to evaluate whether this problem persists when different values of *R* are used, and when different traffic conditions are considered.

Meanwhile, the questionnaire results illustrate, to a large degree, the familiarity advantage enjoyed by the Color-Line visualization, compared to the Distortion visualization. All questions that asked the users to evaluate the performance of the two visualizations (accuracy, agreement, usefulness) for other people indicate that the users believe others to consider the Color-Line visualization to be the better visualization. This is true both for the indirect comparisons (e.g. comparing the two average scores given when answering “How useful do you think others found the \_\_\_ visualization?”), and the direct comparisons (e.g. the average answer to the question “Which visualization do you think others found more useful?”)

However, when users were asked to give their personal opinion, things became more complicated. Although indirect comparisons still turned out to be better for the Color-Line visualization, the difference between the answers was much narrower; meanwhile, the direct comparisons showed that users thought the Distortion visualization may be better than the Color-Line visualization, after all. The perceived accuracy comparison produced a perfect tie, and the perceived usefulness comparison assigned a better result to the Distortion visualization.

Overall, the results of this study are mixed. A significant main effect for Visualization was illustrated for the Time measure, while the Visualization measure fell just short of significance for Accuracy. Not only is the Color-Line visualization perceived as the safer option when it comes to making sure that other people understand a traffic map, the trends in the Accuracy and Time data suggest that Color-Line may be the best visualization of traffic congestion in the general case. However, the users’ own perception of the Distortion visualization indicates that people see it as potentially useful. Additionally, some support was found for the two specific claims about when Distortion visualization seems to be the better option – uniformly-congested and locally-congested traffic maps – but more data is required in each case to support both claims. Furthermore, in the case of uniform congestion, Distortion visualization’s better Accuracy scores seem to be offset by its decreased time performance as compared to the Color-Line visualization. Conversely, in the case of localized congestion, the better Time scores Distortion visualization records do not seem to have an effect on Accuracy.

**8 Future Work**

This study may be built upon in several ways. The conclusions based on the Accuracy measure would be strengthened by the inclusion of more participants. More levels of map density and more levels of traffic would help to inform under precisely what conditions the Distortion visualization is the most accurate, and under precisely what conditions it is the fastest visualization. Even without adding more diversity to map density and traffic levels, the effect of changes in the maximum distortion radius could be explored. The assumption that the *spherical* distortion function is the best function for spatial distortion could be tested, to lend more credibility to the results of the study. A clustering analysis could provide insight about the difference between those users who considered the Distortion visualization useful, and/or more accurate, and whether this perception was based on their actual results. Additionally, exploring the possible improvements offered by the study participants may prove fruitful – especially when it comes to the use of color alongside distortion to either make the 3D analogy more explicit, or more intuitively understandable.

**9 Summary and Impact**

The impact of this paper is twofold – it proposed a novel approach to visualization of traffic congestion that does not rely on color; an open library of visualization tools was created during the process which can be used in further research and development. Furthermore, the efficiency of the distortion visualization was tested in a user study alongside a widely-used alternative, the color-line visualization. By considering a variety of map densities and traffic levels, the study provided ground for judgments to be made about the situations in which the spatial-distortion approach is the best visualization choice. Although more research is needed to provide further support for these results, the value of this research lies, as well, in the fact that it quantifies the efficiency of the color-line visualization with a user study. New areas for research have been identified in order to assess the identified shortcomings the color-line visualization, as well as to improve the distortion visualization in those areas in which it was shown to be strong.

**References**

Arnott, Richard, and Kenneth Small. “The Economics of Traffic Congestion.” *American Scientist* 82.5 (1994): 446-55. Print.

Biersdorfer, J. D. “Q&A: Color-Coded Traffic in Google Maps.” *Gagetwise: The New York Times Blog*. The New York Times, 16Aug, 2010. Web. 15 Dec 2016.

Bruls, Mark, Kees Huizing, and Jarke J. van Wijk. "Squarified Treemaps." *Data Visualization 2000: Proceedings of the Joint EUROGRAPHICS and IEEE TCVG Symposium on Visualization.* By Willem Cornelis de Leeuw, and Robert van Liere. Vienna, Austria: Springer, 2000. 33-42. Print. *Eurographics.*

Card, Stuart K., and Jock Mackinlay. "The structure of the information visualization design space." *Information Visualization*, 1997. Proceedings, IEEE Symposium on. IEEE, 1997.

Cleveland, William S., and Robert McGill. "Graphical perception and graphical methods for analyzing scientific data." *Science* 229.4716 (1985): 828-833.

Fitch, Walter M., and Emanuel Margoliash. "Construction of Phylogenetic Trees." *Science* 155.3760 (1967): 279-84. Print.

