

PROFEEL: LOW COST VISUAL-HAPTIC PERCEPTUALIZATION OF PROTEIN STRUCTURE-STRUCTURE ALIGNMENTS

MARC D. HANSEN, SURESH LODHA, ALEX PANG

*Computer Science Department
University of California, CA 95064
(mhansen, lodha, pang)@cse.ucsc.edu*

We describe *ProFeel*, an application for exploring molecular data from protein structure-structure alignments using a force-feedback joystick. Protein structure analysis is a useful application for multimodal information perceptualization because researchers in this area typically use several independent measures in assessment. A system that allows the user to simultaneously evaluate different criteria would therefore be quite useful. Four variables are assigned three different haptic effects each to allow the user to discern twelve separate data values. The haptic representation is coupled with a traditional molecular graphics visual display.

1 Introduction

In day to day life people use their five senses to obtain information from their environment. We interact with objects and develop intuition about the physical properties which describe them. For example, an average person shown a baseball will know from past experience approximately how heavy it is, how compressible it is, and how it will travel when thrown. Compare this familiarity to the interactions we normally have with a computer. In his 1965 paper, “The Ultimate Display”,¹ Ivan Sutherland pointed out that computers allowed us to create representations for concepts which we may not be able to perceive directly, such as intra-molecular forces. He described how, by adding sound and force-feedback to computer interfaces, it might be easier for us to develop a better understanding of such phenomena. Haptic, or force-feedback devices can push back against the user’s hands, thereby giving the illusion that the user is in contact with an actual surface or object. As an example of how a haptic controller could be used, consider a computer display indicating which parts of a molecule are more charged. We could use a haptic device with a spring-like feel associated with molecules so that when we clicked and dragged on an atom, we could feel the attractive and repulsive forces acting upon it. We could also add the sound of springs being stretched or compressed to enhance the effect.

Until recently, there were powerful economic and technical disincentives to using force-feedback devices. In the past, the cost of high fidelity force-feedback

controllers put them out of the reach of all but the most well-funded research scientists. Now haptic controllers such as Logitech's WingMan ForceTM 2 joystick can be widely purchased for under \$200. Writing programs for haptic devices was also complicated, often requiring the programmer to explicitly define the force functions for each force-feedback effect as well as the many parameters which influence such functions. Currently, graphical user interfaces such as Immersion Corporation's I-FORCE StudioTM 3 exist for easily describing physical models for effects. Instead of writing equations for spring and damper effects, the user can interact with a drawing of a spring, and stretch it to make it bigger or smaller. This representation can then be saved as a resource file to be read into the program and later modified on the fly if desired. At this point the hardware and software support for force-feedback devices are mature enough that researchers can shift their focus from developing realistic force-feedback effects to determining how these effects can be used to provide information effectively to users.

1.1 Protein Structural Alignment

Protein structural alignment is an important tool for analyzing proteins for several reasons. First, it allows us to detect similarities between different protein structures. These similarities can be used to build evolutionary models for how proteins change and develop over time. These models in turn, can help us understand how protein shape and function are related.

Second, studying structural alignments can tell us what amino acids may be substituted at positions in the structure without destroying the shape of the protein. This is so because protein structure is more conserved than the amino acid sequence, (i.e., the shape of a protein is less likely to change than its composition). This information is useful in protein structure prediction methods which rely on similarity between the protein of interest whose structure is not known (called the target) and a protein with a known structure. Using this information, a protein structure prediction may be evaluated according to whether the target protein sequence fits well with the structure if the target sequence is threaded through the structure. This additional data allows the researcher to make a better structure prediction despite possibly having a low amount of sequence similarity between the target and structure.

Finally, structural alignments can be used as test sets to quantify the abilities of protein structure prediction algorithms. For example, ignoring the fact that a sequence has a known structure, run the prediction algorithm using the sequence, and see if it comes up with the other proteins that have been structurally aligned to it.

2 Background and Previous Work

Previous work has shown that people perform well when presented with information in sensory modalities other than vision. In an experiment involving the use of data representation using visual and audio cues, Hansen et al.⁴ used a traditional molecular graphics display and musical parts to represent four variables simultaneously. This application presented information relevant to scientists examining alignments of two protein structures in three dimensional space.

