EVALUATION OF A MULTIMODAL SONIFICATION AND VISUALISATION OF DEPTH OF MARKET STOCK DATA

Keith V. Nesbitt

The University of Sydney
School of Information Technologies
NSW 2006, Sydney, Australia
knesbitt@cs.newcastle.edu.au

Stephen Barrass

CSIRO
Virtual Environments
P.O. Box 664, Canberra 2601, Australia
Stephen.Barrass@csiro.au

ABSTRACT

Day traders make minute by minute decisions from depth of market stock data read from a table. A visualisation has been designed to help traders make better decisions from extra context and history in the display. An expert review of the visualisation identified the need for more dynamic information and higher resolution in critical areas. We designed a sonification to complement the visualisation with this information in a multimodal display. We evaluated the visualisation, sonification and multimodal displays with 15 non-experts. The subjects predicted price movements significantly better than chance, at 61.3% correct with the visualisation, 70% with the sonification and 70% with the multimodal display. The prediction of downward movements was significantly better than upward with a best of 83% for downward with the sonification, compared with 65.5% for upward with the multimodal display. There was no effect due to the order of experience with the different displays. The results suggest that the multimodal display is redundant for downward movements and complementary for upward. The subjects commented that the sonification provides recent trends while the visualisation provides context.

1. INTRODUCTION

Technical analysis of stock market data has the premise that all information about the market is contained in the data. Traders who take this approach make trading decisions from patterns that they observe in the data - for example, a buy decision might be based on the observation of three consecutive peaks in price. Recognition of new patterns may provide an edge over other traders. The domain expert we are working with would like to extend the technical analysis approach to more complex depth of market stock data that has not been analysed in this way before.

Traders currently read depth of market stock data from a numeric table. We hypothesise that perceptual displays could enable a better understanding of more data more quickly. In previous work we developed a 3D visualisation to provide context and history not available in the conventional table [1]. An expert trader reviewed the visualisation and identified extra information that could be useful for making trading decisions. This review was used to specify and design a sonification to complement the visualisation in a multimodal display. The effectiveness of the visualisation, sonification and multimodal displays were tested in an experiment with 15 subjects.

In the following sections we look at previous work on the sonification of stock market data, describe our 3D visualisation of depth of market stock data, present the expert review of the visualisation, and design a sonification to complement the visualisation. In the second part of the paper we evaluate the visualisation, sonification and combined multimodal displays with non-expert subjects. We then analyse and discuss the results and present conclusions and directions for further work.

2. BACKGROUND

The Japanese method of Candlestick charting was used around 1710 in the worlds oldest futures market [2]. Today we commonly use charts to show price bars that record the opening, closing, maximum and minimum price for a period of trading. A graph of closing price over time is used for longer term analysis and often augmented by a volume histogram. Computers have been used to chart larger amounts of data, and to produce charts from live data feeds in real time from the market. New visualisations have been developed to enable the perception and analysis of more information more quickly, such as treemaps for visualising stock portfolios [3] and Wright's visualisation of live data feeds of liquidity information [4]. Wright's work is one of the few attempts to visualise depth of market data, but does not show historical information.

Stock market data has also been a subject for sonification researchers. Stock market price data was directly played back as a sound waveform, using the technique called audification, by Frysinger [5]. The sounds were difficult to interpret and one reason may have been that the stock market does not follow physical-acoustic laws so audification does not produce natural or 'ecological' sounds that can be understood from everyday listening experience.

Stock market price data for a single share was mapped to pitch in an experiment by Brewster et. al [6] to test the idea that sounds can be used in place of limited screen space in a mobile PDA device. The task was to make as much profit as possible by buying and selling shares using either the sonification or a line graph to monitor the price over time. Results showed no difference in performance between the two modes, but subjects reported a significant decrease in workload with the sonification which allowed them to monitor the price while using the visual display to buy and sell shares.

Both price and volume for a single stock were sonified in an experiment on the orthogonality of a multivariate mapping by Nuehoff et al. [7]. Price was mapped to the frequency of a tone and volume to the amplitude. The task was to judge the price and volume from the sound. The stock data were simulated for the experiment. Results show perceptual interactions and asymmetries could potentially distort the interpretation of the display. Judgements of price from frequency were effected by change in amplitude as well. Likewise, judgements of volume from amplitude were effected by change in frequency. Judgements of volume from amplitude were effected more by increases than equivalent decreases.

