Int J Software Informatics, Volume 9, Issue 3 (2015), pp. 355–370
International Journal of Software and Informatics, ISSN 1673-7288

E-mail: ijsi@iscas.ac.cn
http://www.ijsi.org
Tel: +86-10-62661040

An Evaluation of the Utility of Large High-Resolution Displays for Comparative Scientific Visualisation

C. Müller, M. Krone, K. Scharnowski, G. Reina and T. Ertl

(Visualisierungsinstitut der Universität Stuttgart, Germany)

Abstract In many disciplines, such as computer aided drug design, multiple simulation runs are performed with varying parameters, yielding ensembles of data sets. Comparative visualisation of these simulation results can help understanding the influence different parameters have. However, researchers might need to compare large numbers of variants. Single desktop monitors often do not have the resolution and screen size required for showing a whole ensemble at once with sufficient detail. Wall-sized high-resolution displays can be a solution for this problem. Although a number of studies has been conducted on how large high-resolution displays affect the speed and accuracy of certain tasks, only few of them are related to actual scientific visualisation tasks. We built a system for comparative visualisation of simulation results that can be used with conventional desktop monitors and with large high-resolution displays. We conducted a study using biochemical simulation data to evaluate the impact of screen size and 3D stereo output on a comparison task.
Key words: large high-resolution displays; usability; structural biology; comparative visualisation; user study

Müller C, Krone M, Scharnowski K, Reina G, Ertl T. An evaluation of the utility of LHRDs for comparative scientific visualisation. *Int J Software Informatics*, **Vol.9, No.3 (2015): 355–370. http://www.ijsi.org/1673-7288/9/i225.htm**

1 Introduction

A common scenario in computational drug design is the analysis of similar molecules to assess their differences in function or structure. The function is often investigated by conducting simulations under different conditions or with selective mutations of amino acids, resulting in so-called ensembles. Grouping of the simulation results is usually based on the principle of molecular similarity: molecules with similar geometry and biochemical properties are likely to exhibit similar functions[18]. Comparisons are based on a variant about which some knowledge is available, and the task is to find the best possible alternatives in terms of similar properties. Consequently, the visualisation usually depicts the differences between a variant and a selected baseline.

Since small multiples are recognised as a sensible choice for multi-way comparison[42,43], we consider it a reasonable approach to allocate the whole screen

Corresponding author: Christoph Müller, Email: christoph.mueller@visus.uni-stuttgart.de
Received 2015-04-19; Revised 2015-07-05; Accepted 2015-11-24.

space to a grid of identically-sized renderings of pairwise comparisons. The details that can be gleaned from each variant are directly related to the available resolution of the display. Thus, one major challenge of comparative visualisation for large data is scalability. Regular desktop monitors usually lack the resolution for large ensembles. Large high-resolution displays might be an alternative. However, while allowing for improved insight, they require specialised interaction concepts, since mouse and keyboard are impractical when a user should be able to move freely in front of the screen.

We performed an experiment to investigate how the display size and resolution impact participants finding very similar and very different members of an ensemble of molecules. In the experiment, we used a layman task resembling the one of the application domain scientists from the field of biochemistry. Other than broadening the audience significantly, the task offers the additional benefit that we know the ground truth, which allows for a quantitative evaluation. These conditions do not pertain to the fully-fledged application case. We thus only present the results of an informal evaluation with our collaboration partners that provided the data since the analysis cannot be performed by non-experts. This paper is an extension of our previous work[24] presented at the Symposium on Visual Information Communication and Interaction (VINCI).

2 Related Work

Comparative visualization for ensemble data facilitates finding similarities or dissimilarities between data sets. A number of approaches for vector field ensembles (e.g. from CFD simulations) exists[11,14,44], some of which also incorporate uncertainty[30]. In structural biology and computational drug design, similarity-based methods using geometrical properties of the respective molecules are used to examine chemical databases[8]. The widely-used molecular visualisation software Chimera[31] offers a structure-based morphing animation between two input molecules by interpolating the atom positions, which facilitates a comparison of two input data sets. There are a number of comparison algorithms that use perform either shape-based comparisons[13] or focus on physical properties like the electrostatic potential[1]. Scharnowski et al.[38] combined these approaches.

Large high-resolution displays have been built for more than two decades. Early systems, including the first CAVE[6], were exclusively built using video projectors[12]. In contrast, most recent installations use flat panel displays for achieving up to 1.5 billion pixels[32] at a comparatively low price[25,29]. Wall-sized displays are often used to solve tasks in a collaborative manner. Furthermore, they usually provide at least partial immersion by covering the whole field of view[33]. Applications for such large high-resolution displays range from control rooms over geo-spatial visualisations to scientific ones. Ni et al.[28] provide a comprehensive overview of system designs and applications.

The general benefits and drawbacks of large high-resolution displays have been previously studied using navigation, comparison, and search tasks mostly using satellite imagery[5,37] or maps[16,36]. Ball et al.[5] reported improved performance of users for finding and comparison tasks when using large high-resolution displays. They also found benefits from physical navigation in front of the display compared

to panning virtually. Shupp et al.[37] obtained similar results for searching or navigational tasks. In addition, they found that curved displays further decreased the time to complete a task thanks to the reduced cost for physical navigation.