Force-feedback has a long history of use in enhancing virtual reality type interfaces such as flight simulators,⁵ surgical simulations,⁶ and molecular docking.⁷ More recently the use of haptic devices to provide information in scientific visualization applications has been investigated. Hughes and Forrest⁸ developed a vibro-tactile mouse for exploring data presentations in which an additional variable not shown was accessible via the vibrations of the mouse as it was passed over portions of the visual display. Pao et al.⁹ successfully implemented haptic rendering of scalar, vector, and tensor fields. Fritz¹⁰ demonstrated the utility of the PHANTOMTM haptic device for representing data plots as virtual surfaces.

One problem with the the use of haptic rendering to represent information is the difficulty in choosing appropriate parameters that are intuitively associated with the visualizations. It is hoped that using an underlying physical object model might address this concern. One of the first researchers to use such a model was Cadoz.¹¹ He developed a user interaction paradigm based on a musical instrument as the physical object. Dufresne et al.¹² used this idea to develop a multimodal user interface system that combined auditory and haptic feedback. The metaphor of a plucked string was used for window stretching. When users stretched a window it felt as if they were stretching an elastic string and they simultaneously heard the sound of a plucked string rising in pitch. Dufresne et al. found that when compared to the unimodal interface, the bimodal interface allowed subjects to complete tasks more successfully and efficiently. Not surprisingly, the subjects also preferred the bimodal interface. Since users are familiar with the sights and sounds of objects in the real world, it is logical to expect that combining associated visual and haptic cues might result in stronger representations of information.

Minsky et al.¹³ describe a haptic system based on the object metaphor of sandpaper. Most people are familiar with sandpaper: they know what it looks like, and how it feels. Therefore if we use sandpaper to represent data values instead of an arbitrary texture map, we can use its associated feel to reinforce the data mapping. The sandpaper texture map along with its

corresponding feel can be varied so that the roughness varies with the value of the variable we wish to represent. The challenge then is to pick real life objects that have evocative visual and physical features which lend themselves to data perceptualization. Also, we want to pick these objects so that multiple parameters can be mapped with minimal masking of each other. An ancillary benefit of using physical objects and phenomena as a representational paradigm is that this facilitates the inclusion of sonification as a third sensory modality in future work: one simply adds an appropriate sound associated with the physical model.

To the best of our knowledge, no one has investigated the use of haptic textures to represent information in scientific visualizations. A haptic texture is a force-feedback representation of a physical texture, such as the sandpaper example by Minsky. Other textures that could be created using force-feedback include grooved surfaces similar to a vinyl record, or smooth and slippery surfaces like ice. Also, with the exception of Grabowski, et al.¹⁴ there has been little work on combining vision, hearing and touch to create scientific perceptualizations. Grabowski et al. built upon the previously described work of Fritz by enhancing the haptic data plots to use sound (pitch was mapped to the vertical coordinate, z).

3 Development Environment

ProFeel runs under Microsoft's Windows 95, and was written using the combination of Microsoft's Foundation Class library and their Visual C++ version 6.0 development environment. The computer used is a PC laptop running an Intel Pentium processor. The force-feedback device used is Logitech's WingMan Force joystick, which retails for approximately \$130. Immersion's graphical haptic effect development tool, I-FORCE Studio, was used to create resource files containing predefined force-feedback effects. These effects are described in terms of waves. The joystick responds to the waves in a manner analogous to that of a cork bobbing up and down in a tank of water. A sine wave used to describe motion along the y axis will cause the joystick will move smoothly backwards and forwards. A sawtooth wave like the one in Figure 5 describing motion along the x axis will cause the joystick to snap right rapidly and then immediately begin moving left at a slightly slower rate. Immersion's Feel Foundation Classes library was used to actually manipulate the joystick.