Price and volume for two stocks were sonified simultaneously by Ben-Tal et al. [8]. The two stocks were mapped to perceptually distinct vowel-like sounds. The closing price of the day was mapped to the number of bursts of sound and the volume of trade to the duration of the bursts. The data for each day was mapped to a second of sound for periods of up to a year. They informally observed that they could categorise high volume, high price trading days as loud, dense sounds, while low volume, low price days were heard as pulsed rhythmic sounds.

We are not aware of prior work on multimodal display of stock data. We characterise multimodal displays based on McGee et al. [9]

- Conflicting - contradictory information is displayed to each sense. Performance is worse in the combined display.
- Redundant - the same information is displayed to each sense. Performance of single modality and combined displays is the same but there may be a reduction in workload or an increase in confidence.
- Complementary - different information is displayed to each sense. Performance with the combined display is superior to the separate single modality displays.

3. TASK SCENARIO

The *depth of market* refers to the number of buyers and sellers currently trying to trade a financial instrument such as a company share or future contract. A buyer may make a *bid* to purchase a specified volume of shares while at the same time a seller may *ask* a price for some specified volume. The balance of bids and asks determines the state of the current market. The difference between the highest bid and lowest ask is known as the *spread*. The task of buying and selling shares to make a profit on short-term variations in market prices is called discretionary trading. The emphasis is on making a small profit many times during the period of trading. A trader may decide to sell when the volume of bids around the last trade price outweighs the volume of asks, or conversely sell if asks outweigh bids.

The depth of market data is displayed in a table, like Table 1, which updates every 30 seconds or so. The highest bid and lowest ask at the top show the spread. A wider spread usually indicates a lower likelihood that a trade will occur. The price of the last trade is shown for comparison with the current spread. Other more peripheral information includes the volume of stock in bid and ask quotes, and context of lower bids and higher asks.

Although the table is useful for quantitative analysis it requires mental arithmetic to read relations between elements, and a good memory to make comparisons with previous overwritten data. We would like to enable quicker, more informed decisions through perception of higher level information in patterns across a wider range of the data and over longer time periods.

We are particularly interested in temporal patterns that may indicate the trading direction.

Buyers		*Trade*		*Sellers*
Volume	Price	12.03	Price	Volume
14,553	12.03	1	12.04	42,450
28,850	12.02	2	12.06	20,540
23,000	12.02	3	12.07	8,261
2,121	11.99	4	12.09	35,000
41,000	11.98	5	12.10	120,515
17,000	11.97	6	12.11	574

Table 1. *Table of Depth of Market Data*

4. VISUALISATION

The visualisation was designed to provide extra context and historical information that is not easy to understand from a conventional tabular display [1]. The data are represented by a series of surfaces formed by the volume of bids, asks and trades at each price, as shown in Figure 1. Price is on the X axis, volume on the Y axis and updates in time are appended on the Z axis. Bids are coloured yellow and asks are coloured green to separate them. The highest bid and lowest ask are next to each other close to the centre of the display while less important data spread to the periphery. The last trade is shown as a red ribbon that typically tracks the highest bid and lowest ask.

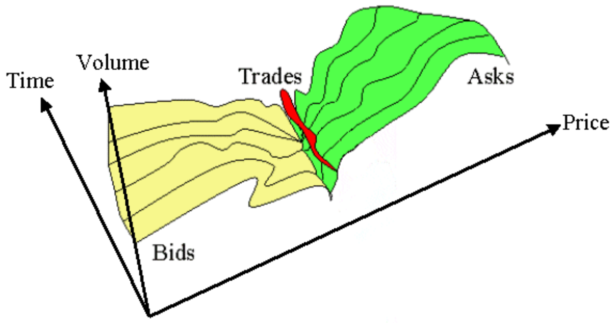

Figure 1. *Visualisation of Depth of Market Data.*

As updates accumulate the visualisation takes on the form of a 'landscape' that looks like a valley between two hills with a river flowing through it. This 'ecological' metaphor may help users interpret the visualisation from familiar natural properties such as 'steepness' of cliffs or 'height' of the hills.

