



University of HUDDERSFIELD

University of Huddersfield Repository

Moore, David J. and Wakefield, Jonathan P.

A Design Tool to Produce Optimised Ambisonic Decoders

Original Citation

Moore, David J. and Wakefield, Jonathan P. (2010) A Design Tool to Produce Optimised Ambisonic Decoders. In: 40th International Audio Engineering Society Conference - Spatial Audio, 8-10 October 2010, Tokyo, Japan.

This version is available at <http://eprints.hud.ac.uk/id/eprint/9097/>

The University Repository is a digital collection of the research output of the University, available on Open Access. Copyright and Moral Rights for the items on this site are retained by the individual author and/or other copyright owners. Users may access full items free of charge; copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational or not-for-profit purposes without prior permission or charge, provided:

- The authors, title and full bibliographic details is credited in any copy;
- A hyperlink and/or URL is included for the original metadata page; and
- The content is not changed in any way.

For more information, including our policy and submission procedure, please contact the Repository Team at: E.mailbox@hud.ac.uk.

<http://eprints.hud.ac.uk/>

A DESIGN TOOL TO PRODUCE OPTIMISED AMBISONIC DECODERS

DAVID MOORE¹, JONATHAN WAKEFIELD²

¹ Department of Computing and Creative Technology, Glasgow Caledonian University, Glasgow, UK
j.d.moore@gca.ac.uk

² Department of Computing and Engineering, University of Huddersfield, Huddersfield, UK
j.p.wakefield@hud.ac.uk

This paper describes a tool for designing Ambisonic surround sound decoders. The tool is highly flexible and provides a decoder designer with powerful features to enable the design of a decoder to their specific requirements. The tool employs computer search to find decoder parameters that best meet design criteria specified in a multiobjective fitness function. Features include: objective range-removal and importance, even performance by angle, performance that correlates with human spatial resolution, and frequency dependent and independent decoders of different orders. Performance can be optimised for a single listener or multiple off-centre listeners. The current tool works for 5.0 surround sound however it can be extended to other horizontal-only and 3D configurations. Results are shown that demonstrate the tool's capability and flexibility for various scenarios.

INTRODUCTION

This paper describes a flexible software-based tool for designing Ambisonic surround sound decoders. The decoder design tool (DDT) provides the user with a high-level interface for executing a search for decoder parameters that best fit a fitness function modelled on auditory localisation theory (the interface is shown in Fig. 1). By adjusting the interface controls the user can produce decoders with different performance characteristics.

The DDT consists of a main user interface and two sub-panels (see Fig. 2). The main user interface is the top level of the application where all of the tool's main functionality can be controlled. The performance panel provides detailed information about decoders produced by the search algorithm (see Fig. 4), and the options panel enables the user to configure search properties (see Fig. 5). The following sections summarises each of the DDT's features.

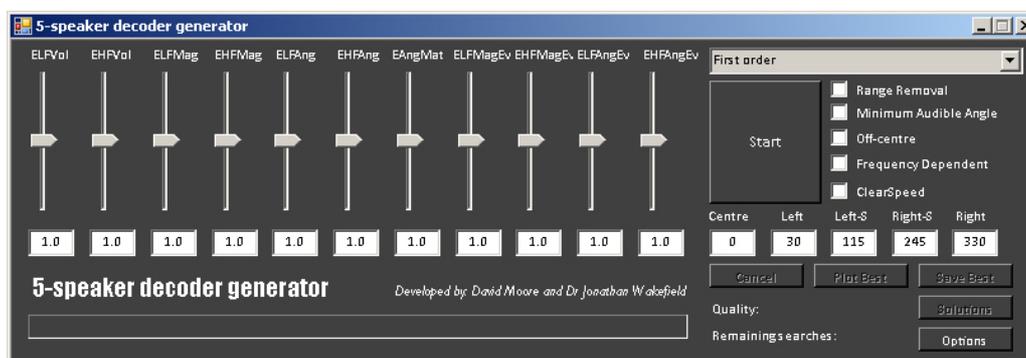


Figure 1: Main user interface of the DDT



Figure 2: DDT structure

1 BACKGROUND

It is well known that the design of Ambisonic decoders for irregular loudspeaker layouts is complicated [1]. A non-linear set of equations needs to be solved in order to produce a suitable set of decoder parameters. In recent years, a number of experiments have been performed using search algorithms as an alternative to solving the

complex decoder equations mathematically [2-5]. In these experiments various design criteria have been used to determine a decoder's localisation performance over the ITU 5-speaker layout. However, these have all been done as ad hoc experiments and it has not been possible to vary or modify design criteria outside of the test environment. Little work has looked at designing a tool for others to use - the only example is the application designed by Wiggins [6]. This current work represents the first attempt to produce a more advanced and flexible decoder design tool.