4 Architecture

ProFeel itself performs no data analyses. Instead, it reads the results from

other programs and turns them into haptic effects. Currently, our system reads structural alignments and their corresponding environment analysis files. The alignments are taken from the FSSP^{15,16,17} (Families of Structurally Similar Proteins) database. This database is widely considered to be an excellent source of structural alignments and has the added benefit of accessibility via the world wide web.^a

The analysis tool *Environments*¹⁸ is used to generate files containing information about each residue position's environment in the parent structure. (An amino acid's environment is its local space.) For each residue, four environmental parameters are determined: secondary structure, polarity of the region, exposure to solvent, and goodness-of-fit. The first three variables as well as the amino acid environment probabilities used in the goodness-of-fit calculation are extracted directly from the output of the *Environments* program. The score associated with each residue, environment pair is given by $\ln(P(i : j)/Pi)$ where $P(i : j)$ is the probability of finding residue i in environment j , and Pi is the probability of residue i in any environment. Bowie determined these probabilities using a database of 16 known structures and their aligned homologous sequences.

In analyzing a protein structure, one common question is: How likely is it for a particular amino acid to be found in its assigned location? To get a handle on this question, we derived a goodness of fit score. For each position in the alignment, goodness-of-fit is calculated taking the score for the parental structure's amino acid in its environment and subtracting the score for the aligned child structure's amino acid in that same environment. A positive score indicates that the child structure's amino acid is at least as likely as the parent structure's amino acid to be found in the given environment.

*Rasmol*¹⁹ provides the molecular graphics visualization. Inter-process communication between *ProFeel* and *Rasmol* is accomplished via a socket connection.

5 Visual-Haptic Mappings

We reimplemented Hansen et al.'s⁴ data sonification application using force-feedback instead of sound. The force-feedback effects may be presented singly or in any combination by clicking on the "haptics" checkbox available for each of the four variables (see Figure 1).

The four variables and their corresponding haptic mappings are:

1. Secondary structure Secondary structure, or local folding pattern (he-

^a <http://www2.ebi.ac.uk/dali/fssp/fssp.html>

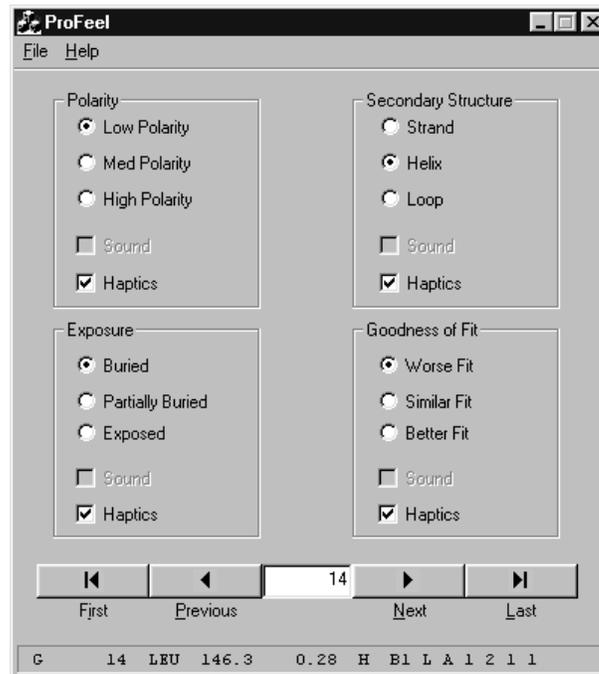


Figure 1: ProFeel User Interface

lix, loop, or beta sheet), is a nominal variable. The object metaphor for secondary structure is based on the physical model provided by the actual protein structure. Visually the secondary structure is represented using Rasmol's cartoon mode (see Figure 2), which uses traditional Richardson ribbons. Helices are shown as corkscrews, sheets as flat ribbons with arrowheads, and loops as thin spaghetti-like squiggles. In the haptic representation of secondary structure, the straight sheets are represented by a linear movement of the joystick from left to right (see Figure 3). The helices are indicated by a rotational movement of the joystick in a circle. Since loop regions are characterized by their lack of regular structure, they are represented by a wriggling movement of the joystick from front to back in which a sinusoidal wave causes the joystick to move smoothly back and forth between the left and right sides. The secondary structure representation is perhaps the most intuitive since in this case the physical structure of the protein lends itself to a direct haptic representation.