We asked a practicing trader for an heuristic review of the visualisation. His overall impression was that it could have advantages for predicting the direction of the next trade because it shows more information more succinctly than a table. He thought that separation of the two hills gave a good indication of spread and that the steepness and slope of the sides were a good indication of the 'demand' for the stock. However, he also thought that the most interesting information could be obscured down in the valley between the hills and at too low a resolution. He also thought that the 15 minute update rate was not dynamic enough and suggested 30 second updates like the web-based table. Overall the review identified that:

- The visualisation could be very useful for predicting the direction of the next trade.
- More information is needed in the valley region.
- The most recently updated information should be provided immediately.

5. SONIFICATION

The review of the visualisation was used to specify the design of a sonification to provide additional information for predicting the direction of the next trade.

Our sonification design method has two levels - a schema level and a perceptual level [10]. At the schema level we identified a 'market place' metaphor from the task scenario and comments from the review of the visualisation. In this market place vendors shout the price of produce and shoppers reply with offers or agree to trade. Listeners may interpret 'direction' of the next trade from this familiar experience. At the perceptual level we map information onto perceptually scaled auditory variables, and address issues of perceptual grouping and segregation in the overall auditory scene. The resulting sonification should enable quick, confident and accurate answers to the global question "what is the direction of the next trade?" - either up or down. It should also enable answers to intermediate questions about relations between data elements such as "how wide is the current spread between bids and asks?" and "where is the current activity relative to the last trade?". At the local level it should allow answers to questions about individual elements such as "are there any bids?" or "what is the volume of the most recent ask?".

The schema level was implemented with samples of a male voice saying the word 'buy' and a female voice saying the word 'sell'. The male voice was pitched at E5 and the female voice at G5. The samples were edited to 0.25 seconds duration. The difference in timbre, vowel formant, and pitch of the samples allows both sounds to be heard when played simultaneously.

The data consists of price and volume of bids, asks and trades. Price data were mapped into seven ordered information categories representing the importance of the bid or ask relative to the last trade, as shown in Table 2. This fish-eye mapping emphasizes bids and asks close to the price of the last trade.

Bid/Ask– last trade (cents)	*Price Importance*
-4 or more	- low
-2,-3	- mid
-1	- **high**
0	0 **high**
1	+ **high**
2,3	+ mid
4 or more	+ low

Table 2. *Mapping from Price Data to Price Importance*

Volume data was mapped to three ordered levels of importance using a logarithmic scale as shown in Table 3.

Volume Data	*Volume Importance*
<= 10,000	low
<= 100,000	mid
> 100,000	**high**

Table 3. *Mapping from Volume Data to Volume Importance*

The combined mapping from *Price Importance* and *Volume Importance* to overall *Importance* is shown in Table 4. This mapping reflects that volume close to the last trade price can strongly influence the direction of the next trade, but that volume is not important in the price periphery.

Importance	*Volume*		
Price	***High***	*Mid*	*Low*
- low	- low	- low	- low
- mid	- mid	- low	- low
- high	**- high**	**- high**	**- high**
0 high	**0 high**	**0 high**	**0 high**
+ high	**+ high**	**+ high**	**+ high**
+ mid	+ mid	+ low	+ low
+ low	+ low	+ low	+ low

Table 4. *Mapping from Price and Volume to Importance.*

Next we map the combined *Importance* information into the auditory display, as shown in Table 5. The seven ordered information categories are mapped to seven ordered pitch categories C5-B5 within the same octave. This mapping allows the perception of order without overlapping harmonic or octave pitches. However pitch does not allow the perception of the central zero. We addressed the zero by a redundant mapping to seven categorical spatial locations in the stereo display, ordered from left to right with the zero in the middle. The listener can identify the 'zero' elements at the spatially absolute centre of the display. As price diverges from the last trade the sound moves further from the central position. The mapping of information categories to both pitch and space increases the perceptual segregation and perceived order of the categories. The three ordered categories of Volume information are mapped to three equal steps in Loudness.

Importance	Pitch (notenum)	Pan (degrees)	Loudness (range)
- low	C5	-90	low
- mid	D5	-60	low, mid
- high	E5	-30	low,mid,high
0 high	F5	0	low,mid,high
+ high	G5	30	low,mid,high
+ mid	A5	60	low,mid
+ low	B5	90	low

Table 5. *Mapping from Importance to Pitch, Pan and Loudness.*

Bids and asks are heard as they are made and come from the left or right as though from a crowd. If the bid or ask is lower than the last trade it is heard to the left, if it is the same then it comes from the centre and if it is higher it comes from the right. A flurry of bids to the right could indicate demand to buy at a higher price than the last trade and could indicate upward movement in trade price. A pattern of bids to the left mixed with asks to the right might indicate equilibrium in the market.