2 DECODER GENERATION

2.1 Search algorithm

The DDT uses a heuristic search algorithm known as the Tabu Search for finding 'good' decoder parameters according to a fitness function [7]. A heuristic search was employed because searching exhaustively for the 'best' set of decoder parameters using a parameter resolution comparable to previously published work is infeasible using currently available computer processing power.

The Tabu Search is regarded as a good local search algorithm that is capable of producing solutions in a short amount of time. It enhances its performance by using memory structures. One of these memory structures is known as the Tabu list - a list of previous moves which are designated out-of-bounds, or Tabu (hence the name). The Tabu list is used to guide the search away from previously visited areas in the search space preventing the algorithm from getting stuck in a local minimum of the search domain.

2.2 Fitness function

The fitness function used for guiding the search is based upon two models defined in Gerzon's Metatheory of Auditory Localisation: the velocity vector and energy vector [8]. The velocity vector is able to quantify a decoder's localisation performance at low frequencies, whereas the energy vector is able to quantify a decoder's localisation performance at mid/high frequencies. The magnitude of both vectors indicates the sound source image quality, and the angle of both vectors indicates the perceived location of the sound source for a listener.

Eleven objectives are included in the fitness function to measure a decoder's performance. A full mathematical definition of the objectives is given in [9]. In summary they aim to meet the following:

- Equal low frequency volume as a source is panned around the listener (objective ELFVol)

- Equal mid/high frequency volume as a source is panned around the listener (objective EHFVol)
- Velocity vector magnitude is as close to the optimum magnitude as possible (objective ELFMag)
- Energy vector magnitude is as close to the optimum magnitude as possible (objective EHFMag)
- Velocity vector angle is as close to the correct angle as possible (objective ELFAng)
- Energy vector angle is as close to the correct angle as possible (objective EHFAng)
- The velocity vector and energy vector angles are as closely matched as possible (objective EAngMat)
- Equal velocity vector magnitude error around the listener (objective ELFMagEv)
- Equal energy vector magnitude error around the listener (objective EHFMagEv)
- Equal velocity vector angle error around the listener (objective ELFAngEv)
- Equal energy vector angle error around the listener (objective EHFAngEv)

Each of the objectives is checked at a number of angles around the soundstage (specified by the user in the DDT's option panel). In previous work, the angles have been checked from 0 degrees to 180 degrees in 1 degree steps as only half the soundstage needed to be evaluated when the speaker array is left-right symmetrical. The relative influence of these objectives in the fitness function can be adjusted using the relevant sliders (this will be returned to in section 3.1.1).

3 USER INTERFACE

3.1 Main user interface

The main user interface has a number of controls that can be set before starting a search for decoder parameters (each control is labelled in Fig. 3).

The user can enter the order of the required decoder by selecting from the 'decoder order' drop down box (see control 2). This drop down box gives the option of deriving decoders from first order to fourth order. Higher order decoders are able to produce better

localisation performance (according to the velocity and energy vectors) but can take longer to derive because there are a greater number of decoder parameters.

The checkboxes on the right of the DDT (control 3) enable the user to switch a number of important components on and off in different combinations. These components manage the ability to: apply range-removal, set a minimum audible angle weighting scheme, optimise for off-centre listeners, produce a frequency dependent or independent decoder and run multiple searches using High Performance Computing hardware. A summary of each component follows.

3.1.1 Range-removal and Importance

Range-removal is a technique that prevents certain fitness function objectives from biasing the search. Each objective has a different range of potential values which can result in the objective with largest range biasing the search for a decoder. When range-removal is switched on, all objectives are scaled to the same range so that they have an equal impact in the search. The algorithm for range-removal is described in [10].

A companion technique for range-removal is ‘Importance’. This technique simply involves applying a weighting to range-removed objectives. Applying

logical weightings to objectives that have been equalised using range-removal gives a decoder designer a fine level of control over a decoder’s performance characteristics. The user can adjust the relative ‘Importance’ of the fitness function objectives by using the onscreen sliders. By choosing importance values for each objective the user can effectively put the different objectives in rank order of importance. For example, if a user seeks to prioritise good performance for the energy vector, then the objectives EHFMag and EHFAng would receive higher importance weightings when compared to the other objectives.