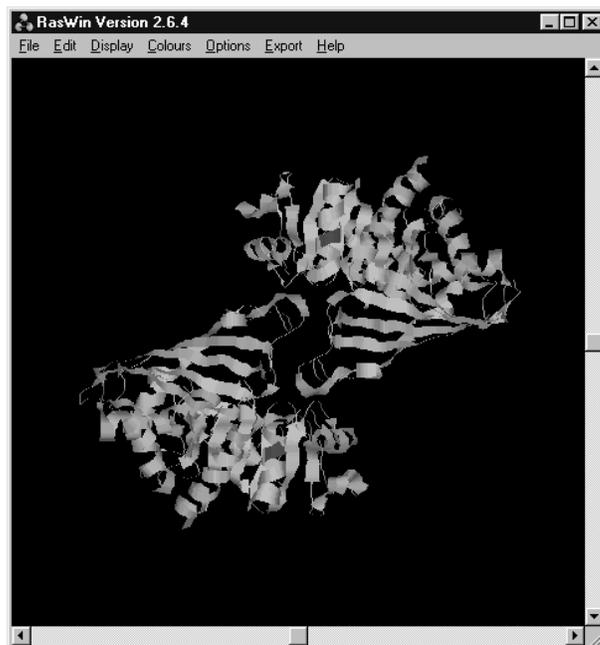


Figure 2: Rasmol User Interface

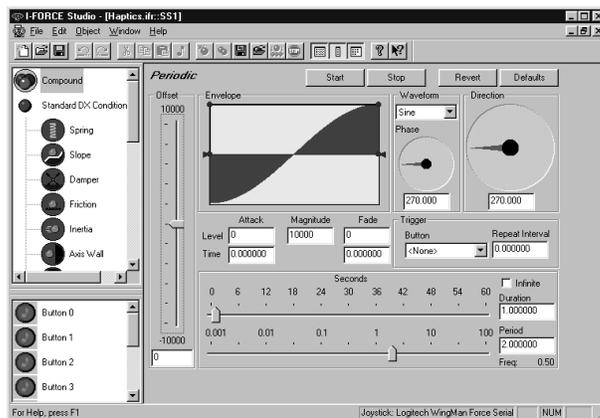


Figure 3: Example Force-Feedback settings for Secondary Structure (Strand)

- 2. Exposure** Amount of exposure (buried, partially buried, exposed), represents the proximity of the residue highlighted in color to the solvent on the outside of the protein, and is an ordinal variable. As in the case of secondary structures, the object metaphor for exposure is based on the physical properties of exposure in nature. A noise occurring in the middle of a large auditorium will likely be associated with more resonance than a noise occurring in a completely exposed environment such as the middle of an open field. Therefore, an amino acid found in the middle of the protein will be associated with higher resonance. Resonance is indicated through force-feedback by a 25 Hz vibration that builds and decays over a two second interval (see Figure 4). A maximally exposed residue is associated with little resonance, and the amplitude of the vibration is diminished. As the exposure decreases, the amplitude of the vibration increases. Visually, the amount of exposure of a residue may be estimated by rotating the molecule and observing the highlighted residue's position relative to the protein's outer surface.

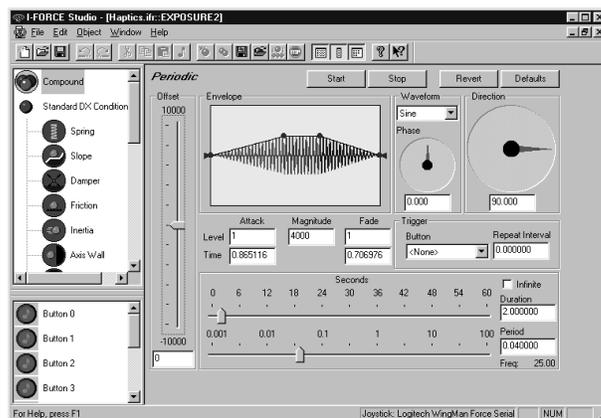


Figure 4: Example Force-Feedback settings for Exposure (Partially Buried)

- 3. Polarity** Polarity, or the amount of charge in the residue's local environment (low, medium, or high), is an ordinal variable. The model for the haptic representation of this variable is a Geiger counter. In a low polarity environment, a series of widely separated clicks are produced. A sawtooth wave function was used to make the clicks feel "sharp": the joystick moves in one direction and then quickly snaps back. As the polarity increases, the clicks increase in frequency from 4 Hz to 6.5 Hz and

finally 10 Hz (see Figure 5). Visually, polarity is related to exposure. More exposed environments tend to be highly polar due to the charged water molecules present at the exterior of most proteins. In future work, the frequency of the clicks could be indicated by a striped texture map in which there are more stripes associated with more frequent clicks.