The table of bid and ask data is updated every 30 seconds. In our display of historical data we play an hour of data per minute with data updates every 0.5 seconds. The sounds are scheduled at 0.01 second intervals for up to 50 new entries each frame. A highly active market is a continuous hubbub 'texture'. A mid-range of activity of around 20 events per update sounds intermittent and overlapping. When market activity is low individual events can be heard, while a lack of activity is silent.

Data per frame	Activity	Sound
none	none	Silence
<5	low	Separate events
<20	mid	Overlapping events
>20	**high**	Hubbub texture

Table 6. *Mapping of temporal data to Activity and Sound*

6. MULTIMODAL

The multimodal depth of market display was implemented on a BARCO stereo Projection Table, shown in Figure 2, in the CSIRO Virtual Environment Lab. The BARCO screen is connected to an SGI Onyx2 computer with synced shutter glasses and Polyhemus head-tracking. The sound is provided over Sennheiser HD540 headphones. The visualisation is built on the Avango Virtual Reality framework [11] running on an SGI Onyx2 computer. Data updates are sent from the visualisation to the sonification using UDP network protocol. The sonification was built using the Avango sound server [12] and Max synthesis system [13].

Figure 2. *Experimental apparatus.*

7. EXPERIMENT

Our null hypothesis is that subjects cannot predict the direction of the next trade from these displays. The alternative hypothesis that subjects can predict depends critically on the technical trading hypothesis that the information needed is contained in the data. Since we rely on this premise the data we used was recorded from real trading data for two shares over a day of trading on the Australian Stock Exchange. The data were divided into 6 subsets, three for training and three for evaluation. These subsets were randomly allocated to visual training, visual evaluation, auditory training, auditory evaluation, multimodal training and multimodal evaluation for each Subject. The Order of presentation of the different Modes of display (V,A,M) was also randomised.

The experiment was designed to answer the questions:

- Can people use the visual (V), auditory (A) and combined multimodal (M) displays to predict the direction of the next trade from depth of market data?
- What differences are there in performance with the visual (V), auditory (A) and multimodal (M) displays?
- Do people find consistent patterns in the data?
- How do people interpret and make decisions from these displays?

Each Subject carried out the experiment one at a time with the researcher. At the start the Subject was given a written introduction to trading with depth of market stock data, and allowed to ask questions. They then carried out a training session followed by an evaluation session for each Mode. In the training session the Subjects were shown a display of historical data which was paused at 10 random points where they were told the direction of the next trade - up or down. In the evaluation the Subjects were asked to predict the direction of next trade at 10 random pause points as either up or down. The up and down movements were between 1 and 7 cents, with 80% of the decision points involving only 1 or 2 cent changes. After each evaluation the Subjects were asked for comments.

8. RESULTS

There were 13 male and 2 female subjects between the ages of 20 and 42. Only one subject had any familiarity with depth of market data and none had traded on the stock market. We recorded 10 predictions for each Subject in each of the 3 Modes (V, A, M). The experiment typically took 45 minutes.

We analysed totals correct out of 10 by regression analysis and found no significant effect for variation in Subject, Mode or Order, as seen in Table 6. This lack of effect led us to analyse proportions of correct predictions for trades that went up in price separately from trades that went down using generalised linear models [14]. In the up direction we again found no significant effects. However, in the down direction we found that some Subjects performing significantly better than others (P=0.014), and a significant variation with Mode (P=0.029). However the Order was still not significant, as in the up direction.

Direction	*Subject*	*Mode*	*Order*
all	0.884	0.397	0.953
up	0.981	0.812	0.793
down	**0.014**	**0.029**	0.653

Table 6. *P values for significance (P<0.05) of main variables.*

Next we compared how well the Subjects were actually predicting the direction of the next trade, using the regression model to estimate mean proportions correct (PC) and standard error (s.e.) for one binomial trial, and using an exact two-tailed binomial to calculate probability (P) of the result occurring by chance. This analysis shows Subjects predict the direction of the next trade at levels significantly above chance in all three Modes, with A (70%), M (70%) and V (61.3%), as shown in Table 7.