Gastner, Michael T., and M. E. J. Newman. "Diffusion-based Method for Producing Density-equalizing Maps." *Proceedings of the National Academy of Sciences* 101.20 (2004): 7499-504. Print.

Geipel, Markus M. "Self-organization applied to dynamic network layout." *International Journal of Modern Physics C* 18.10 (2007): 1537-1549.

Guo, Hanqi, et al. "TripVista: Triple Perspective Visual Trajectory Analytics and its application on microscopic traffic data at a road intersection." *Visualization Symposium (PacificVis).* Hong Kong, China: IEEE Pacific, 2011. 163-70. Print.

Horvitz, Eric J. et al. "Prediction, Expectation, and Surprise: Methods, Designs, and Study of a Deployed Traffic Forecasting Service." *Proceedings of the Twenty-First Conference on Uncertainty in Artificial Intelligence (UAI2005).* Arlington, VA, USA: AUAI, 2005. 1-8. Print.

Johnson, B., and B. Shneiderman. " Tree-maps: a space-filling approach to the visualization of hierarchical information structures." *Visualization, 1991. Visualization '91, Proceedings., IEEE Conference on.* San Diego, CA, USA: IEEE, 1991. 284-91. Print.

Lu, Chang-Tien, Arnold P. Boedihardjo, and Jinping Zheng. "AITVS: Advanced Interactive Traffic Visualization System." *Proceedings of the 22nd International Conference on Data Engineering (ICDE’06).* Atlanta, GA, USA: IEEE, 2006. 167. Print.

Munzner, Tamara. "Process and Pitfalls in Writing Information Visualization Research Papers." *Information Visualization: Human-centered Issues and Perspectives.* By Andreas Kerren, John T. Stasko, Jean-Daniel Fekete, and Chris North. Vol. 4950. Berlin, Germany: Springer, 2008. 134-53. Print.

---. *Visualization Analysis and Design*. Boca Raton: CRC Press, 2014. Print.

North, Stephen C. "Drawing graphs with NEATO." *NEATO User Manual 11* (2004).

Oh, Seung Man, et al. “A Dissimilarity Measure for Comparing Origami Crease Patterns.” *Proceedings of the 4th International Conference on Pattern Recognition Applications and Methods (ICPRAM).* Lisbon, Portugal: SCITEPRESS, 2015. 386-93. Print.

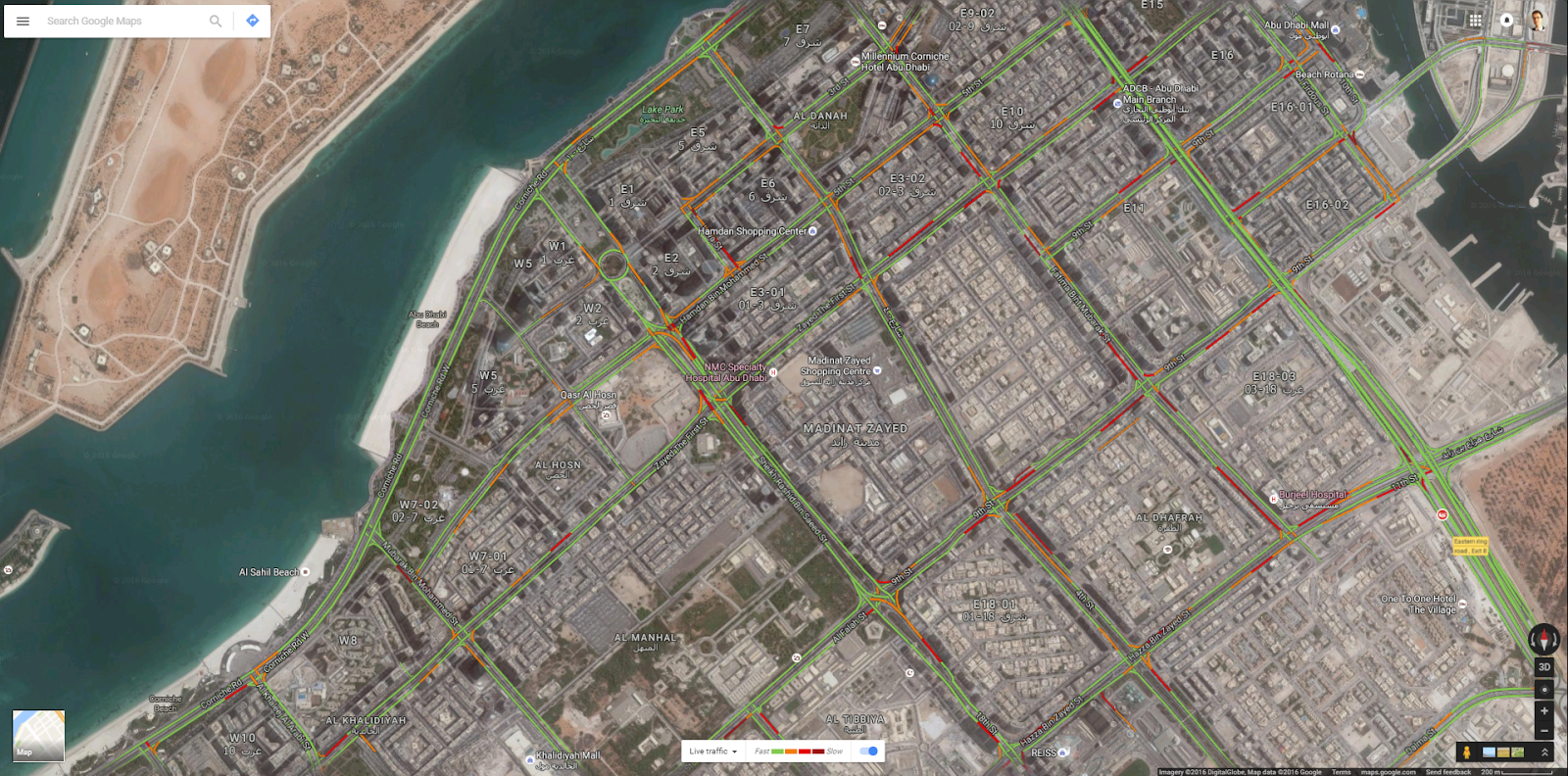
Schleiffer, Ralf. "Intelligent agents in traffic and transportation." *Transportation Research Part C: Emerging Technologies* 10.5 (2002): 325-329.