However, some studies suggest that user performance does not generally improve with resolution or wider views: Jakobsen and Hornbæk[16] explained this discrepancy with the fact that certain methods, like focus and context techniques, do not work well with large high-resolution displays. They reason that interaction techniques like zooming and panning increase the mental effort as these affect the whole content of the large display. They conclude that, in order to be useful on large high-resolution displays, visualisation techniques need to be carefully chosen and adapted. Similarly, Yost and North[47] showed that displays beyond human visual acuity make sense only if the employed visualisation method is designed appropriately. Another factor when using physically large displays is that changing positions of the viewer can make it difficult to grasp certain mappings of data correctly. Bezerianos and Isenberg[3] compared how users were able to judge angles, areas, and length from different positions in front of a wall-sized display. Jakobsen and Hornbæk[17] investigated how scaling the information space with the available display space adds to the problem of designing the right visualisation. When testing a variable information space, they found that the increasing screen real estate does not outweigh the increasing complexity of the task. For a fixed information space, like we use, tasks could be completed faster with larger displays, but the effect was clearly not linear, i.e. adding more resolution to an already large display space had a rather small benefit. Ruddle et al.[36] emphasised that the advantages of physical navigation compared to virtual navigation only emerge if the entire data set is visible on the large high-resolution display and there is, thus, no need for panning and zooming, as is the case with our approach.

Although Moreland[25] states that research on large high-resolution displays should focus on "how to use displays" and that this research should be application-driven, only few specific application areas have been tested. It has been shown that comparative genomics analysis can make effective use of the high amount of pixels available[34]. Gjerlufsen et al.[9] identified a biological comparison task as application case for large high-resolution displays, but did not conduct a user study. A study comparing the display of long documents on a large screen with a desktop monitor and printed paper yielded mixed results[46]; the large high-resolution display only showed significant benefits for searching differences. An informal long-term study[4] has been conducted to investigate benefits and drawbacks when performing other everyday tasks on a 3×3 array of LCDs. Perceived advantages included more space for quicker access to running applications and an increased awareness for secondary tasks. Endert et al.[7] present further design considerations for employing large high-resolution displays in day-to-day office use. They conclude that the physical form factor and the placement of interaction devices plays an important role to make this a suitable scenario.

The need for interaction techniques beyond keyboard and mouse for wall-sized displays is widely known[22]. While we emulate a pointing device using ray-casting[2,23,45], other techniques using enhanced tablet devices[15] and using the whole body[19] have also been investigated. Although it has been shown that 3D

navigation tasks can generally benefit from physically large screens[41], the performance of large high-resolution displays for scientific visualisation applications has not yet been investigated sufficiently. For example, Nam et al.[27] mainly discuss technical issues like frame rates and delays, while Scheidegger et al.[39] already hint at a small multiples scenario, but do not offer any details or evaluation. Most importantly, none of the aforementioned studies specifically addressed three-dimensional or even stereo visualisation, which is the focus of our experiment.

3 Application Scenario

3.1 Comparative visualisation

We used a modified version of the MegaMol system[10] to implement a tile-based comparison of the following techniques. These two comparative visualisation methods address analysis tasks from the domain of structural biology:

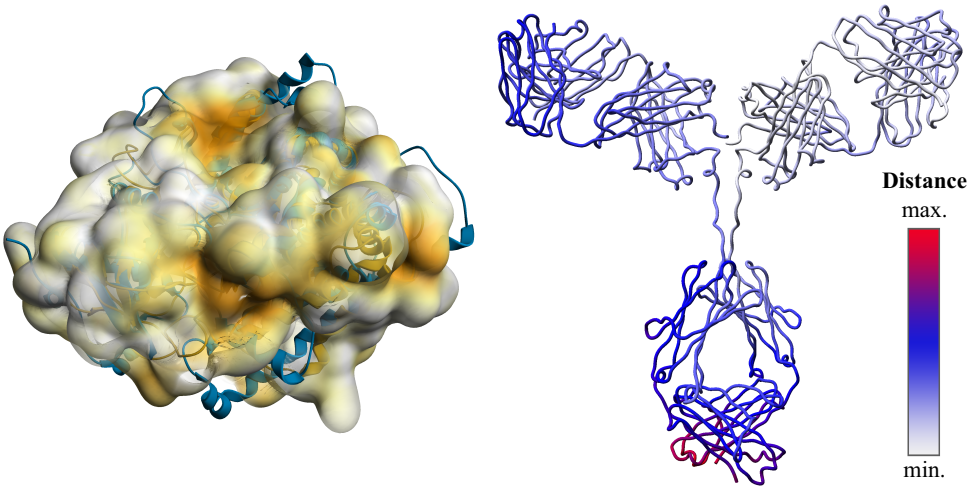

Figure 1. Examples of the two visualizations used in our comparison. Left: The pairwise surface-based comparative visualisation[38]. Orange surface parts indicate a high difference in the electrostatic potential, whereas the opacity of the surface depicts the geometric difference. Right: The structure-based comparative visualisation applied to two conformations of an antibody (data set D1). The colour gradient to the right was used to illustrate the structural deviation.

Surface-based comparison

A biomolecule's function is heavily influenced by both shape and physico-chemical properties of its surface. For example, many enzymatic reactions are triggered by smaller molecules docking to the surface. Not only the geometric surface shape has to fit, but also the electrostatic potential. Hence, one of the approaches used in our setup is the surface-based comparative visualisation presented by Scharnowski et al.[38]. Here, a point-to-point-mapping between two given input surfaces is defined by using a deformable model approach. This mapping is then used to compute local difference measurements for both the local surface geometry and the electrostatic

potential at the surface. Figure 1 shows a resulting rendering. The potential difference is encoded as colour saturation of the surface. Geometrical differences are mapped to transparency to provide a measurement for the uncertainty of the comparison as comparisons between geometrically very different surface parts are less meaningful. The rendering consequently focuses on geometrically similar regions and highlights regions of high potential differences, and hence allows to draw conclusions about functional regions of the surface.