Please note that although this is intended as a companion technique for range-removal, importance weightings can also be applied to non range-removed objectives if needed.

3.1.2 Minimum Audible Angle Optimisation

When the ‘Minimum Audible Angle’ checkbox is switched on an angle dependent weighting is applied to the velocity vector and energy vector objectives in order to improve performance in areas of the sound stage where humans are more sensitive to sound localisation (i.e. in the front and the rear). This technique is described in [11].

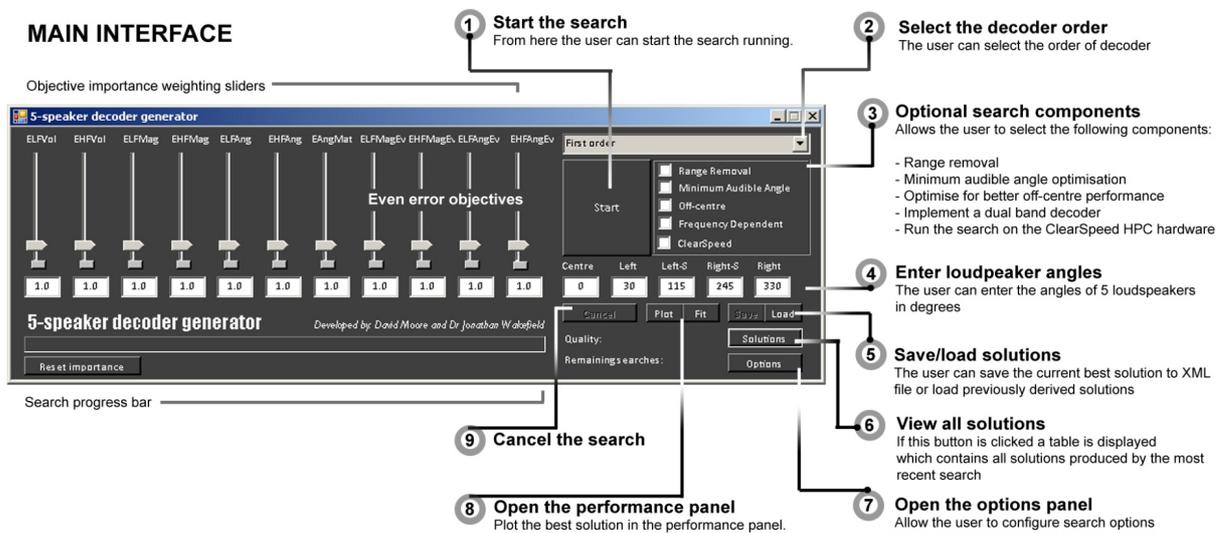


Figure 3: Main user interface with main controls labelled

3.1.3 Off-centre optimisation

The ‘Off-centre’ checkbox gives the user the option of optimising a decoder’s localisation performance in off-centre listening positions. When selected, 9 listening

positions are checked in the fitness function (the centre position and 8 equally spaced off-centre positions).

This function is important when playing audio to a distributed audience (for example and audience in a

cinema or auditorium). The reader is referred to [12] for more information.

3.1.4 Frequency dependent decoders

Toggling the ‘Frequency Dependent’ checkbox allows the user to optimise for a frequency dependent decoder or a frequency independent decoder. The former results in a separate set of optimised parameters for high and low frequencies, whereas the latter is a compromise solution that uses a single set of parameters for high and low frequencies.

3.1.5 High performance computing

Toggling the final checkbox labelled ClearSpeed allows multiple searches to be run in parallel on remotely accessed high performance computing hardware manufactured by ClearSpeed [13]. Incorporating this feature increases the chance of finding a good decoder within a set time because significantly more potential solutions can be evaluated. Alternatively, it enables the tool to derive good decoders more quickly making the tool more interactive for a decoder designer. This is described in [14].

3.1.6 Utility functions

The buttons labelled as control 5 in Fig. 3 allow the user to load or save solutions produced by the search. When loading a solution the user has the option to use it as the

starting point of the search (rather than a random start point). This can be useful improving existing decoders. The button labelled as control 6 allows the user to view a list of all solutions produced by the search from the most recent search run. Finally, the user can input the angles of the speakers using the edit boxes labelled as 4.