Skog, Tobias, Sara Ljungblad and Lars E. Holmquist. "Between aesthetics and utility: designing ambient information visualizations." *Proceedings of the IEEE Symposium on Information Visualization 2003 (INFOVIS’03).* Seattle, WA, USA: IEEE, 2003. 233-40. Print.

Wang, Zuchao, et al. "Visual Traffic Jam Analysis Based on Trajectory Data." *IEEE Transactions on Visualization and Computer Graphics* 19.12 (2013): 2159-68. Print.

|  |  |
| --- | --- |
| horwitz1.tif | horwitz3.tif |

**Figure 1.** Map visualization approach used by Horvitz et al. (4), including the additional features.



**Figure 2.** The traffic congestion overlay model of Google Maps Traffic.

cartograms.tif

**Figure 3.** Cartograms illustrating the results of the 2000 US presidential election, from Gastner and Newman (7502). A map of the continental United States (c) is distorted according to population distribution data at two different levels of granularity (d, e), and according to the distribution of electors in the US Electoral College (f).

|  |  |
| --- | --- |
| https://lh5.googleusercontent.com/v-EEHemrGYKdytG7dExGbTYJXYuwbkTVoFQJGP_IpuVZddTmf6wYXn1U5nmXkYn4zKIpnSVJ44MCEqzTU4_9ZNZrYJTN1_cS7z4052qStpMXgclQGhs-yy2HlwPHp-gLTUxYkUptz3c | https://lh5.googleusercontent.com/v-EEHemrGYKdytG7dExGbTYJXYuwbkTVoFQJGP_IpuVZddTmf6wYXn1U5nmXkYn4zKIpnSVJ44MCEqzTU4_9ZNZrYJTN1_cS7z4052qStpMXgclQGhs-yy2HlwPHp-gLTUxYkUptz3c |

**Figure 4.** An illustration of the problems with the Spring Network approach. Although both streams of the highway – pictured at the bottom edge of the network – are experiencing the same amount of traffic, the stream that is attached to the dense traffic network is dragged along and significantly shortened. This produces a false impression that traffic on the affected stream is much heavier than it should be.



**Figure 5.** A gravity visualization. A gridded rubber sheet is pushed down by a heavy object. Seen from above, the lines would seem to converge due to the object’s mass; seen from below, the lines would seem to bulge out around the object.

|  |  |  |  |
| --- | --- | --- | --- |
| **Distortion Function** | **Graph of Function**  **(Cross-Section View)** | **Visualization**  **(Top-Down View)** | **Problems** |
| ***Refraction*** | lensing distortion.tiff | google.colormap.png | The edges of the visualization are overlapping in one curve. |
| ***Gravity*** | gravity distortion.tiff | google.colormap.png | Overlapping at the center of the visualization. |
| ***Logistic*** | logistic distortion.tiff | google.colormap.png | Awkward shape. |
| ***Cosine*** | cosine distortion.tiff | google.colormap.png | Overlapping at the edges of the visualization. |
| ***Linear*** | linear distortion.tiff | google.colormap.png | Awkward shape. |
| ***Conical*** | linear distortion.tiff | google.colormap.png | Top-down view appears curved even though the cross-section is linear. |
| ***Convex*** | linear distortion.tiff | google.colormap.png | The distortion is sloping too gently. |
| ***Spherical*** | linear distortion.tiff | google.colormap.png |  |

