Pacific Symposium on Biocomputing 4:368-379 (1999)

PROMUSE: A SYSTEM FOR MULTI-MEDIA DATA PRESENTATION OF PROTEIN STRUCTURAL ALIGNMENTS

MARC D. HANSEN, ERIK CHARP, SURESH LODHA,
DOANNA MEADS, ALEX PANG

Computer Science Department
University of California
Santa Cruz, CA 95064
(mhansen,echarp,lodha,doanna,pang)@cse.ucsc.edu

We present and evaluate PROMUSE: an integrated visualization/sonification system for analyzing pairwise protein structural alignments (superpositions of two protein structures in three-dimensional space). We also explore how the use of sound can enhance the perception and recognition of specific aspects of the local environment at given positions in the represented molecular structure.

Sonification presents several opportunities to researchers. For those with visual impairment, data sonification can be a useful alternative to visualization. Sonification can further serve to improve understanding of information in several ways. One use for data sonification is in tasks such as background monitoring, in which case sounds can be used to indicate thresholding events. With PROMUSE, data represented visually may be enhanced or disambiguated by adding sound to the presentation. This aspect of data representation is particularly important for showing features that are difficult to represent visually, due to occlusion or other factors. Another feature of our system is that by representing some variables through sound and others visually, the amount of information that may be represented simultaneously is extended. Our tool aims to augment the power of data visualization rather than replace it.

To maximize the utility of our sonifications to represent data, we employed musical voices and melodic components with unique characteristics. We also used sound effects such as panning a voice to the left or right speaker and changing its volume to maximize the individuality of the sonification elements. By making the sonification parameters distinct, we allow the user to focus on those portions of the sonification necessary to resolve possible ambiguities in the visual display.

Sonifications of low level data such as raw protein or DNA sequences tend to sound random, and not very musical. We chose instead to sonify an analysis of data features, and thereby present a higher level view of the data. We also used brief melodic phrases rather than single notes in order to generate sounds that were more pleasing and musically idiomatic.

To validate the utility of our system, we present the results of an experiment in which PROMUSE was used to test the use of sound as an aid for clarifying visual information. We also compare the overall effectiveness of visual versus aural information delivery.

1 Introduction

The application of music to data representation has notable benefits. Not only can multiple values be given simultaneously, but the result can sound pleasant. Unfortunately, as the number of sonified variables increases, it can become increasingly difficult to accurately detect any single value. In order to minimize this problem, we decided to sonify a maximum of four parameters simultaneously. We based the parts for the sonifications in PROMUSE on a typical jazz quartet consisting of a bass line, drums, comp (chordal accompaniment), and a lead instrument for the melody. To increase the utility of our sonifications in representing data, we used voices and melodic components that were very distinct (see Table 1). We also employed sound effects such as panning a voice to the left or right speaker, and changing a voice's volume to maximize the individuality of the sonification elements. We used basic music theory (arranging, voice leading, development of complete melodic phrases, etc.) as the basis for our sonification parameters.

In order to test the utility of our program, we conducted an experiment in which we sonified certain aspects of individual amino acid local environments in a pairwise structural alignment (the superposition of two protein structures in three dimensional space). Visualizations were also presented using the molecular graphics program RasMol. [1] We then set out to examine possible interactions between presenting the data visually and aurally.

2 Previous Work

Several authors have previously pointed out the similarities between music and bio-sequences. Lubert Stryer [2] has compared the staggered array of tropocollagen molecules in a collagen fiber to a musical fugue. In his Pulitzer Prize-winning book *Gödel, Escher, Bach*,[3] Douglas Hofstadter likens mRNA to a musical tape, and amino acids to the "notes" resulting from translation.