Structure-based comparison

Not only the surface structure, but also the underlying internal structure of a protein is essential for its function. Proteins are linear chains of amino acids. These chains fold into a specific, three-dimensional structure. This so-called tertiary structure is often visualised using a spline that closely follows the amino acid chain[35]. The spline is usually decorated with 3D geometry to enhance perceptibility, e.g. using a tube that surrounds the spine. As proteins are flexible, there can be differences in the structures. These differences are of interest since misfolded proteins can cause illnesses like Alzheimer's disease. Hence, the second application we use in our setup is a comparison of the tertiary structure. Here, two proteins with equal chain length are superimposed as good as possible using the established RMSD alignment algorithm[20]. The Euclidean distance of two splines that represent the proteins is computed. These distances are then colour-coded onto the tertiary structure as shown in Figure 1. The geometry shader implementation presented by Krone et al.[21] is used for fast rendering. This visualisation allows biologists to analyse structural differences and to assess their impact on the protein function. It can also be used to analyse intrinsically disordered proteins, which have no fixed three-dimensional functional structure as long as they are not bound to another molecule.

3.2 Large high-resolution display

Our powerwall is a tiled large high-resolution display made of five portrait-oriented strips[26]. Two 4K LCoS projectors are projecting the images for the left and the right eye on each strip. The image of each projector is deflected by a mirror in order to achieve the portrait orientation. The setup yields a total net resolution of $10{,}800 \times 4{,}096$ pixels for each eye, which are projected onto a physical screen size of about 6×2.2 metres. A pixel is about 0.5 millimetre in size, which corresponds to about 50 ppi. Stereo separation is done using interference filters (Infitec). The lack of colour fidelity is the main disadvantage of this technique as it filters the colour spectrum to achieve the channel separation. As our comparative visualisation makes use of colour coding, we expect this to be disadvantageous, although we chose the colour table such that is also works in stereo mode. For displaying mono images, the stereo filters in the projectors can be completely removed, which leads to increased colour fidelity in this case.

A cluster of ten nodes is used to create the imagery for the ten projectors. Each node is equipped with two Intel Xeon X5650 CPUs and 24 GB of RAM and comprises two NVIDIA Quadro 6000 GPUs. The cluster nodes are running Windows HPC Server 2008 R2. We use MPI to make use of the DDR InfiniBand interconnect between

the nodes. The machine that controls the whole application and processes the user input is connected to the cluster via Gigabit ethernet. In our experiment, the users did not interact directly with the this computer. Instead they used a 6DOF mid-air pointing device (wand, see Figure 2), which was tracked using an optical tracking system from NaturalPoint. The wand has two buttons, one on the top and one on the button, which can be used like the buttons of a mouse.

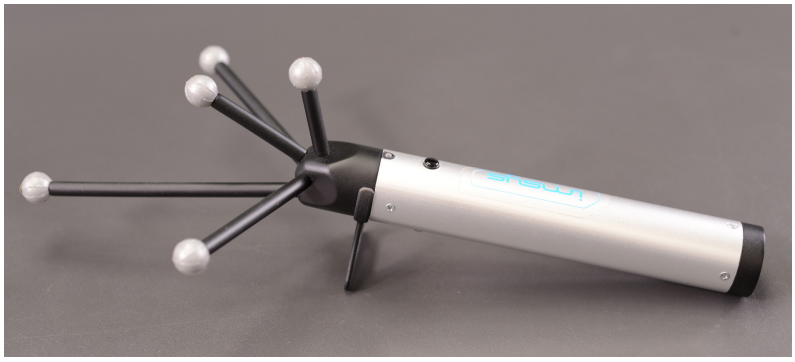

Figure 2. The wand used in our tests. The device has two buttons that trigger the interaction modes for rotation, translation, and dragging views.

4 User Study

4.1 Test scenarios

The goal of our user study was to test whether an large high-resolution display is applicable for comparative scientific visualisation. Although combining the comparison of different display scenarios and interaction techniques into one experiment makes it difficult to isolate the effects of either factor, we opted for a setup that is as close to a real application scenario as possible. We tested three data sets under the following display conditions.

Large high-resolution display

Two of the three display conditions that we tested have been performed on the large high-resolution display described Section 3.2. One was performed in stereo mode (Infitec filters active), while the other was done without colour filters. In both cases, the user was presented a grid of 10×4 visualisations using the same baseline molecule (Figure 3). Although the grid is freely configurable, the views could not be resized by the user. However, their order in the grid could be changed using the wand. For doing so, the user had to point the wand at one of the views, which indicated the selection by a brighter border line. By pressing both buttons of the wand at the same time, this view then could be dragged to a different location. Pressing the buttons separately allowed for rotating (top button) and zooming (bottom button) according to the indirect HOMER technique[2]. All transformations were synchronised over all views. While performing the task, the participants were free to move anywhere inside the tracking volume, which covered the whole area in front of the screen up to four metres distance from the wall.

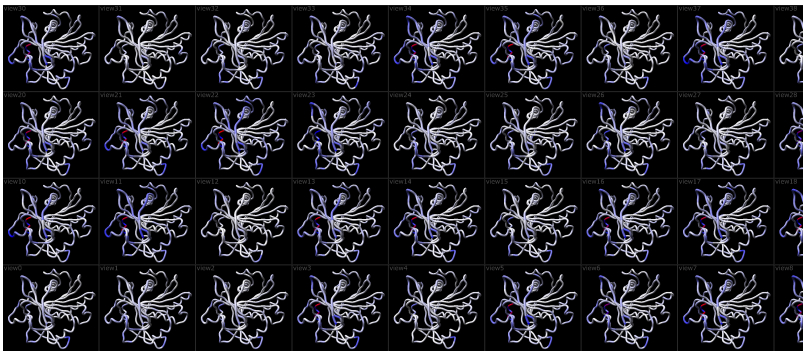

Figure 3. Layout of the application on the tiled display: all 40 variants of data set D2 are visible. Each of the views covers approximately 1000×1000 px.