3.2 Performance panel

When opening the performance panel from the main user interface, the localisation performance of the best decoder produced by the search is detailed (see Fig.4). There are four plots showing the following information:

Plot 1 (labelled 10 in Fig. 4) shows the velocity vector response around the 360° sound stage. Velocity vector magnitudes are shown at each angle and velocity vector angles are displayed every 30 degrees (starting from 0 degrees at the front of the system). Ideal vector magnitudes and angles are shown in light grey.

Plot 2 (labelled 13) shows the energy vector response around the 360° sound stage. Energy vector magnitudes and angles are displayed with ideal magnitudes and angles in light grey.

Plot 3 (labelled 11) shows the low frequency virtual microphones and pressure around the listener, whereas plot 4 (labelled 14) shows the mid/high frequency virtual microphones and energy around the listener.

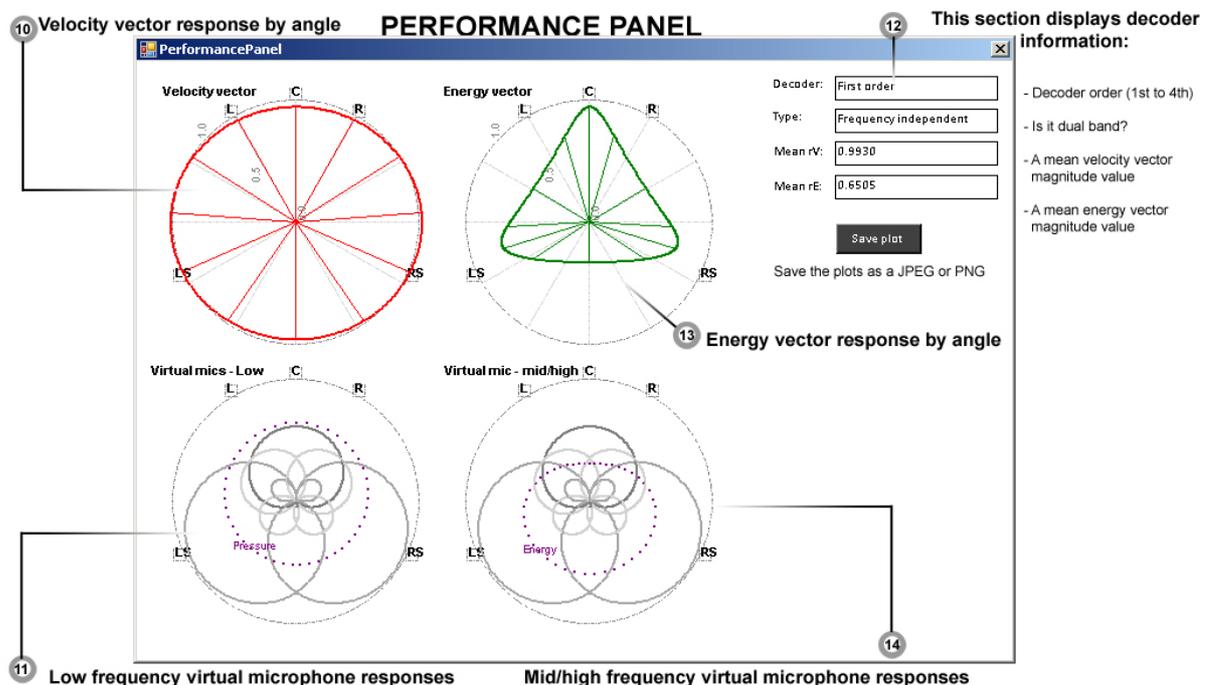


Figure 4: Performance panel with main features labelled. The plots can be saved as high quality image files.

3.3 Options panel

The options panel allows the user to set the main properties of the Tabu Search (see Fig. 5). By using the edit box (label 15) the user can enter the number of bad moves before the Tabu Search stops running. A higher value for this parameter might lead to a better solution being found as the search could potentially reach a better local minimum, however, a higher number of bad moves is likely to have an impact on time-to-solution.