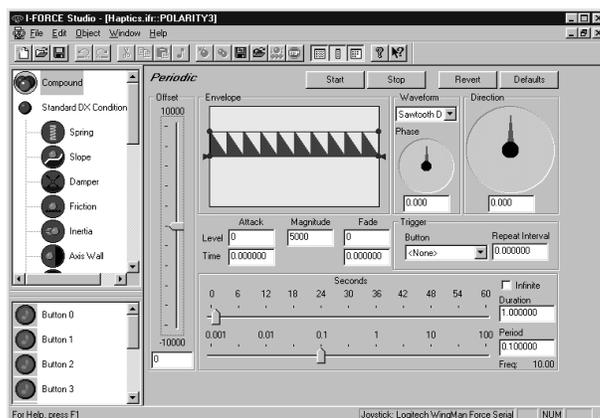


Figure 5: Example Force-Feedback settings for Polarity (Exposed)

4. Goodness-of-fit The “goodness-of-fit” of the amino acid substitution (worse, similar, or better) is an ordinal variable. The physical model for this variable is temperature. The visual representation for this variable is given by the color used to highlight the residue: “hot” red for worse, “cool” blue for better, and “warm” purple (a mixture of blue and red) for similar. Lack of data is indicated by yellow highlighting. Since temperature is related to the speed at which molecules vibrate, the haptic effect for this variable is a constant high frequency vibration. The “hotter” the color, the higher the frequency of the vibration, progressing from 50 Hz (See Figure 6), to 75 Hz, and then to 100 Hz.

6 Structural Alignment Data

6.1 Description and sources of data sets

The structural alignment consisted of the G chain of lobster D-glyceraldehyde-3-phosphate dehydrogenase (1gpd-G) superimposed on the salmonella typhimurium strain LT2 galactose-binding protein (1gca). The alignment was ob-

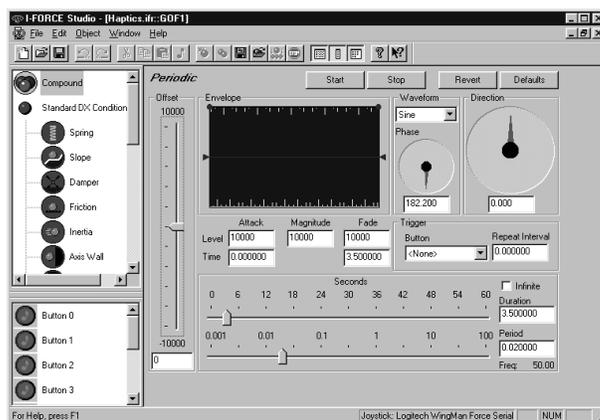


Figure 6: Example Force-Feedback settings for Goodness-of-Fit (Better Fit).^b

tained from the FSSP database. This particular alignment was chosen because the parent protein is an example of an alpha/beta structure, thereby allowing us to test all three of our secondary structure mappings. Also, since the sequences only have 56% identity, the alignment contained a nice spread of goodness-of-fit scores.

7 Conclusions and Future Work

In informal testing, force-feedback appears to have a useful role in disambiguating data which may be unclear if only presented visually. It remains to be seen whether masking effects will diminish the effectiveness of presenting multiple haptic effects simultaneously. One or two effects could be discriminated rather easily; discriminating three or four effects was much more difficult.

Current work on *ProFeel* is progressing on several fronts. In order to demonstrate the utility of this tool, user testing will be conducted over the next several months. The experimental design will be similar to the one discussed in our PROMUSE⁴ paper. The main enhancement for this project is the addition of data sonification capability. We are currently investigating sounds which could be paired with haptic effects to enhance the user's perception of the data values. An example would be a scratching sound paired with a sandpaper texture. One difficulty with using haptic effects such as the ones

^bThe reason the wave function appears solid in the picture is that the frequency is high enough that the drawings of the waves overlap almost completely.

described here is that they make it difficult to use the joystick as a pointing device. We are hoping that haptic textures will alleviate this problem. Using haptic textures, a user could move a mouse over a picture and feel textures like ice, sandpaper or grooves. Immersion corporation's FEELit mouse will support haptic textures. It is due for release before the end of the year.