Desktop monitor

In the single monitor case using a 24 inch WUXGA display, we split the screen into two views (Figure 4), similarly to what van den Elzen and van Wijk[43] used for visual data exploration via information visualisation techniques. Our first view showed the same grid of 40 visualisations that is used on the large high-resolution display, only at a much smaller scale. The second one allowed the user to compare two of these visualisations side-by-side to overcome the limited resolution available for each member of the ensemble view. As in the large high-resolution display case, the user had the option to rotate the data set and to zoom into the views, in this case using two different mouse buttons. Again, all views were synchronised. The desktop application provided two additional modes of mouse interaction: one that allowed reordering the small multiples by exchanging two variants with mouse clicks and another for selecting the content of the two detail views.

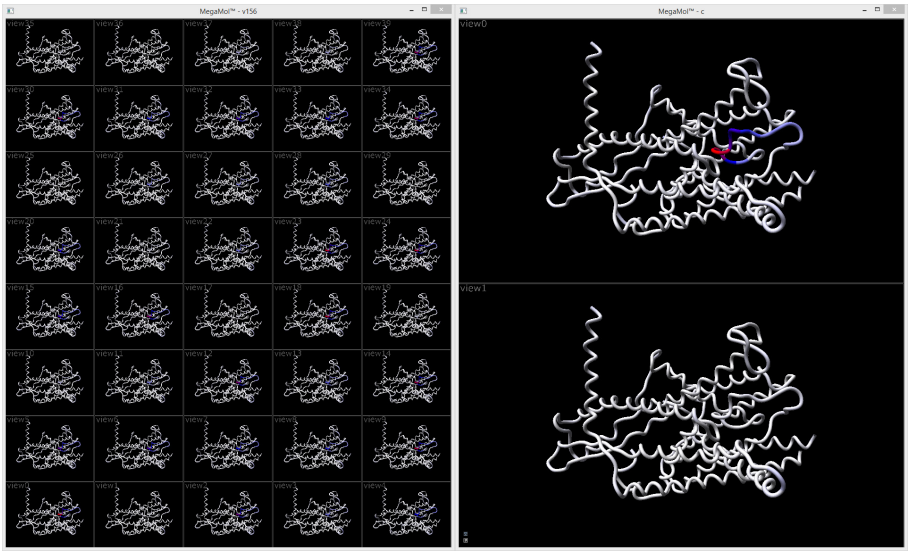

Figure 4. Layout of the application on the desktop monitor: the left view shows an overview of all variants while the right view enables side-by-side comparison of two user-selected variants (data set D0).

We used a 24 inch WUXGA (1,920 × 1,200 px) display from Dell for this test, which has approximately 96 ppi. The display is powered by an NVIDIA Quadro K6000 GPU in a dual Xeon E5-2637 node equipped with 128 GB RAM.

4.2 Experiment

Task

We used the simplified structure-based comparison for two reasons: First, a user who is no expert in structural biology must be able to complete the task. Second, the task should not be exploratory, but rather yield a quantitative result so we can compute its distance to the ground truth.

The structure-based comparison can be used without prior knowledge in structural biology, since it only requires assessing distances which are colour-coded on the geometry. Selecting the colour coding proved to be difficult due to the reduced colour spectrum available when using Infitec stereo separation, in particular for red tones. However, in order to exploit the wider view of a large high-resolution display, important parts should be salient, which suggests the use of red. We finally decided on the compromise depicted in Figure 1, which is not optimal, but still retains the saliency of red for emphasising important parts.

For each test, we showed 40 views each displaying a comparison with a different point in time of the simulation. The geometry of the baseline variant was used in all views, i.e. participants did not have to compare different structures. They had to identify the view that showed the protein with the smallest deviation from the baseline and the one with the largest deviation. As there can be multiple similar variants, we told the participants that they did not have to find the absolute minimum and maximum, but a view that was as close as possible. For the evaluation, we computed how far the chosen solution was from the correct one.

The three data sets we used were a transmembrane protein (D0, Figure 4), an antibody (D1, Figure 1), and a globular protein (D2, Figure 3). D0 is difficult, because it has a relatively high number of candidates for both minimum and maximum deviation. The antibody (D1) is geometrically most complex, but has three quite obvious candidates for the minimum deviation. D2 has many good candidates for the smallest deviation having almost equally good scores. Finding the maximum deviation is difficult, because medium deviations, which are hard to assess, occur in several regions.

Participants

The study was performed by 18 volunteers including one of the application domain experts (two females and 16 males) aged between 24 and 41 (avg. 31, $\sigma = 4.06$). We recruited participants that rated themselves either expert on the field of 3D visualisation (avg. 4 on a five-level Likert scale, $\sigma = 1.25$) or on the field of structural biology (avg. 1.5, $\sigma = 0.83$). The average experience with 3D stereo output devices was 3.28 ($\sigma = 1.28$), with 3D input devices like the wand used in our setup it was only 2.61 ($\sigma = 1.01$).

Procedure

Each participant first performed the experiment with an easy test data set in each of the abovementioned scenarios to familiarise themselves with the task and the respective interaction methods. Then, all data sets were tested in all scenarios. The order of the scenarios and of the data sets was different for each participant to account for learning effects. Also, the variants displayed in each view have been assigned randomly for every single test.

The participants could freely choose how they performed the task, i.e. no specific strategy was prescribed. They could also decide for each test in which order they searched for the view with the minimal and maximal deviation. The time was stopped once both views had been named. After the user completed a single task, we asked them to rate on an unlabelled five-level Likert scale how confident they are that they found the correct answers and how easy it was finding them. The participants had no time limit for completing the task, but were told that they should stop once they felt that they could not find any better candidate. In total the whole procedure took a bit more than an hour on average, but a few participants required almost two hours.

4.3 Hypotheses

Based on the outcome of previous studies, it was difficult to predict which setup would perform best. On the one hand, it has been shown that physical navigation in front of large displays has advantages over virtual navigation on small displays. On the other hand, users could try to solve the task without any or with very little (virtual) navigation, have more experience using the mouse, and the mouse is also more accurate than the wand. Nevertheless, we hypothesised that the large display has slight advantages.