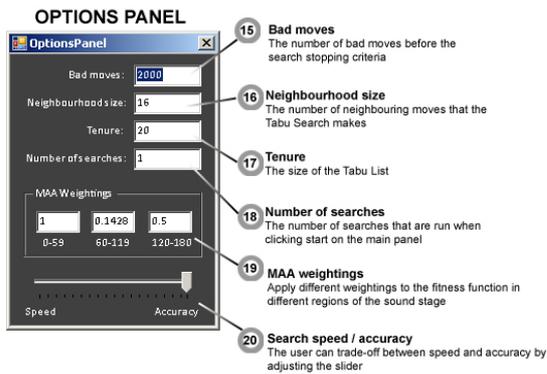


Figure 5: Options panel with main options labelled

The edit box labelled as 16 allows the Tabu Search neighbourhood size to be set. The neighbourhood size is the number of local solutions the Tabu Search generates when searching around the current best solution. The default neighbourhood size is twice the number of parameters so a positive and negative step can be made for each parameter. For example, a first order frequency-independent decoder requires 8 parameters so the default neighbourhood size will be 16.

The edit box labelled as 17 allows the user to set the Tabu Tenure (i.e. the size of the Tabu List). A larger tenure will result in slower search times as the search has to traverse the list for 'Tabu' solutions at each iteration of the algorithm. However, a larger tenure will reduce the chance of the search returning to the same local minimum. On the other hand, a smaller tenure will result in the algorithm running faster but may prevent the search from visiting a wider area of the search space.

The edit box labelled as 18 allows the user to set the total number of sequentially run searches both on the host machine or the ClearSpeed hardware.

The edit box labelled as 19 provides the user with the ability to set their own MAA weightings in the fitness function at the front, sides and rear. Finally, the slider highlighted as 20 allows the user to trade-off between search speed and solution accuracy. If the user chooses

speed over accuracy fewer angles are checked in the fitness function resulting in each solution being evaluated more quickly and vice versa.

4 RESULTS

To demonstrate the capability of the DDT a range of decoders will be presented. The decoders were derived using different settings on the main user interface.

Fig. 6 shows the performance of a typical first order frequency independent decoder optimised for the ITU 5-speaker layout (Decoder 1). When deriving this decoder all checkboxes were turned off, and each objective had an importance value of 1 (apart from the even error objectives which had values of 0). For comparison, Fig. 7 and Fig. 8 show the performance of a typical first order decoder (Decoder 2) and a typical fourth order decoder (Decoder 3) derived using the same settings, but with range-removal switched on.

Note that for Decoder 1 the velocity vector is ideal and the pressure (low frequency volume) is even around the listener. For Decoder 2 the velocity vector performance is reduced, but the energy vector has been improved showing the impact range-removal has had in the search. When comparing Decoder 3 with Decoder 2 it can be seen that Decoder 3 has better vector magnitudes, particularly around the front of the system. This demonstrates the advantage of using higher orders.

Fig. 9 shows the performance of a typical fourth order decoder derived with the even error objectives included in the search. When deriving this decoder a higher importance weighting was given to the energy vector objectives (reflected by its performance). When comparing this decoder with Decoder 3 it is clear that performance is also more even around the listener. The extent of even performance can be controlled using importance. However, it should be noted that there is a direct trade off between even performance by angle and overall performance by angle.

When deriving the decoder shown in Fig. 10 range-removal and the MAA were switched on and the objectives were all given equal importance. It is clear that this decoder has much better performance at the front of the system when compared to the other decoders (i.e. the vector magnitudes are very close to their ideal value of 1 between $\pm 30^\circ$). This demonstrates the affect the MAA component has in the search.

Results showing the impact of the off-centre component are not shown here for brevity. A recent paper describes the performance of off-centre optimised decoders in detail [12]. The reader is also referred to [15] which details the DDT's extensive testing.