**Figure 6.** A comparison of the different distortion functions considered for use in the spatial-distortion visualization.

|  |  |  |  |
| --- | --- | --- | --- |
| **Color-Line** | google.colormap.png | google.colormap.png | google.colormap.png |
| **Spatial-Distortion** | google.colormap.png | google.colormap.png | google.colormap.png |

**Figure 7.** Three additional traffic networks visualized with the spatial-distortion algorithm employing the *spherical* distortion function. Corresponding color-line visualizations are provided for comparison.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **High Traffic** | **Medium Traffic** | **Low Traffic** |
| **Dense Map**  **(Roads: 100 %)** | *V*: 0.12 s/car  *R*: 400 m | *V*: 0.18 s/car  *R*: 467 m | *V*: 0.24 s/car  *R*: 533 m |
| **Medium Map**  **(Roads: 80%)** | *V*: 0.36 s/car  *R*: 300 m | *V*: 0.42 s/car  *R*: 367 m | *V*: 0.48 s/car  *R*: 433 m |
| **Sparse Map**  **(Roads: 60%)** | *V*: 0.60 s/car  *R*: 200 m | *V*: 0.66 s/car  *R*: 267 m | *V*: 0.72 s/car  *R*: 333 m |

**Table 1.** The values of constants for the nine unique combinations of traffic networks investigated in the user study. *V* (amount of time between vehicle additions) increases by 0.06 s/car per step as one progresses from High to Low Traffic. At the same time, *R* (maximum distortion radius) increases by 66.6 m per step. As one progresses from Dense to Sparse Map, *V* increases by 0.24 s/car per step, while *R* decreases by 100 m per step.

|  |  |  |  |
| --- | --- | --- | --- |
| **Color-Line** | **High Traffic** | **Medium Traffic** | **Low Traffic** |
| **Dense Map** | google.colormap.png | google.colormap.png | google.colormap.png |
| **Medium Map** | google.colormap.png | google.colormap.png | google.colormap.png |
| **Sparse Map** | google.colormap.png | google.colormap.png | google.colormap.png |

|  |  |  |  |
| --- | --- | --- | --- |
| **Spatial-Distortion** | **High Traffic** | **Medium Traffic** | **Low Traffic** |
| **Dense Map** | google.colormap.png | google.colormap.png | google.colormap.png |
| **Medium Map** | google.colormap.png | google.colormap.png | google.colormap.png |
| **Sparse Map** | google.colormap.png | google.colormap.png | google.colormap.png |

**Figure 8.** The 18 unique color-line and spatial-distortion visualizations presented to user study participants.

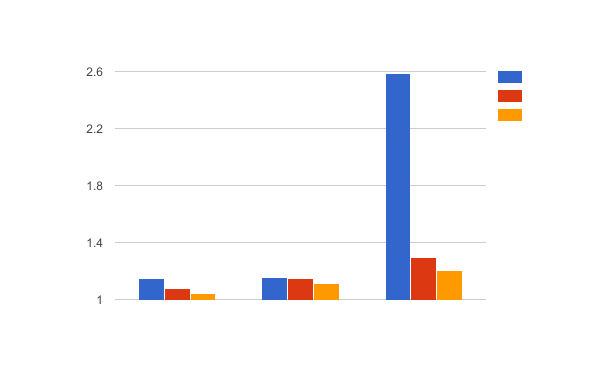
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **(AB)** | **HI** | **ME** | **LO** | **mean** |
| **DE** | 1.130772885 | 1.066477203 | 1.073911982 | 1.090387357 |
| **ME** | 1.147030453 | 1.116275559 | 1.114498874 | 1.125934962 |
| **SP** | 2.402581524 | 1.26653318 | 1.201898623 | 1.623671109 |
| **mean** | 1.560128287 | 1.149761981 | 1.130103159 | 1.279997809 |

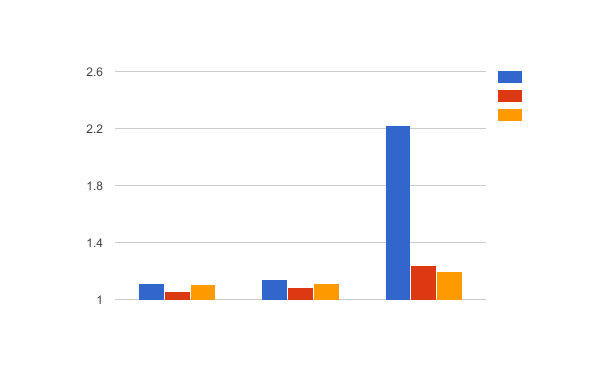
|  |  |  |  |
| --- | --- | --- | --- |
| **(AC)** | **DI** | **GO** | **mean** |
| **DE** | 1.088536376 | 1.092238337 | 1.090387357 |
| **ME** | 1.139380126 | 1.112489798 | 1.125934962 |
| **SP** | 1.694453586 | 1.552888632 | 1.623671109 |
| **mean** | 1.307456696 | 1.252538923 | 1.279997809 |