However, as we required colour coding, we believed that the stereo setup would be less advantageous for three reasons: the stereo technology used for our tiled display relies on Infitec, which has less colour fidelity than other technologies like polarisation or active shutter glasses. The glasses themselves also limit the field of view. Furthermore, due to the simplification of the task for average users, no comparison of the spatial structure is necessary for solving the task. Thus, we did not expect significant advantages from the improved depth cues when using stereo; in contrast, the negative effects should predominate, particularly as the need for interaction remains: occlusion, which is inherent to 3D visualisation, is not neutralised by stereo projection. The interesting question was whether the colour distortion would stay in an acceptable range for the actual application.

4.4 Results

Accuracy and speed

The root-mean-square deviation (RMSD) value[20] is commonly used in structural biology as a ground truth for the structural difference of two proteins. However, this single value has no spatial interpretation in relation to the tertiary structure of the molecule. Therefore, we computed a piecewise distance for every pair of protein variants and mapped it to the tertiary structure. The sum δ of these distances serves as

a localised variant of the RMSD. The error of the participants' answer was calculated as the difference between the correct variant's δ_c and the chosen variant's δ_u. It was normalised to the global difference range of the respective data set to make the results comparable. Two outliers – one for the large high-resolution display (mono) and one for the desktop – have been excluded from the evaluation, because the error was an order of magnitude larger than for the rest of the participants.

We evaluated whether the display scenario is relevant for the relative error, which is depicted in Figure 5. Normal distribution of the data was rejected by the Shapiro-Wilk test. We therefore performed a Kruskal-Wallis test, which did not reveal significance ($H = 0.87$, 2 d.f., $p = 0.65$). Although participants were on average slightly faster on the desktop (see Figure 6), the difference is not significant ($H = 1.76$, 2 d.f., $p = 0.41$). Performing the tests for the data sets individually led to analogous results. We only found statistically significant differences for the correctness of answers between the three data sets ($H = 14.29$, 2 d.f., $p = 0.00079$). This is in accordance with our estimation of the respective difficulty. The timings, however, did not differ significantly between the data sets ($H = 0.38$, 2 d.f., $p = 0.83$).

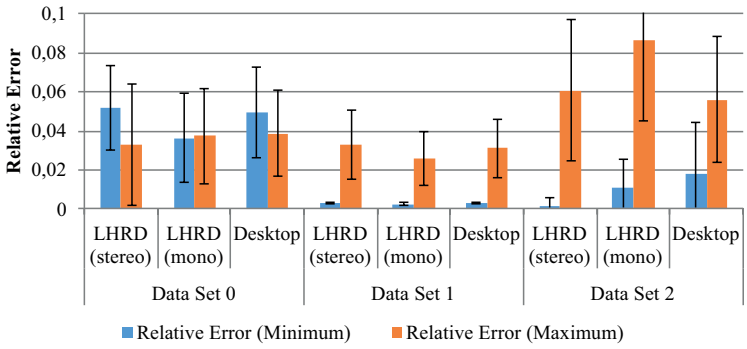

Figure 5. Averaged relative error that participants made when searching the variant with the minimum deviation and the maximum one from the baseline variant on the large high-resolution display (LHRD) and the desktop monitor. Error bars indicate standard deviation.

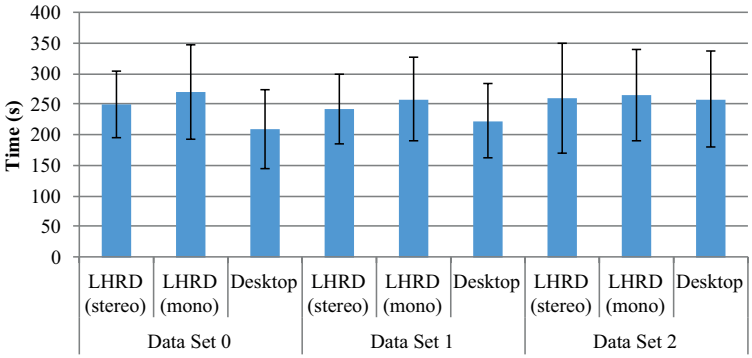

Figure 6. Performance averages for all data sets on the large high-resolution display (LHRD) and the desktop monitor. As the participants could search the least and most deviation in an arbitrary order, only the overall time for finding both is given.

Participants were generally very confident that they chose the correct variants (see Figure 7). In almost all cases, they found it easier finding the view with the minimum deviation and in turn also were more confident about this task. However, we found that there is a tendency towards anti-correlation (−0.0968) between measured error and confidence rating as well as between error and perceived easiness (−0.0973). That is, participants slightly underestimated the problem independently from the three display scenarios. This finding is independent from the three display scenarios.

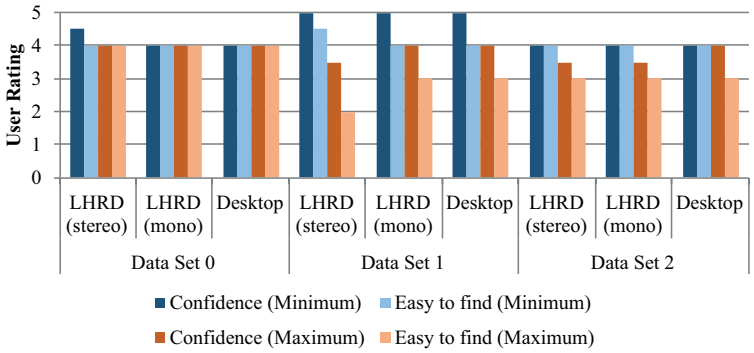

Figure 7. Medians of the confidence ratings by the users for finding the minimum and maximum deviation on the large high-resolution display (LHRD) and the desktop. Numbers are given on a five-level Likert scale.