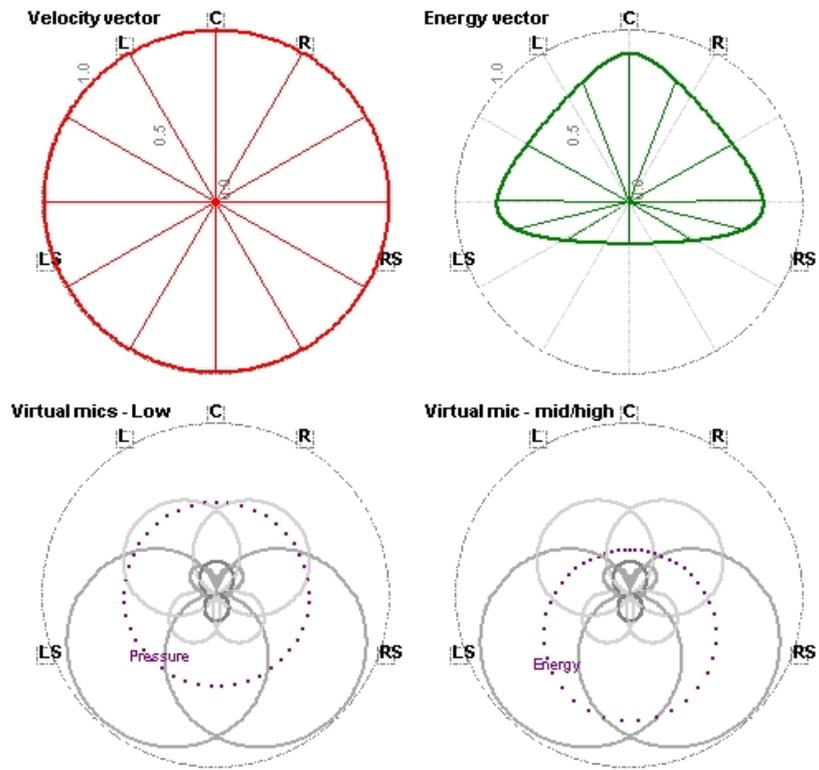


Figure 6: First order decoder derived without range-removal

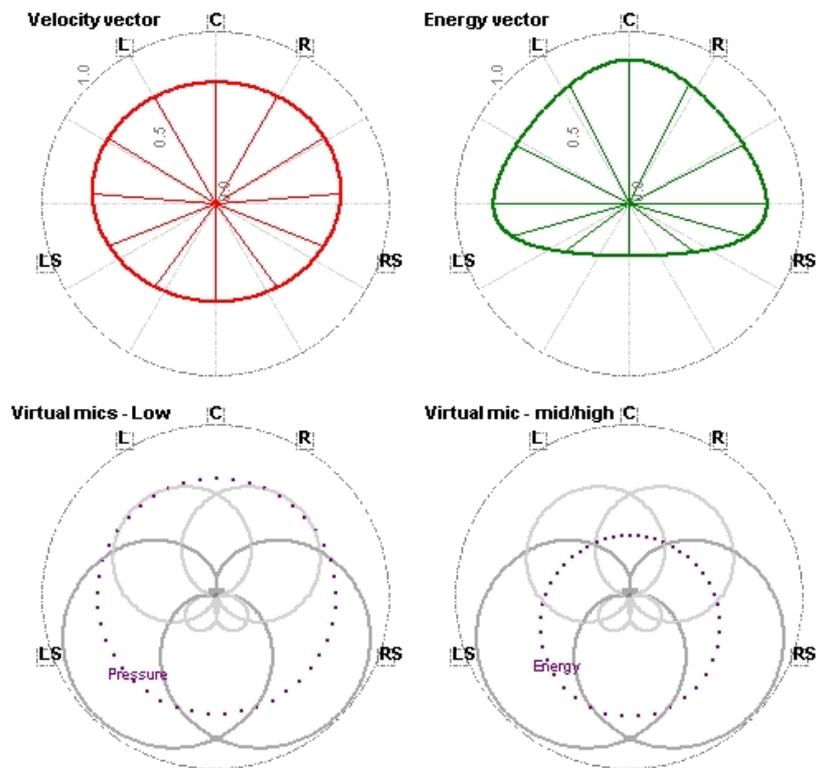


Figure 7: First order decoder derived using range-removal

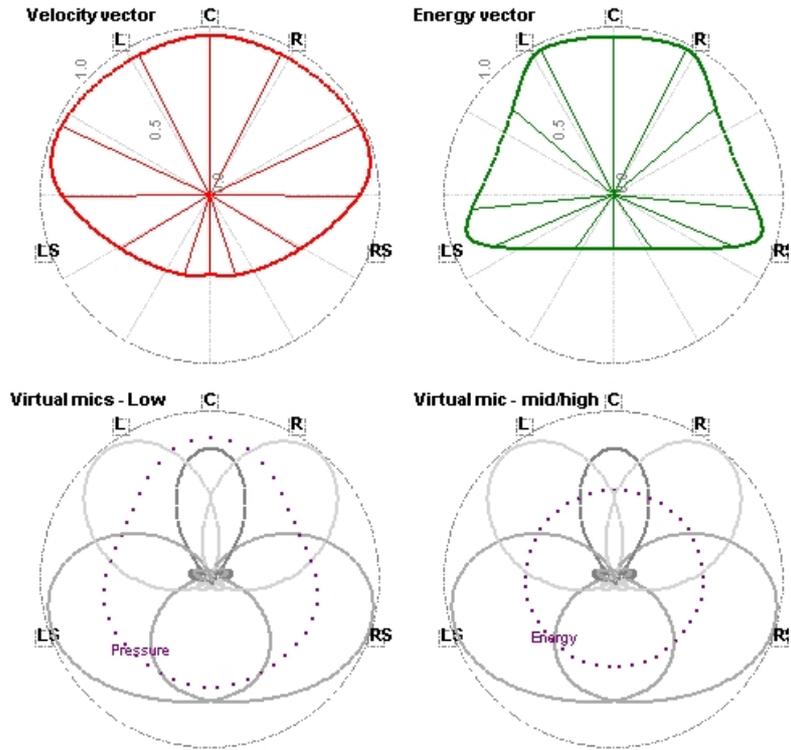


Figure 8: Fourth order decoder derived using range-removal

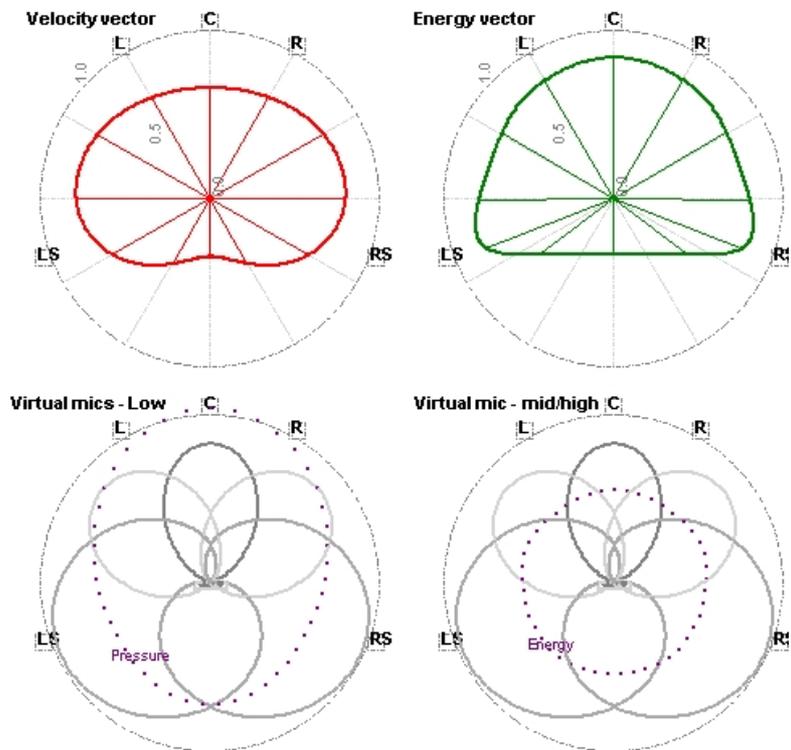


Figure 9: Fourth order decoder optimised for even performance by angle

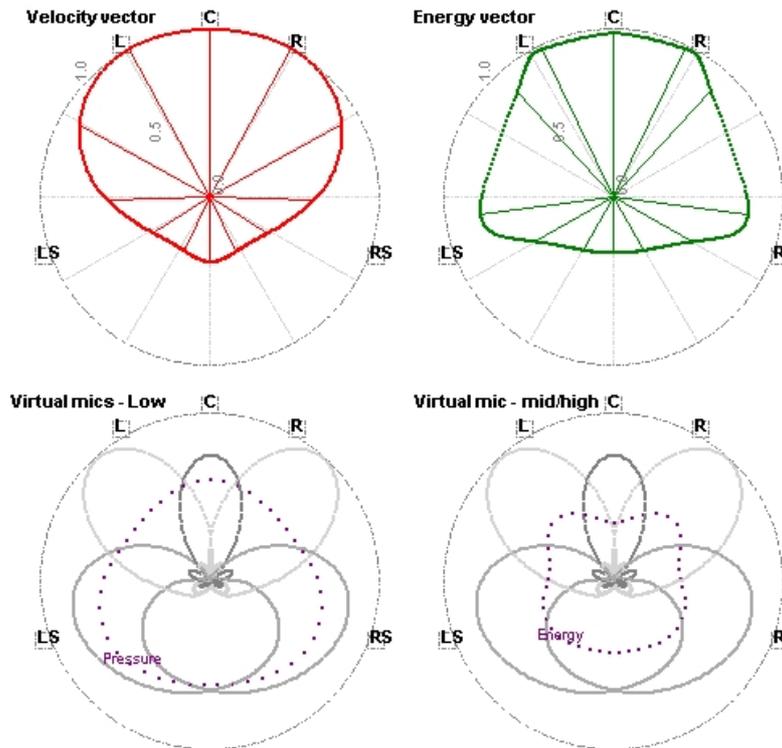


Figure 10: Fourth order decoder derived using the MAA weighting

5 CONCLUSIONS

This paper has presented a design tool for producing Ambisonic decoders for 5-speaker surround sound layouts. The tool has a variety of features that can be used in isolation or combination when fine-tuning the performance of a decoder.

An overview of each of the DDT's features was provided with a selection of results to demonstrate its capability and flexibility when producing decoders for the ITU 5-speaker layout.

6 FUTURE WORK

The DDT currently produces decoders for 5-speaker horizontal-only systems. However, there are plans to extend it to decoders for 3D layouts in the near future in order to take into account height.

REFERENCES

[1] M. A. Gerzon and G. J. Barton, "Ambisonic Decoders for HDTV", Presented at the 92nd Audio Engineering Society Convention, Vienna, Austria (1992).