Although there has been much less research focused on data sonification than visualization, there have been several important efforts in this area. A good source of reference for some early work on data sonification by several researchers including Sara Bly, Bill Gaver, Carla Scaletti, and Elizabeth Mynatt is the collection of papers edited by Kramer. [4] Recent applications of musical sounds for data exploration and visualization include efforts by Brady et al. [5] and Lodha et al. [6]

Hayashi[7] and Munakata[8] were among the first to publish on the possibility of using music to represent DNA sequences. Ohno [9,10] and Pickover [11] have also written about mapping DNA sequences to music.

Protein sequences are a bit more difficult to map to music because while there are 20 different amino acids which occur in naturally occuring proteins, in Western music there are only 12 different notes within an octave. Another difficulty in sonifying protein sequences is that most people do not have perfect pitch (the ability to name any note played) and would find it difficult to tell which of a possible 20 different notes was presented. For those visiting the San Francisco Bay area, the Exploratorium museum is currently showing an exhibition, "Musical Mutants",[12] in which protein multiple alignments are sonified by mapping notes to amino acids according to one of several different classification schemes. The protein sequences are played on different instruments to facilitate discrimination of the data. King et al.[13] sonified proteins with their DNA code sequences. In their work the melody line was derived from the four notes they chose for the DNA nucleotides. The bass line was formed by assigning 7 notes to amino acid characteristics (polarity, size, etc.) and sonifying the appropriate properties in sequence. PROMUSE[14] extends this idea of using musical parts to sonify protein data in several ways. First, our tool was designed not to replace information visualization with sonification, but rather to augment the power of visual representations of data. We therefore provide visual information along with our sonifications. Second, to make the aural information easier to interpret, each data value is assigned both its own voice as well as its own musical pattern. The use of complete melodic phrases instead of single notes results in a more musical and less random sounding composition. Third, compared to King et al., we double the amount of sound data available by using four separate musical parts instead of two.

3 Architecture

PROMUSE itself performs no data analyses. Instead, it reads the results from other programs and turns them into music. Our program displays information visually via RasMol, but this output can be turned off to enable PROMUSE to be used in conjunction with existing analysis and visualization tools such as DINAMO [15] and **ProtAlign** [16] which do not output any sound. PROMUSE can sonify a protein from start to finish, or a subset chosen via either a command line interface or a dialog box.

Currently, our system reads structural alignments and their corresponding environment analysis files. The alignments are taken from the FSSP [17,18,19] (Families of Structurally Similar Proteins) database. This database is generally accepted as containing excellent structural alignments and has the added benefit of accessibility via the world wide web.[a] The analysis tool *Environ-*

[a] http://www2.ebi.ac.uk/dali/fssp/fssp.html

ments [20] is used to generate files containing information for each position in the parent structure.

Four environment parameters are examined: secondary structure, polarity, exposure, and goodness-of-fit. The first three variables as well as the amino acid environment classifications used in the goodness-of-fit calculation are extracted directly from the output of the *Environments* program. 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 by taking the log-odds ratio for finding the parental structure's amino acid in its environment and subtracting the log-odds ratio for finding 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.

4 Audio-Visual Design

PROMUSE uses the RasMol molecular graphics program to visually represent the data. The protein is displayed in cartoon mode to explicitly show secondary structure (see Color Plate immediately preceding article). For structure-structure alignments, the cartoon display mode is useful because it allows the overall features of the protein to be shown with minimal clutter. We use red highlighting to indicate which section of the protein is being examined. The remainder of the protein is colored light gray.

Each residue environment is mapped to a musical interval of eight measures (lasting about 20 seconds). The overall tempo of the music is constant. We use the musical qualities of melody, rhythm, timbre and dynamics to create mappings of music to the values of local environment variables. The auditory mappings for PROMUSE are based on the idea of musical parts. Since we are sonifying four parameters, we based our arrangements on a typical jazz quartet consisting of a solo instrument to play the melody line, a drum part, a bass line, and a harmonic comp part (i.e., a rhythmic accompaniment consisting of the chords of the piece played on a keyboard instrument or guitar).