The results of our user study back our hypothesis regarding pros and cons of the desktop and the large high-resolution display. As mentioned above, there is no significant difference in accuracy and time. This is in line with previous findings[36] as our tasks did not require virtual navigation in either setup and the whole data set could be visualised on all screens. Surprisingly, the participants did not perform significantly worse using the stereo display, despite the colour distortion.

User feedback and general observations

We asked participants to order the three test scenarios according to their preference for solving the task (cf. Figure 8). The desktop and the large display without stereo glasses clearly performed best with each of them being nominated seven times for the first place. The large high-resolution display in stereo mode was ten times on the last place. One of the most frequent complaints and the reason for choosing the powerwall in stereo mode as least preferred scenario was the lack of colour fidelity. Although three participants explicitly liked stereo and preferred the better spatial impression, there have also been complaints about the glasses being annoying by one subject and four participants said the stereo display was stressful and useless for the task. We believe that this led to a better rating (median 2.5 on a five-level Likert scale) in user fatigue for the stereo scenario. The difference between the other setups was smaller: median 1.5 for large high-resolution display (mono) and median 1 for the desktop monitor. Interestingly, eyestrain was a frequent reason for people disliking the desktop application. The size of the comparison matrix as well as the pixel density was oftentimes considered too low, thus requiring the user to concentrate more on the task, which in turn was perceived as exhausting. A

quarter of the participants chose the desktop as favourite scenario mainly due to familiarity. Users also felt being faster, mostly because they did not have to move physically and could move the mouse very quickly from one view to the next.

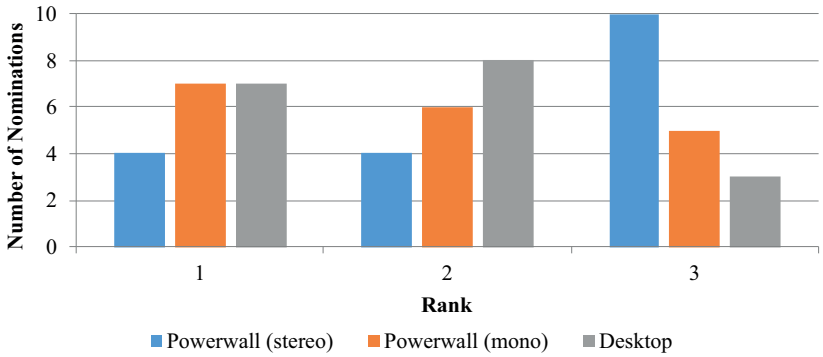

Figure 8. Histogram of how often participants chose a scenario as their preferred one (rank 1) and their least preferred one (rank 3).

Three participants explicitly noted a specific and fast problem solving strategy as reason why they preferred the desktop: By quickly exchanging the content of one of the detail views, one could examine very subtle colour changes on the (static) geometry quite easily. Twelve participants discovered this approach, which resembles the folding technique for comparison suggested by Tominski et al.[40], immediately or after some time. Only six of them realised that this was also possible on the large display. Some users further refined this strategy (on the desktop as well as on the powerwall) by systematically ordering the candidates around the best one and comparing them afterwards or by moving always the best candidate of a pairwise comparison further over the screen. People using this pattern on the desktop frequently used the second of the detail views as "external memory" that stored their currently best candidate.

Although some participants compared the variants in an almost chaotic pattern, eight of them sorted potential answers, mostly on a line in the middle of the screen. Users were most likely to align the important views on their eye level. Several suggestions have been made for supporting this initial sorting task, including a special area – preferably in the middle of the screen – that can be used to remember and compare all candidates. Furthermore, an additional interaction mode was suggested that enabled selection by clicking a button on the wand while it is in front of a certain view. This way, the user could walk once along the wall-sized display, select all candidates, and then compare only those. In the large high-resolution display cases, seven participants immediately started wandering around as the test started, while three participants did not move at all. Users generally did well with the wand, although they had little experience in using it. We believe that the relatively large targets that had to be hit for dragging the views were helpful here. Only two participants disliked the *object in your hand* metaphor and one suggested that the rotation angle should be exaggerated in order to enable faster rotations.

5 Expert Feedback

We showed the application employed for the user study and the actual surface-based comparison application (similar to Figure 9) to three researchers working in the field of technical biochemistry. They often work with large ensembles of simulated molecular structures. A common task for them is to find differences or similarities in the molecular structure that indicate whether the function of the proteins is similar or not. After familiarising with our setup, they were confident that their exploratory data analysis tasks would benefit from large high-resolution displays. They also believe that the possibility to see a comparison of many different data sets at once in full detail facilitates making unexpected scientific discoveries and observing correlations between multiple data sets. Especially for tasks that require more complex visualisations like the surface-based comparison explained above, they found that stereo output makes it easier to discern structural differences.

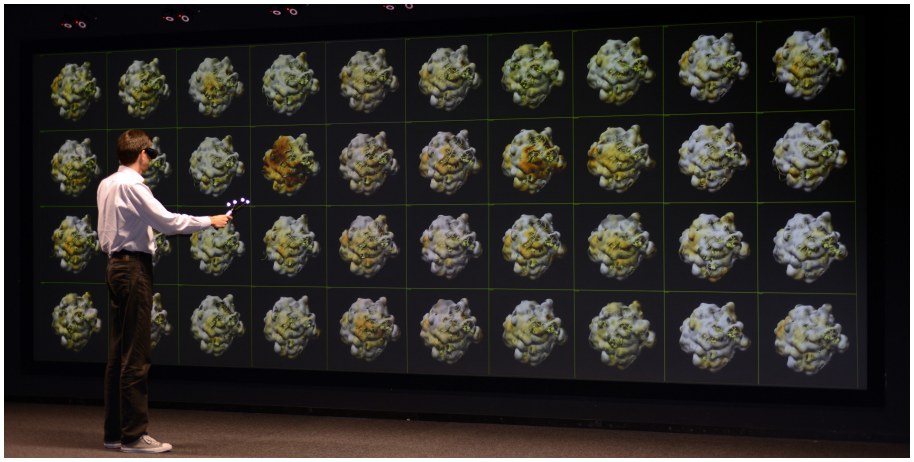

Figure 9. A selection of 10×4 views of the surface-based comparison visualized on our powerwall.