|  |  |  |  |
| --- | --- | --- | --- |
| **(BC)** | **DI** | **GO** | **mean** |
| **HI** | 1.629007468 | 1.491249106 | 1.560128287 |
| **ME** | 1.172919703 | 1.126604259 | 1.149761981 |
| **LO** | 1.120442916 | 1.139763403 | 1.130103159 |
| **mean** | 1.307456696 | 1.252538923 | 1.279997809 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Source of Variation** | **SS** | **df** | **MS** | **F** | **P-value** | **F Crit** |
| **A (Density)** | 34.1374 | 2 | 17.0687 | 137.3392 | **0** | 3.0119 |
| **B (Traffic)** | 22.6373 | 2 | 11.3187 | 91.0731 | **0** | 3.0119 |
| **C (Visualization)** | 0.4343 | 1 | 0.4343 | 3.4945 | 0.0621 | 3.8582 |
| **A \* B** | 35.9408 | 4 | 8.9852 | 72.2974 | **0** | 2.3879 |
| **A \* C** | 0.5630 | 2 | 0.2815 | 2.2651 | 0.1048 | 3.0119 |
| **B \* C** | 0.5975 | 2 | 0.2987 | 2.4038 | 0.0913 | 3.0119 |
| **A \* B \* C** | 0.7461 | 4 | 0.1865 | 1.5009 | 0.2005 | 2.3879 |
| **Error (Within)** | 69.3489 | 558 | 0.1243 |  |  |  |
| **Total** | 164.4054 | 575 |  |  |  |  |

**Table 2.** Results of the 3 x 3 x 2 ANOVA with replication for Accuracy.



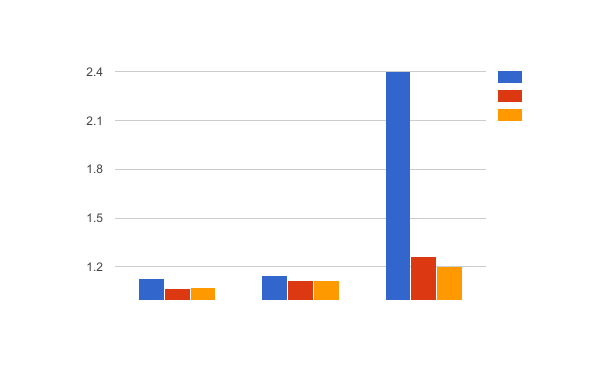


**Figure 9.** 3-factor ANOVA Accuracy graphs. The top chart presents Distortion Visualization values; the bottom chart presents Color-Line Visualization values. Colors distinguish Traffic: High/Medium/Low corresponds to Blue/Red/Yellow. Clusters of bars distinguish map Density: Dense/Medium/Sparse corresponds to 1st/2nd/3rd.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **(AB)** | **HI** | **ME** | **LO** | **mean** |
| **DE** | 1.130772885 | 1.066477203 | 1.073911982 | 1.090387357 |
| **ME** | 1.147030453 | 1.116275559 | 1.114498874 | 1.125934962 |
| **SP** | 2.402581524 | 1.26653318 | 1.201898623 | 1.623671109 |
| **mean** | 1.560128287 | 1.149761981 | 1.130103159 | 1.279997809 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Source of Variation** | **SS** | **df** | **MS** | **F** | **P-value** | **F Crit** |
| **A (Density)** | 17.0687 | 2 | 8.5343 | 97.8086 | 0 | 3.0281 |
| **B (Traffic)** | 11.31867 | 2 | 5.6593 | 64.8593 | 0 | 3.0281 |
| **A \* B** | 17.9704 | 4 | 4.4926 | 51.4879 | 0 | 2.4040 |
| **Error (Within)** | 24.3443 | 279 | 0.0873 |  |  |  |
| **Total** | 70.7021 | 287 |  |  |  |  |

**Table 3.** Results of the 3 x 3 ANOVA with replication for the Density \* Traffic interaction in Accuracy data.



**Figure 10.** A graph of the 2-factor ANOVA for Density \* Traffic interaction in Accuracy data. Colors distinguish Traffic: High/Medium/Low corresponds to Blue/Red/Yellow. Clusters of bars distinguish map Density: Dense/Medium/Sparse corresponds to 1st/2nd/3rd.