We map the four data parameters under investigation such that each parameter can take on three possible values. These values are indicated by the use of distinct voices and musical patterns (see Table 1). We made a conscious effort to make the mappings from the environment variables to music as intuitive as possible. For example, an exposed environment produces a brighter, sharper and busier sounding drum part; an environment with low polarity re-

Table 1: Auditory Mapping

parameter	*data value*	*part*	*voice*	*volume* [b]	*pan* [c]
Secondary Structure	Sheet	Bass	acoustic bass	127	0
	Helix		twangy bass	127	0
	Loop		slap bass	127	0
Polarity	Low	Comp	electric piano	127	127
	Medium		marimba	70	127
	High		electric guitar	70	127
Exposure	Buried	Drums	brush	127	64
	Partially Buried		cymbals	127	64
	Exposed		full kit	127	64
Goodness of Fit	Poor	Melody	trumpet	80	33
	Medium		saxophone	70	33
	Good		synthesizer	70	33

sults in a duller, softer, and more sparse piano part. Our primary variable of interest is goodness-of-fit. For this reason we chose to represent it using the melody line. We anticipated that the melody line would be the easiest for users to pick out when all four parts are combined. The drum voices are easy to vary from dull to bright, a sound mapping we associated with exposure. We mapped the bass line to secondary structure because secondary structure appears for several contiguous residue positions and the ear can tolerate a repeated bass pattern quite easily. Finally, we assigned the polarity variable to the remaining comp part.

We found that even at the same volume level, some voices sounded louder than others. To compensate for this we set individual volume levels for all the instruments. Some instruments (such as the saxophone) entered too softly, causing us to increase their note attack rates as a countermeasure. We also noticed that it was difficult to pick out individual parts when all the voices were sent to both the left and right audio channels. To correct this problem we moved the bass line to the left speaker and the comp part to the right. The melody tended to get lost in the comp part, so we moved it to the left where it could stand against the bass line. The drum part was relatively easy to hear, so we let it come out of both channels equally. To reiterate, we placed an emphasis on making the musical nature of these patterns both strongly distinguishable, and suggestive of the values they represent.

[b] Volume level may vary between 0 (min) and 127 (max).
[c] Pan setting may vary between 0 (left channel only) and 127 (right channel only)

5 Development Environment

We made three design choices that should make PROMUSE easy to port to other UNIX systems. First, to display the protein molecules we use RasMol. Second, we read music files in the General MIDI standard file format (MSF). Rather than using platform specific MIDI libraries, we make a call to the SGI MIDI player. This takes only one line of code, and will be easy to change for other systems. Third, the graphical user interface for the project was developed using the Forms Library for X.[21] RasMol, MIDI players, and the Forms Library for X are available on all the major versions of UNIX.

To create the MIDI files, we used Mark of the Unicorn Performer version 5.0, a multi-tracking sequencer for the Macintosh. The drum tracks were taken from the CD: "DrumTrax MIDI drum libraries".[22]

6 Experimental Design

6.1 Subjects, Collection Environment, and Tasks

In order to verify the utility of our program for enhancing visual representations of data, we conducted a controlled experiment on a total of 18 subjects. Most of the subjects were undergraduate and graduate computer science students from the University of California at Santa Cruz. Seven subjects rated themselves as having better than adequate musical abilities. Two rated themselves as having better than adequate experience with protein structures. Subjects were given headphones and allowed to use a workstation reserved for the experiment. The subjects then viewed or listened to protein data features and clicked on radio buttons in a graphical user interface to indicate which value they detected for each parameter presented. The subject responses were recorded, as was the length of time taken to answer.

6.2 Hardware and Software Specifications

We ran PROMUSE on a Silicon Graphics *Octane* running the IRIX6.4 operating system. The workstation contained a 195 MHz MIPS R10000 processor and 128 MB of RAM. Attached to the workstation was a 19 inch color monitor also from SGI. The headphones we used were Sony Digital headphones, model MDR-V6. To create our data visualization we used RasMol version 2.6.