D	*M*	*#*	*T*	*%*	*PC*	*s.e.*	*P*
all	V	92	150	61.3	0.613	0.047	**0.0069**
	A	105	150	70.0	0.070	0.047	**0.0000**
	M	105	150	70.0	0.070	0.047	**0.0000**
up	V	50	85	58.8	0.583	0.077	0.1284
	A	46	78	59.0	0.595	0.079	0.1405
	M	42	64	65.6	0.658	0.085	**0.0169**
down	V	43	64	67.2	0.649	0.050	**0.0081**
	A	59	72	81.9	0.831	0.036	**0.0000**
	M	63	86	73.3	0.734	0.040	**0.0000**

Table 7. *Experimental results for all, up and down trades. Table shows Mode (M), number of correct predictions (#), total number of predictions (T), percentage correct prediction (%), estimate of mean proportions correct (PC) and standard error (s.e.) for one binomial trial, and the exact two-tailed binomial of calculated probability (P).*

Subjects were best at predicting down trades from A (83.1%) followed by down trades from M (73.4%), up trades from M (65.8%), and down trades from V (64.9%). The prediction of up trades from V (58.8%) and A (59.0%) separately was not significantly above chance. The effectiveness of M in the up direction where the individual V and A are not so effective indicates that complementary information from both Modes is needed to predict up trades. Subjects predict down trades much better than up trades, at levels so far above chance that the overall analysis shows significant prediction for all 3 Modes despite the up results. The prediction of down trades was best with the A (81.9%), followed by M (73.3%), and then V (67.2% correct). These results suggest that A is the best source of information for predicting down trades, and that the combination of A with V reduces performance in the M display.

One Subject commented that there were certain points in the display where they felt very certain about their decision, and others where they were not so sure. This could indicate that there are patterns at some places in the data that are particularly clear to understand. We followed this lead by analysing the frequency for the proportion of correct responses at each decision point, as shown in Figure 3. While there were too few predictions made at each decision point to statistically validate the presence of patterns, indications are promising. With 10 points consistently predicted and 1 point consistently mis-predicted.

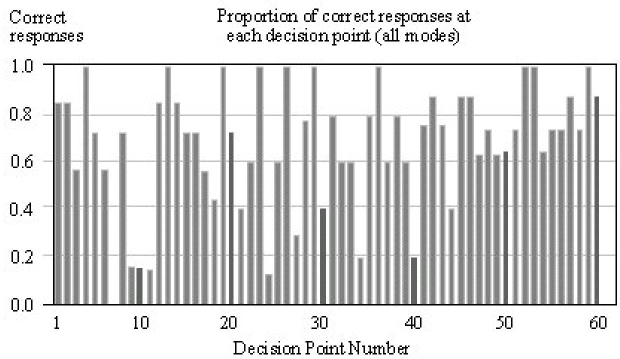

Figure 3. *Proportion Correct (PC) at the 60 decision points.*

The comments from each Subject were recorded after the evaluation in each Mode. In V nine Subjects commented that they made decisions based on 'size, height, slope and steepness of cliffs', 'how close the peaks were to the center', and 'bending and trends in the river or valley'. Three said they could not understand how to make decisions. One commented that it was not clear whether the 'forces shown by hills' were pushing or pulling against each other. Overall three Subjects said they preferred the V display.

In A the Subjects made decisions from 'frequency of calls', 'closeness to the center', and 'loudness'. Six said they found it easy to understand the A display. One commented that the words 'buy' and 'sell' could be interpreted as commands rather than labels which could lead to a prediction in the opposite direction. One commented that information about the last trade price was missing from the A display.

In M Subjects commented that the combined display contained more information than V or A alone, and that V provided 'context, history, past and general trends', while A provided 'most recent trends', 'focus', 'eagerness' and 'presence'. Four mentioned that when V became occluded they used A to make the decision. Four commented that V and A were sometimes in conflict, and 3 of these relied on the A while the other relied on the V in this situation. One found the conflicts in M made it more ambiguous and preferred the A alone. Another found A distracting and preferred V over M.