[2] P. G. Craven, "Continuous Surround Panning for 5-speaker reproduction" Presented at the 24th

International Audio Engineering Society Conference, Banff, Canada (2003).

[3] B. Wiggins, I. Paterson-Stephens, V. Lowndes, and S. Berry, "The Design and Optimisation of Surround Sound Decoders using Heuristic Methods", Proceedings of UKSim 2003 pp. 106-114, (2003).

[4] P. W. Tsang, W. K. Cheung, and C. S. Leung, "Decoding Ambisonic Signals to Irregular Loudspeaker Configuration Based on Neural Networks", *Neural Information Processing*, pp. 273-280, Springer (2009).

[5] D. Moore and J. P. Wakefield, "The Design and Detailed Analysis of First Order Ambisonic Decoders for the ITU Layout", Presented at the 122nd Audio Engineering Society Convention, Vienna, Austria (2007).

[6] B. Wiggins, *An Investigation into the Real-time Manipulation and Control of Three-Dimensional Sound Fields*, PhD. University of Derby (2004).

[7] F. Glover, "Tabu Search – Part 1", *Journal of Computing*, vol. 1, no. 3, pp. 190-206, (1989).

- [8] M. A. Gerzon, “General Metatheory of Auditory Localisation”, Presented at the 92nd Audio Engineering Society Convention, Vienna, Austria (1992).