6.3 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 obtained 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.

6.4 Experimental Flow Detail

Each subject trial took about 45 minutes, and consisted of five phases:

1. **Introduction** Subjects were given a two page overview of the experiment design and purpose. A general explanation of the experiment followed. This was followed by a brief question and answer session. After starting the program, subjects were given brief instructions on the basic layout and functions of the relevant controls on the user interface. Subjects then put on a pair of high-fidelity headphones. Subjects were allowed to adjust the volume settings, but not the left—right balance.

2. **Presentation of audio and visual mappings** Subjects pushed a "play" button to cause the program to simultaneously present a data sonification and its corresponding visualization (see section 4). Each of the four sonification parameters: bass, drums, comp, and melody, were presented to each subject in random order using a latin square design. The use of a latin square allowed us to minimize any systematic learning effects that might have otherwise appeared (for example if subjects consistently perform better on the drums tests only after having undergone the bass tests immediately prior). For each of the four sonification parameters the three allowable data values were presented in order from first to last. The appropriate radio buttons were lit to indicate which data values were currently being presented. The entire process was repeated twice for each of the four sonification parameters.

3. **Training** As with the presentation stage, all training was conducted with both the visual and auditory stimuli presented in tandem. A latin square was used to vary the order of the parameters. Four levels from each parameter were presented in random order: all three levels were tested at least once, with one level being tested twice. For bass and comp, the

third level was repeated. For drums the first level was repeated, while for melody it was the second level.

After being presented with a sonification/visualization pair, the subjects indicated which information they detected by clicking on a radio button. If the subject answered correctly, a green "Y" light came on. An incorrect selection caused a red "N" light to turn on. In either case, the auditory and visual information corresponding to the subject's pick were then presented, thereby allowing the subject to see and hear any potential differences between the data given and the choice.

4. **Testing** Data for the testing phase was presented in three modes: audio, visual, and audio+visual. A latin square was used to vary the order of the modes between subjects. Each mode was tested through to completion before proceeding to the next one. For each presentation mode, testing of individual parameters was followed by testing of all four parameters in combination. Subjects were exposed to the single parameters for 10 seconds. Presentations of all four parameters in combination lasted 20 seconds. The same randomization scheme used for the training sessions was used for testing the parameters individually. For testing the parameters simultaneously, we chose 8 combinations from random samplings of actual protein data to create a well represented parameter/value mixture. For each mode, all 8 combinations were presented in random order.

5. **Exit Questionnaire** Subjects were asked to complete a brief exit questionnaire for the purpose of obtaining feedback on qualitative aspects of the experiment. We were particularly interested in whether the sonifications intuitively matched the data represented.

7 Results

7.1 Overall Results

In order to test whether our sound mappings were as discernable and intuitive as we had hoped, we compared accuracy scores (i.e., what fraction of the time did subjects pick the correct data value) for the three different presentation modes. Overall, accuracy scores were much higher for the audio and audio+visual modes than for the visual mode (see Figure 1). In general, adding a visual component to the audio had a minimal impact on accuracy. As expected, secondary structure had the highest accuracy scores for the visual mode, while goodness-of-fit had the lowest. Our most striking success occurred with the exposure data. We anticipated that this information would be fairly

easy to classify in the visual mode. Instead we found that subjects had a very difficult time performing this task. Sonifying this information produced a drastic improvement in the accuracy scores.