9. CONCLUSION

The results do not support the null hypothesis. Subjects can predict the direction of the next trade significantly better than chance. There was no difference due to the Order of experience with the displays. The finding that the performance with the sonification and multimodal displays was similar indicates the combined display is redundant for this task. However, the results for prediction of up and down trades show the situation is not so straightforward. Subjects were much better at predicting down trades than up trades. Some Subjects were much better than others at predicting down trades, but there was little difference on up trades. The Subjects were able to predict down trades from the sonification over 80% of the time, and three actually scored 100%. The lower prediction for downward movements using the multimodal display indicates that the visualisation may have introduced noise or conflict when combined with the sonification. Yet in the upward direction the combined multimodal display was the only Mode where performance was significant, indicating that the combination of the sonification and visualisation was complementary for predicting upward movements. This result was supported by comments that the sonification was used for most recent trends while the visualisation was used for context. Sometimes there seemed to be a conflict between the information in the different Modes in which case most users went with the sonification which contained the most recent data. The sonification was also used redundantly when the visualisation was obscured. Subjects were more confident of some decisions than others, and an analysis shows that, at least anecdotally, there are certain decision points that were much more consistently predictable than others.

10. FUTURE WORK

1. Why do Subjects predict down trades so well from the sonification, but not up trades? Perhaps the sonification is biasing toward down decisions or some feature is highlighted by the sonification.
2. Why do some perform better with the sonification than others, yet there is little difference with the visualisation or multimodal displays?
3. Why is the multimodal display complimentary for up trades but redundant for down trades?
4. What are the significant patterns in the visual and sound displays? Indications are some decision points are much more easily and correctly predicted than others.
5. How well do the predictions from these displays compare with a table? How do predictions from experts compare with non-experts? These questions could be explored with an experiment to measure trading performance of experts on real data.

Answers to these questions could provide more general insights and lead to principles for designing more effective sonifications, visualisations and combined multimodal displays.

11. ACKNOWLEDGEMENTS

We would like to thank Bernard Orenstein of Agents Pty. Ltd. for his expertise as a trader in the task analysis and heuristic evaluation. We would like to thank the CSIRO Virtual Environments group for assistance with programming and using the VR system. Finally, thanks to Warren Muller from CSIRO Mathematical and Information Sciences for help with experimental design and statistical analysis.

12. REFERENCES

[1] K. V. Nesbitt, "Interacting with Stock Market Data in a Virtual Environment", Joint Eurographics IEEE TCVG Symposium on Visualisation, Ascona, Switzerland, 2001, SpringerWein New York.

[2] C. Nicholson, "Specialised Techniques in Technical Analysis". Course Notes E114. Sydney, Australia, Securities Institute and Australian Technical Analysts Association, 1999.

[3] W. Jungmeister and D. Turo, "Adapting Treemaps to Stock Portfolio Visualisation". Technical Report Dept. of Computer Science, University of Maryland, CARTR-648, CS-TR-2996, SCR-TR-99-120, 1992.

[4] W. Wright, "Information Animation Applications in the Capital Markets". IEEE Information Visualisation 95, New York. 1995.

[5] S.P. Frysinger, "Applied Research in Auditory Data Representation". In E. Farrell (Ed.). *Extracting meaning from complex data: processing, display, interaction.* Proceedings of the SPIE, Vol 1259, 130-139, 1990.

[6] S. A. Brewster and R. Murray. "Presenting Dynamic Information on Mobile Computers". Personal Technologies, 4(2), pp 209-212. 2000.

[7] J. G. Neuhoff, G. Kramer, and J. Wayand. "Sonification and the Interaction of Perceptual Dimensions: Can the Data Get Lost in the Map?" International Conference on Auditory Display, 2000. Atlanta, Georgia, USA.

[8] O. Ben-Tal, M. Daniels, and J. Berger. "De Natura Sonoris: Sonification of Complex Data". World Scientific and Engineering Society, 2001.

[9] M.R. McGee, P. D. Gray, and S.A. Brewster. "Communicating with Feeling". First Workshop on Haptic Human-Computer Interaction, Glasgow, UK, 1998.

[10] S. Barrass. "Auditory Information Design". Ph.D. Thesis, Australian National University, 1997 http://thesis.anu.edu.au/public/adt-ANU20010702.150218/

[11] H. Tramberend. "Avango: A Distributed Virtual Reality Framework". IEEE Conference on Virtual Reality, Houston, USA, 1999.

[12] Gerhard Eckel, "A Spatial Auditory Display for the CyberStage". International Conference on Auditory Display, Glasgow, UK, 1998.

[13] M. Puckette, "Combining Event and Signal Processing in the Max Graphical Programming Environment". Computer Music Journal 15/3, 1991.

[14] A.J.Dobson. "An Introduction to Generalized Linear Models", Chapman and Hall, London, 1990.