|  |  |  |  |
| --- | --- | --- | --- |
| **Density** | **Distortion** | **Color-Line** | **T-test** |
| **Dense** | 1.09 | 1.09 | 0.85 |
| **Medium** | 1.14 | 1.11 | **0.04** |
| **Sparse** | 1.69 | 1.55 | **0.05** |

|  |  |  |  |
| --- | --- | --- | --- |
| **Traffic** | **Distortion** | **Color-Line** | **T-test** |
| **High** | 1.63 | 1.49 | **0.05** |
| **Medium** | 1.17 | 1.13 | **0.00** |
| **Low** | 1.12 | 1.14 | 0.24 |

**Table 4.** Average Accuracy values for Density \* Visualization and Traffic \* Visualization interactions.

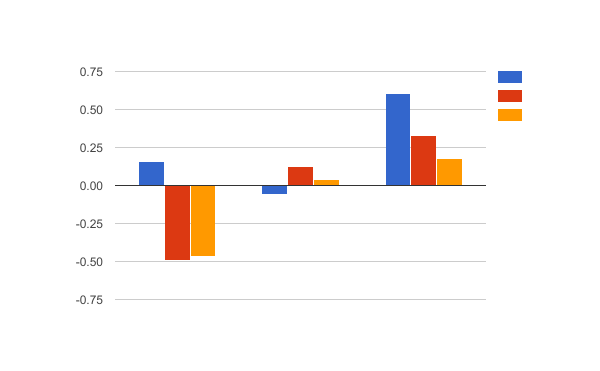
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **(AB)** | **HI** | **ME** | **LO** | **mean** |
| **DE** | -0.02 | -0.54 | -0.50 | -0.35 |
| **ME** | 0.02 | -0.02 | 0.01 | 0.01 |
| **SP** | 0.65 | 0.26 | 0.13 | 0.35 |
| **mean** | 0.22 | -0.10 | -0.12 | 0.00 |

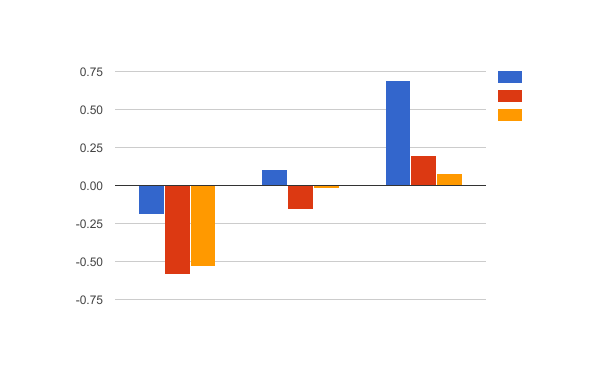
|  |  |  |  |
| --- | --- | --- | --- |
| **(AC)** | **DI** | **GO** | **mean** |
| **DE** | -0.27 | -0.44 | -0.35 |
| **ME** | 0.04 | -0.02 | 0.01 |
| **SP** | 0.37 | 0.32 | 0.35 |
| **mean** | 0.05 | -0.05 | 0.00 |

|  |  |  |  |
| --- | --- | --- | --- |
| **(BC)** | **DI** | **GO** | **mean** |
| **HI** | 0.24 | 0.20 | 0.22 |
| **ME** | -0.01 | -0.18 | -0.10 |
| **LO** | -0.08 | -0.16 | -0.12 |
| **mean** | 0.05 | -0.05 | 0.00 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Source of Variation** | **SS** | **df** | **MS** | **F** | **P-value** | **F Crit** |
| **A (Density)** | 46.8368 | 2 | 23.4184 | 104.1147 | **0** | 3.0119 |
| **B (Traffic)** | 13.9081 | 2 | 6.9540 | 30.9167 | **0** | 3.0119 |
| **C (Visualization)** | 1.2127 | 1 | 1.2127 | 5.3914 | **0.0206** | 3.8582 |
| **A \* B** | 6.5100 | 4 | 1.6275 | 7.2356 | **0** | 2.3879 |
| **A \* C** | 0.4253 | 2 | 0.2126 | 0.9453 | 0.3892 | 3.0119 |
| **B \* C** | 0.4828 | 2 | 0.2414 | 1.0732 | 0.3426 | 3.0119 |
| **A \* B \* C** | 2.2828 | 4 | 0.5707 | 2.5373 | **0.0392** | 2.3879 |
| **Error (Within)** | 125.5102 | 558 | 0.2249 |  |  |  |
| **Total** | 197.1686 | 575 |  |  |  |  |