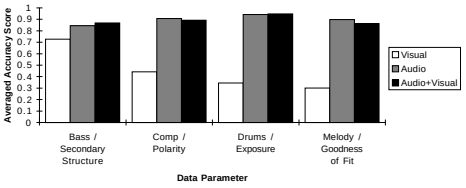

Figure 1: Overall results for the four data parameters in visual, audio, and audio+visual modes

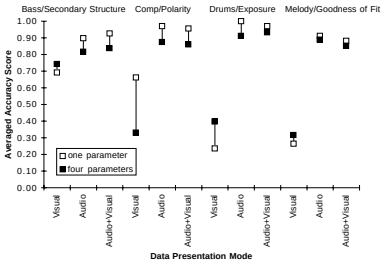

Figure 2: Accuracy of discrimination for single parameter presentations versus all four parameters in combination

7.2 Accuracy of Discrimination

One of our goals in this project was to determine whether we could develop a data-to-sound mapping in which the values of different variables could be distinguished while playing all four sonifications simultaneously. Since the presentations consisting of all four parameters in combination only lasted twice as long as the single parameter presentations despite containing four times as much information, we expected these presentations would produce lower accuracy scores due to the higher information to time ratio. We were therefore pleasantly surprised to see no such drop for the melody parameter (see Figure 2). The drops in accuracy that we did find for the other parameters were not nearly as steep as we had feared. One reason for the drops in accuracy being rather small may be that there are correlations in the data. If so, a subject able to correctly identify one item would have an easier task of correctly choosing the others due to the additional information available from the correlations.

7.3 Effect of Experience with Protein Structures

Two of our subjects had prior experience in visualizing protein structures depicted as cartoons. As expected, these subjects had a very easy time using visual information to determine the secondary structure of the highlighted location (see Figure 3). Their accuracy levels for this task were near 100% in the visual and audio+visual modes. Surprisingly, their results showed little if any improvement over the other subjects in identifying any of the other parameters.

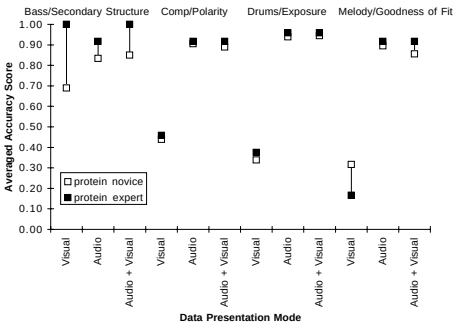

Figure 3: Effect of protein experience

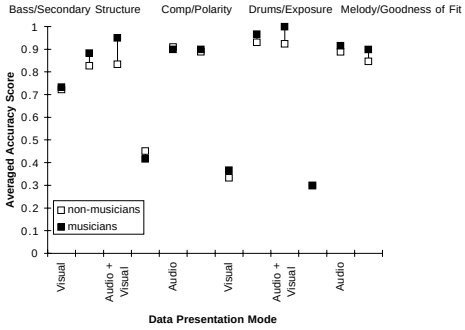

Figure 4: Effect of musical ability

7.4 Effect of Musical Ability

We expected that when it came to extracting information from the sonifications, subjects with a high self-rating in musical ability would perform better than subjects who rated themselves lower. It was anticipated that this difference would be most evident in the audio mode. In fact, the greatest improvements we found were in the audio+visual mode (see Figure 4). This finding was somewhat puzzling. Perhaps the subjects with greater musical ability had an easier time discriminating the musical information and were therefore able to concentrate more on the visual information to extract additional information in the audio+visual mode.

8 Conclusions and Future Work

Sonification appears to have a useful role in disambiguating data which may be unclear if only presented visually. We observed that the declines in classification scores resulting from simultaneous variable presentation can be minimized if care is taken to ensure that distinctive voices, rhythms, and melodic patterns are used for the different data parameters and their levels. For sonifications based on musical patterns, melody has the benefit of standing out well when multiple sounds are presented. Based on the results of this experiment, we are currently working on integrating PROMUSE with DINAMO and **ProtAlign**, our tools for performing visual analyses of structure-sequence alignments.

Currently PROMUSE sonifies four variables simultaneously. Through initial experimentation we determined that panning the sonification parameters to the left or right speakers facilitated discrimination. A logical extension of this finding would be to output the sound in quadrophonic rather than stereo,

and assign each parameter its own sound channel.