- [9] D. Moore and J. P. Wakefield, “The Design of Ambisonic Decoders for the ITU 5.1 Layout with Even Performance Characteristics”, Presented at the 124th Audio Engineering Society Convention, Amsterdam, The Netherlands (2008).
- [10] D. Moore and J. P. Wakefield, “The Design of Improved First Order Ambisonic Decoders by the Application of Range-Removal and Importance in a Heuristic Search Algorithm”, Presented at the 31st International Audio Engineering Society Conference, London, UK (2007).
- [11] D. Moore and J. P. Wakefield, “Exploiting Human Spatial Resolution in Surround Sound Decoder Design”, Presented at the 125th Audio Engineering Society Convention, San Francisco, USA (2008).
- [12] D. Moore and J. P. Wakefield, “Optimisation of the Localisation Performance of Irregular Ambisonic Decoders for Multiple Off-Centre Listeners”, Presented at the 129th Audio Engineering Society Convention, London, UK (2010).
- [13] ClearSpeed 2008, ClearSpeed – Home. Available at: <http://www.clearspeed.com/> [Accessed March 2010]
- [14] D. Moore and J. P. Wakefield, “The Potential of High Performance Computing in Audio Engineering”, Presented at the 126th Audio Engineering Society Convention, Munich, Germany (2009).
- [15] D. Moore, *The Development of a Design Tool for 5-speaker Surround Sound Decoders*, PhD, University of Huddersfield (2009).