Visit the following URL for current information on this work:
<http://www.cse.ucsc.edu/research/slv/bio.html>.

Acknowledgements

We would like to thank James Bowie and the UCLA-DOE Lab of Structural Biology and Molecular Medicine for allowing us to use their protein environment analysis software: *Environments*. Marc Hansen is supported by a GAANN fellowship. This project is supported by DARPA grant N66001-97-8900, ONR grant N00014-96-1-0949, NSF grant IRI-9423881, NASA grant NCC2-5281, and LLNL Agreement No. B347879 under DOE Contract No. W-7405-ENG-48.

1. Ivan E. Sutherland. The ultimate display. In *Proceedings of the IFIP Congress 65*, pages 506–508, 1965.
2. Logitech wingman force joystick home page. http://www.logitech.com/usr/joysticks/js01_100.html, 1999.
3. Immersion I-FORCE Home Page. <http://www.force-feedback.com/iforce/main.html>, 1999.
4. Marc D. Hansen, Erik Charp, Suresh Lodha, Doanna Meads, and Alex Pang. PROMUSE: A system for multi-media data presentation of protein structural alignments. In *Pacific Symposium on Biocomputing*, 1999.
5. Michael W. Haas. Virtually-augmented interfaces for tactical aircraft. *Biological Psychology*, 40:229–238, 1995.
6. Elaine Chen and Beth Marcus. Force feedback for surgical simulation. *Proceedings of the IEEE*, 86(3):524–530, March 1998.
7. Frederick P. Brooks Jr., Ming Ouh-Young, James J. Batter, and P. Jerome Kilpatrick. Project GROPE – haptic displays for scientific visualization. *Computer Graphics*, 24(4):177–185, August 1990.
8. R. G. Hughes and A. R. Forrest. Perceptualization using a tactile mouse. In *IEEE Visualization 96*, pages 181–187, 1996.
9. L. Y. Pao and D. A. Lawrence. Synergistic visual/haptic computer interfaces. In *Proceedings of the Japan/USA/Vietnam Workshop on Research and Education in Systems, Computation, and Control Engineering*, pages 155–162, Hanoi, Vietnam, May 1998.

10. Jason P. Fritz. Haptic rendering techniques for scientific visualization. Master's thesis, University of Delaware, 1996.
11. C. Cadoz, L. Lizowski, and J. L. Florens. Modular feedback keyboard. *Computer Music Journal*, 14(2):47–50, 1990.
12. Aude Dufresne, Odile Martial, and Christophe Ramstein. Multimodal user interface system for blind and “visually occupied” users: Ergonomic evaluation of the haptic and auditive dimensions. In Knut Nordby, Per Itelmersen, David J. Gilmore, and Svein A. Arnesen, editors, *Human-Computer Interaction Interact 95*, pages 163–168. International Federation for Information Processing (IFIP), Chapman & Hall, 1995.
13. Margaret Minsky, Ming Ouh-Young, Oliver Steele, and Frederick P. Brooks. Feeling and seeing: Issues in force display. *Computer Graphics*, 24(2):235–243, March 1990.
14. Nikolai A. Grabowski, M. Waqas Asghar, and Kenneth E. Barner. Joint haptic and aural methods for data visualization. In *Proceedings of the RESNA '98 Annual Conference*, pages 103–105, 1998.
15. L. Holm, C. Ouzounis, C. Sander, G. Tuparev, and G. Vriend. A database of protein structure families with common folding motifs. *Protein Science*, 1:1691–1698, 1992.
16. Lisa Holm and Chris Sander. The FSSP database of structurally aligned protein fold families. *Nucleic Acids Research*, 22:3600–3609, 1994.
17. Lisa Holm and Chris Sander. Fold classification based on structure-structure alignments of proteins (FSSP). *Nucleic Acids Research*, 26:316–319, 1998.
18. James U. Bowie, Roland Luthy, and David Eisenberg. A method to identify protein sequences that fold into a known three-dimensional structure. *Science*, 253:164–170, 1991.
19. Roger Sayle and E.J. Milner-White. RasMol: Biomolecular graphics for all. *Trends in Biochemical Sciences*, 20:374–376, 1995.