For researchers who spend a lot of time analyzing proteins, it is possible that listening to similar music all day might become tedious. Natural sounds might be a viable alternative in this case. Rather than using instrument voices, data parameters could be mapped to sounds like ocean waves, babbling brooks, wind, rain or bird calls. These types of sounds might be especially good for background monitoring tasks.

To read more about PROMUSE and to hear the sound files used visit the following URL http://www.cse.ucsc.edu/research/avis/bio.html.

Acknowledgments

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, and NASA grant NCC2-5281.

References

1. Roger Sayle and E.J. Milner-White. RasMol: Biomolecular graphics for all. *Trends in Biochemical Sciences*, 20:374–376, 1995.
2. Lubert Stryer. *Biochemistry*, page 270. W.H. Freeman and Company, 1988.
3. Douglas Hofstadter. *Gödel, Escher, Bach: an Eternal Golden Braid*, page 519. Basic Books, 1979.
4. G. Kramer. *Auditory Display, Sonification, Audification, and Auditory Interfaces*. Addison-Wesley, 1994.
5. R. Brady, R. Bargar, I. Choi, and J. Reitzer. Auditory bread crumbs for navigating volumetric data. In *Proceedings of the Late Breaking Hot Topics of IEEE Visualization '96*, pages 25–27. IEEE Computer Society Press, 1996.
6. S. K. Lodha, T. Heppe, J. Beahan, A. Joseph, and B. Zane-Ulman. MUSE: A musical data sonification toolkit. *Proceedings of the International Conference on Audiotry Display (ICAD)*, pages 36–40, November 1997.
7. Kenshi Hayashi and Nobuo Munakata. Basically musical. *Nature*, 310:96, Jul 1984.

8. Nobuo Munakata and Kenshi Hayashi. Gene music: Tonal assignments of bases and amino acids. In Clifford A. Pickover, editor, *Visualizing Biological Information*. World Scientific Publishing Co. Pte. Ltd., 1995.
9. S. Ohno and M. Ohno. The all persuasive principle of repetitious recurrences governs not only coding sequence construction but also human endeavor in musical composition. *Immunogenetics*, 24:71–78, 1986.
10. S. Ohno. Of words, genes, and music. In E. Sercarz, editor, *The Semiotics of Cellular Communication in the Immune System, NATO ISI Series*, volume H32, pages 131–147. Springer, 1988.
11. Clifford A. Pickover. There is music in our genes. In *Mazes for the Mind: Computers and the Unexpected*. St. Martin's Press, 1992.
12. Musical mutants. The Exploratorium Museum: 3601 Lyon Street, San Francisco, CA 94123 Tel: 415-563-7337, 1998. Also visit http://www.exploratorium.edu/genepool/exhibits.html.
13. Ross D. King and Colin G. Angus. PM — Protein music. *CABIOS*, 12(3):251–252, 1996.
14. Marc Hansen and Erik Charp. Multi–modal visualization of local environment data for protein structural alignments. Technical Report UCSC–CRL–98–08, UCSC Computer Science Department, 1998.
15. Marc Hansen, Jesse Bentz, Albion Baucom, and Lydia Gregoret. DINAMO: a coupled sequence alignment editor/molecular graphics tool for interactive homology modeling of proteins. In *Pacific Symposium on Biocomputing*, volume 3, pages 106–117, 1998.
16. Marc Hansen, Doanna Meads, and Alex Pang. Comparative visualization of protein structure–sequence alignments. In *IEEE Information Visualization*, 1998. To appear.
17. 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.
18. Lisa Holm and Chris Sander. The FSSP database of structurally aligned protein fold families. *Nucleic Acids Research*, 22:3600–3609, 1994.
19. Lisa Holm and Chris Sander. Fold classification based on structure–structure alignments of proteins (FSSP). *Nucleic Acids Research*, 26:316–319, 1998.
20. 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.
21. T.C. Zhao. XFORMS home page. URL: http://bragg.phys.uwm.edu/xforms.
22. DrumTrax. DrumTrax MIDI drum libraries, 1997.