**Table 5.** Results of the 3 x 3 x 2 ANOVA with replication for normalized Times.



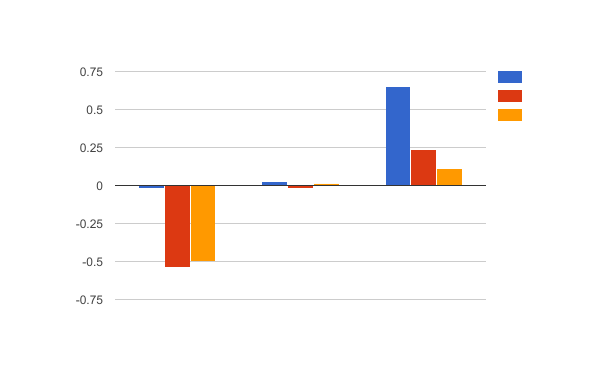


**Figure 11.** 3-factor ANOVA Times graphs. The top chart presents Distortion Visualization values; the bottom chart presents Color-Line Visualization values. Colors distinguish Traffic: High/Medium/Low corresponds to Blue/Red/Yellow. Clusters of bars distinguish map Density: Dense/Medium/Sparse corresponds to 1st/2nd/3rd.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **(AB)** | **HI** | **ME** | **LO** | **mean** |
| **DE** | -0.02 | -0.54 | -0.50 | -0.35 |
| **ME** | 0.02 | -0.02 | 0.01 | 0.01 |
| **SP** | 0.65 | 0.26 | 0.13 | 0.35 |
| **mean** | 0.22 | -0.10 | -0.12 | 0.00 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Source of Variation** | **SS** | **df** | **MS** | **F** | **P-value** | **F Crit** |
| **A (Density)** | 22.5402 | 2 | 11.2701 | 93.3743 | **0** | 3.0281 |
| **B (Traffic)** | 7.2268 | 2 | 3.6134 | 29.9377 | **0** | 3.0281 |
| **A \* B** | 3.3357 | 4 | 0.8339 | 6.9091 | **0.0000** | 2.4040 |
| **Error (Within)** | 33.6747 | 279 | 0.1207 |  |  |  |
| **Total** | 70.7021 | 287 |  |  |  |  |

**Table 6.** Results of the 3 x 3 ANOVA with replication for the Density \* Traffic interaction in Times data.



**Figure 12.** A graph of the 2-factor ANOVA for Density \* Traffic interaction in Times data. Colors distinguish Traffic: High/Medium/Low corresponds to Blue/Red/Yellow. Clusters of bars distinguish map Density: Dense/Medium/Sparse corresponds to 1st/2nd/3rd.

|  |  |  |  |
| --- | --- | --- | --- |
| **Density** | **Distortion** | **Color-Line** | **T-test** |
| **Dense** | -0.27 | -0.44 | **0.05** |
| **Medium** | 0.04 | -0.02 | 0.45 |
| **Sparse** | 0.37 | 0.32 | 0.54 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Traffic** | **Distortion** | **Color-Line** | **T-test** |
| **High** | 0.24 | 0.20 | 0.68 |
| **Medium** | -0.01 | -0.18 | **0.01** |
| **Low** | -0.08 | -0.16 | 0.34 |

**Table 7.** Average Times values for Density \* Visualization and Traffic \* Visualization interactions.

|  |  |
| --- | --- |
| **Question:** | **Result:** |
| In general, how close do you think your paths in the color-line visualization were to the ideal path? | 3.25 |
| In general, how close do you think your paths in the distortion visualization were to the ideal path? | 3.16 |
| In general, which visualization do you think is better when it comes to helping you get close to the ideal path? | 3.00 |
| In general, how close do you think other people's paths in the color-line visualization were to the ideal path? | 3.44 |
| In general, how close do you think other people's paths in the distortion visualization were to the ideal path? | 3.06 |
| In general, which visualization do you think is better when it comes to helping other people get close to the ideal path? | 2.81 |
| In general, how close do you think the paths you provided in the color-line visualizations were to paths provided by other people taking this survey? | 3.25 |
| In general, how close do you think the paths you provided in the distortion visualizations were to paths provided by other people taking this survey? | 3.06 |
| In general, which visualization do you think is better when it comes to helping people identify paths that are close to the paths identified by other people taking this survey? | 2.97 |
| In general, how useful do you think the color-line visualization was? | 3.53 |
| In general, how useful do you think the distortion visualization was? | 3.50 |
| In general, which visualization do you was more useful? | 3.03 |
| In general, how useful do you think other people found the color-line visualization to be? | 3.66 |
| In general, how useful do you think other people found the distortion visualization to be? | 3.31 |
| In general, which visualization do you think other people found to be more useful? | 2.75 |

**Table 8.** The questionnaire questions and results. Where users were asked for opinion on an individual visualization, the scale ranged from 1 (“not at all”) to 5 (“extremely”). Where users were asked to compare the two visualizations, the scale ranged from 1 (color-line visualization) to 5 (distortion visualization).