The discussion with the experts led to several ideas for extensions that would improve the usability in their application domain. One idea was to use the setup for parameter studies, for example for molecular simulations where a protein is simulated with two parameters – like the concentration of the solvents – that are varied independently. The small multiples view could be used as a 2D matrix to directly see the impact of the two parameters. This could be used to assess changes in binding sites and other functional areas or to see differences in the behaviour of the solvent.

Like the participants, the expert users also expressed that it would be helpful to reject samples immediately from the tiled visualisation. This would require reconfiguring the tile layout at runtime, which is possible using our application. The seamless nature of our projection-based display is beneficial for freely (re-)configuring the tiles as there is no physical subdivison of the display area.

6 Conclusions & Future Work

Comparative visualisation of ensembles can consume a lot of screen space for displaying all variants, making large high-resolution displays an obvious choice. Using a sample application from structural biology, we investigated the feasibility and potential benefits and drawbacks of such an approach in a user study. As this scientific visualisation application uses inherently spatial data, which suggest a 3D representation, we specifically included 3D stereo visualisation in our experiment.

We found that accuracy and timings do not significantly differ between our scenarios. Users had no clear preference towards large high-resolution display or desktop monitor, making the latter the economical choice. We believe that our findings can be transferred to ensembles of 3D objects that differ in colour, but not in geometry. In contrast to our expectations, we did not see significant negative effects of the stereo glasses. If geometrical differences are compared, the increased layers of depth in stereo large high-resolution displays[26] might be beneficial. An evaluation of this issue remains for future work, however, as well as the investigation whether a stereo large high-resolution display has significant benefits for spatial comparison tasks.

Expert feedback was very positive concerning the large high-resolution display. One reason for that is that an actual scientific question can require comparison of more than 100 variants. However, for our user study, we restricted the number of variants to 40 in order to allow our participants to complete the task in a reasonable amount of time. A user study with expert users performing actual exploratory analysis tasks could identify the benefits for the application domain.

Recently, displays with high pixel density – like 24 inch UHD ($3{,}840 \times 2{,}160$ px) displays – are gaining wider availability. Such devices might boost user performance on the desktop as they provide a resolution in a single device that has only been possible in the form of a tiled display before. It has also not yet been investigated how a desktop scenario using such a monitor stacks up against a large high-resolution display.

Acknowledgements

We want to thank Jürgen Pleiss, Tobias Kulschewski and Łukasz Gricman for their valuable feedback, as well as all participants of the user study. This work was partially funded by Deutsche Forschungsgemeinschaft as part of SFB 716.

References

[1] Anzali S, Barnickel G, Krug M, Sadowski J, Wagener M, Gasteiger J, Polanski J. The comparison of geometric and electronic properties of molecular surfaces by neural networks. J. Comput. Mol. Des., 1996, 10(6): 521–534.

[2] Bowman DA, Hodges LF. An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments. Proc. I3D. 1997. 35–ff.

[3] Bezerianos A, Isenberg P. Perception of visual variables on tiled wall-sized displays for information visualization applications. IEEE Trans. Vis. Comput. Graphics, 2012, 18(12): 2516–2525.

[4] Ball R, North C. An analysis of user behavior on high-resolution tiled displays. Proc. Interact 2005. 2005. 350–363.

[5] Ball R, North C, Bowman DA. Move to improve: Promoting physical navigation to increase