**Figure 13.** The traffic network of Abu Dhabi Island when distortion is based on *tobserved* instead of *ta*. The roads that experience heavy congestion are clustered in the center; the comparatively free roads inside blocks overflow over the main roads that used to encircle them. The result is an indecipherable tangle of roads.

|  |  |
| --- | --- |
| oogle Maps Traffic.png | oogle Maps Traffic.png |

**Figure 14.** The Dense-Low traffic map in both visualizations.

|  |  |
| --- | --- |
| oogle Maps Traffic.png | oogle Maps Traffic.png |
| oogle Maps Traffic.png | oogle Maps Traffic.png |

**Figure 15.** The Medium-High traffic map in both visualizations (top row), followed by the Sparse-High traffic map in both visualizations (bottom row).

**Appendix A: Alternative Traffic Node Distortion Functions**

In addition to the *spherical* spatial-distortion function, seven other functions were considered.

**A.1 Refraction Function**

The *refraction* function was the first to be investigated:



This function mimics the effect of using a magnifying glass of a constant radius centered at the distorting node; while this is a readily understandable visual metaphor, it is problematic because it requires many nodes in the distortion radius to be displaced to the same distance from the distorting node (just like a magnifying glass does). When applied to traffic visualization, this requirement produces tangles of low-congestion streets that are impossible to understand visually, as illustrated in Figure A1. Thus, the refraction algorithm was discarded.

**A.2 Gravity Function**

A distortion function that emulates the effects of negative *gravity* followed:



where



The gravity function is remarkable among the other functions in that it is continuous; it does not need require a maximum radius *R* to be specified; instead it can be fine-tuned by adjusting the value of the gravity constant *G*. Nevertheless, the function is not ideal, as it causes the visualized surface around the congested node to assume a hyperboloidal, volcano-like shape if G is too high. Mathematically, the *b* nodes closest to *a* are assigned *b*’ coordinates that are farther out than *b* nodes that start farther away from *a*. When visualized, this looks like nodes that are more central skip over nodes the farther nodes; with edges (streets) then drawn between their respective nodes, this overlap produces crossings that should not be part of the visualization. Of course, the overlap can be eliminated by drastically reducing the value of the G constant; then, however, most nodes do not experience a perceptible displacement. This conundrum stems from the fact that the derivative of gravity function decreases extremely rapidly.

**A.3 Logistic Function**

Thus, another function was investigated – the *logistic* function. Unfortunately, it produces the shape of a cylinder with smooth edges:



The steepness coefficient of 0.0159 ensures no overlaps at the function’s inflection point. Similar to the gravity function, the value of *r* is not used to determine the maximum extent of the distortion effect. Instead, *r*/2 is the inflection point’s distance from node *a*, and the a small amount of distortion extends beyond *r*.

**A.4 Cosine Function**

The *cosine* function does produce a nearly perfect hemispherical distortion; however there is a small amount of overlap at the edges.



**A.5 Linear Function**

The overlap is eliminated in the *linear* function, but the shape of the distortion function is still not a perfect hemisphere.



**A.6 Conical Function**

The *conical* distortion has a straight-line cross-section, which produces a cone-like surface that protrudes up from the plane with a vertex above the distorting node:



**A.7 Convex Function**

Compared to the *spherical* function, *convex* has the opposite profile, with curving sides that meet at a sharp vertex above the distorting node:





**Figure A1.** The tangling effect produced by a spatial-distortion visualization using the *refraction* distortion function.

1. Arnott and Small, 446 [↑](#footnote-ref--1)
2. ibid. [↑](#footnote-ref-0)
3. ibid. [↑](#footnote-ref-1)
4. Schleiffer, 327 [↑](#footnote-ref-2)
5. e.g. Guo et al.; Lu, Boedihardjo and Zheng; Wang et al. [↑](#footnote-ref-3)
6. e.g. Bruls, Huizing and van Wijk; Fitch and Margoliash; Johnson and Schneiderman; Oh et al.; Skog, Ljungblad and Holmquist [↑](#footnote-ref-4)
7. Munzner, *Process*, 4 [↑](#footnote-ref-5)
8. Arnott and Small 446 [↑](#footnote-ref-6)
9. Munzner, *Visualization*, 1 [↑](#footnote-ref-7)
10. Horvitz et al., 4 [↑](#footnote-ref-8)
11. Biersdorfer [↑](#footnote-ref-9)
12. Gastner and Newman, 7502 [↑](#footnote-ref-10)
13. Cleveland and McGill, 830 [↑](#footnote-ref-11)
14. Card and Mackinlay, 5 [↑](#footnote-ref-12)
15. Schleiffer, 327 [↑](#footnote-ref-13)
16. Geipel, 1538 [↑](#footnote-ref-14)
17. North [↑](#footnote-ref-15)
18. Biersdorfer [↑](#footnote-ref-16)