user performance with large displays. Proc. CHI. 2007. 191–200.
[6] Cruz-Neira C, Sandin DJ, DeFanti TA, Kenyon RV, Hart JC. The CAVE: Audio visual experience automatic virtual environment. Commun. ACM, 1992, 35(6): 64–72.
[7] Endert A, Bradel L, Zeitz J, Andrews C, North C. Designing large high-resolution display workspaces. Proc. AVI. ACM. 2012. 58–65.
[8] Finn PW, Morris GM. Shape-based similarity searching in chemical databases. Wiley Interdiscip. Rev. Comput. Mol. Sci., 2013, 3(3): 226–241.
[9] Gjerlufsen T, Klokmose CN, Eagan J, Pillias C, Beaudouin-Lafon M. Shared substance: Developing flexible multi-surface applications. Proc. CHI. 2011. 3383–3392.
[10] Grottel S, Krone M, Müller C, Reina G, Ertl T. Megamol – a prototyping framework for particle-based visualization. IEEE Trans. Vis. Comput. Graphics, 2015, 21(2): 201–214.
[11] Guo HQ, Yuan XR, Huang J, Zhu XM. Coupled ensemble flow line advection and analysis. IEEE Trans. Vis. Comput. Graphics, 2013, 19(12): 2733–2742.
[12] Hereld M, Judson IR, Paris J, Stevens RL. Developing tiled projection display systems. Proc. Immersive Projection Technology Workshop. 2000.
[13] Hofbauer C, Lohninger H, Aszódi A. SURFCOMP: A novel graph-based approach to molecular surface comparison. J. Chem. Inf. Comput. Sci, 2004, 44: 837–847.
[14] Hummel M,Obermaier H,Garth C, Joy KI. Comparative visual analysis of lagrangian transport in cfd ensembles. *IEEE Trans. Vis. Comput. Graphics*, 19(12): 2743–2752, 2013.
[15] Jansen Y, Dragicevic P, Fekete JD. Tangible Remote Controllers for Wall-size Displays. Proc. CHI. 2012. 2865–2874.
[16] Jakobsen MR, Hornbæk K. Sizing up visualizations: Effects of display size in focus+context, overview+detail, and zooming interfaces. Proc. CHI. 2011. 1451–1460.
[17] Jakobsen MR, Hornbæk K. Interactive visualizations on large and small displays: The interrelation of display size, information space, and scale. IEEE Trans. Vis. Comput. Graphics, 2013, 19(12).
[18] Johnson MA, Maggiora GM. Concepts and applications of molecular similarity. Wiley, 1990.
[19] Jakobsen MR, Haile YS, Knudsen S, Hornbæk K. Information visualization and proxemics: Design opportunities and empirical findings. IEEE Trans. Vis. Comput. Graphics, 2013, 19(12): 2386–2395.
[20] Kabsch W. A discussion of the solution for the best rotation to relate two sets of vectors. Acta Crystallogr. Sect. A, 1976, 34(5): 922–923.
[21] Krone M, Bidmon K, Ertl T. GPU-based visualisation of protein secondary structure. Proc. TPCG. 2008. 8. 115–122.
[22] Lee B, Isenberg P, Riche NH, Carpendale S. Beyond mouse and keyboard: Expanding design considerations for information visualization interactions. IEEE Trans. Vis. Comput. Graphics, 2012, 18(12): 2689–2698.
[23] Mine MR. Isaac: A virtual environment tool for the interactive construction of virtual worlds [Technical report]. Chapel Hill, NC, USA, 1995.
[24] Müller C, Krone M, Scharnowski K, Reina G, Ertl T. On the utility of large high-resolution displays for comparative scientific visualisation. International Symposium on Visual Information Communication and Interaction (VINCI). ACM. 2015. 8.
[25] Moreland K. Redirecting research in large-format displays for visualization. Proc. LDAV. 2012. 91–95.
[26] Müller C, Reina G, Ertl T. The VVand: A two-tier system design for high-resolution stereo rendering. CHI POWERWALL Workshop, 2013.
[27] Nam S, Jeong B, Renambot L, Johnson A, Gaither K, Leigh J. Remote visualization of large scale data for ultra-high resolution display environments. Proc. UltraVis. 2009. 42–44.
[28] Ni T, Schmidt GS, Staadt OG, Livingston MA, Ball R, May R. A survey of large high-resolution display technologies, techniques, and applications. IEEE Virtual Reality 2006. 2006. 223–236.
[29] Navrátil PA, Westing B, Johnson GP, Athalye A, Carreno J, Rojas F. A practical guide to large tiled displays. Proc. Advances in Visual Computing. 2009. 970–981.
[30] Otto M, Germer T, Theisel H. Uncertain topology of 3D vector fields. Proc. PacificVis. 2011. 67–74.

[31] Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera – a visualization system for exploratory research and analysis. J. Comput. Chem., 2004, 25(13): 1605–1612.

[32] Papadopoulos C, Petkov K, Kaufman A. Building the Reality Deck. CHI POWERWALL Workshop, 2013.

[33] Petkov K, Papadopoulos C, Zhang M, Kaufman AE, Gu XF. Interactive visibility retargeting in vr using conformal visualization. IEEE Trans. Vis. Comput. Graphics, 2012, 18(7): 1027–1040.

[34] Ruddle RA, Fateen W, Treanor D, Sondergeld P, Ouirke P. Leveraging wall-sized high-resolution displays for comparative genomics analyses of copy number variation. Proc. BioVis. 2013. 89–96.

[35] Richardson JS. The anatomy and taxonomy of protein structure. Advances in Protein Chemistry, 1981, 34: 167–339.

[36] Ruddle RA, Thomas RG, Randell RS, Quirke P, Treanor D. Performance and interaction behaviour during visual search on large, high-resolution displays. Inf. Vis., 2013.

[37] Shupp L, Andrews C, Dickey-Kurdziolek M, Yost B, North Chris. Shaping the display of the future: The effects of display size and curvature on user performance and insights. Human-Computer Interaction, 2009, 24(1–2): 230–272.

[38] Scharnowski K, Krone M, Reina G, Kulschewski T, Pleiss J, Ertl T. Comparative visualization of molecular surfaces using deformable models. Comput. Graph. Forum, 2014, 33(3): 191–200.

[39] Scheidegger L, Vo HT, Krüger J, Silva C, Comba JLD. Parallel large data visualization with display walls, 2012.

[40] Tominski C, Forsell C, Johansson J. Interaction support for visual comparison inspired by natural behavior. IEEE Trans. Vis. Comput. Graphics, 2012, 18(12): 2719–2728.

[41] Tan DS, Gergle D, Scupelli PG, Pausch R. Physically large displays improve path integration in 3D virtual navigation tasks. Conference on Human Factors in Computing Systems. 2004. 439–446.

[42] Tufte E. Envisioning Information. Graphics Press, Cheshire, CT, USA, 1990.

[43] van den Elzen S, van Wijk JJ. Small multiples, large singles: A new approach for visual data exploration. Comput. Graph. Forum, 2013, 32(3pt2): 191–200.

[44] Verma V, Pang A. Comparative flow visualization. IEEE Trans. Vis. Comput. Graphics, 2004, 10(6): 609–624.

[45] Ware C, Jessome DR. Using the bat: a six-dimensional mouse for object placement. IEEE Comput. Graph. Appl. Mag., 1988, 8(6): 65–70.

[46] Yang S, Chung H, North C, Fox EA. The effect of presenting long documents with large high-resolution displays on comprehension of content and user experience. Proc. Intl. Symposium on Electronic Theses and Dissertations. 2010.

[47] Yost B, North C. The perceptual scalability of visualization. IEEE Trans. Vis. Comput. Graphics, 2006, 12(5): 837–844.