ICSV14
Cairns • Australia
9-12 July, 2007

14th International Congress on Sound & Vibration

AMBISONIC LOCALISATION - PART 2

Richard Lee[1] and Aaron Heller[2]

[1]Pandit Littoral, Cooktown, Queensland 4895, OZ
ricardo@justnet.au
[2]SRI International, Menlo Park, CA 94025, USA
heller@ai.sri.com

Abstract

An earlier paper [1] examined the performance of "Classic" horizontal Ambisonic surround sound systems with speakers arranged in regular polygons or as diametrically opposite pairs. We extend the work to non-traditional Ambisonic systems; ITU-R BS775 5.0 layouts supposedly used for domestic surround sound and also non-optimal decode, where the speaker array does not correspond to the decoder. Our interest is in simple enhancements possible with present day technology that could benefit the listener at home.

1 INTRODUCTION

We can produce Ambisonic recordings to play on home theatre systems, by decoding to 4 or 5 speakers and re-encoding to "one channel / speaker" formats like Dolby Digital or DTS. Nimbus Records [2] have been very successful with this G-format approach, using "Classic" decode to a square speaker array; probably the most important Ambisonic distribution format at present. Its success depends on the type of surround setup in the average home. ITU-R setups are almost unknown. Living rooms are better suited to simple rectangular or square speaker arrangements; "Real World Systems" in [3] and this is what's found in most homes.

We analyse and listen to three decodes on two speaker arrays, matched and unmatched to the decodes, checking "robustness" and size of sweet spot, to find the best general purpose decoder.

A simple computer simulation gives insights into how these systems behave but the valuable work here is our listening test results. Our calculations and simulations are only because we believe they have some bearing on the subjective experience.

2 DECODES & SPEAKER ARRAYS

Our previous paper [1] describes the experimental method for the listening tests. It allows synchronized playback of the decoded signals and instant switching. Segments can be isolated for repeated playback & detailed comparisons. The main difference is the use of large Revel Studio Monitors in place of small JBL LSR25s. These were positioned with a laser protractor for angular position and ultrasonic rangefinder for distance.

From [1], we have two "Classic" Ambisonic Decoders.

a) **Rect**angular array with speakers at ±30° & ±150°. Aspect ratio 1.732

Our previous tests show this has very good performance for frontal sounds; nearly as good as a regular Hexagon system. Images to the side are less stable than on Square or Hexagonal systems. The decoding equations by Heller [1] are an exact solution of Gerzon's Diametric Opposite Decoder equations.[4] A narrow rectangle is convenient in many rooms.

b) **Square** array decode

Many domestic surround systems have speakers close to a square. Usually there is a centre speaker between the front speakers and the listener is somewhat behind the centre. On these "Real World Systems", if the listener moves forward to the central position, he is in the optimum position for this simple decode and layout. Listening tests on this system are in our previous paper. Nimbus suggest this is the best decode for domestic surround material because of its robust performance with different speaker arrays.[2]

c) **Wiggins** decode & **ITU** speaker array

The new layout is the ITU-R recommendation with speakers at ±30° and ±110°. Speaker signals are by Wiggins [5] who used a Tabu search to find decoder coefficients. We follow Wiggin's [6] suggestion that no Shelf Filters [9] are used for this decode. The numeric methods used to determine decodes for irregular layouts like ITU, often come up with HF & LF solutions that are so different that Shelf Filters are inappropriate for ITU layouts. Instead, these numeric methods optimise on a number of weighted criteria, including rE & rV.[7][11]

Gerzon & Barton [7] claim that for irregular layouts (like ITU), there is no analytic solution to maximising the Ambisonic rE & rV factors in all directions, having both Energy & Velocity models indicate the same directions, and a flat frequency response.

No centre front, CF speaker is used. Adriaensen [8], Craven [10] and Wiggins, report that their independent numeric methods for ITU layouts tend to turn off the CF speaker. Often the CF speaker is very different from the other speakers and draws attention to itself.

True ITU layouts are rare. Is the complicated decode for this rare layout worthwhile, compared to simpler square decode for general distribution of surround material?

We use a '**Decoder on Speaker Array**' naming system. Square on ITU is Square decoder with ITU speaker array. '**System**' refers to a decoder with its matched speaker array.

2.1 Decoders and gain

Tables 2.2 - 2.4 are the speaker decoding equations[1]. The first two use Shelf Filters [1][9] to maximise the Ambisonic rE & rV factors at different frequencies making them Classic Ambisonic Decoders. Linear phase 2048 point FIRs achieve the essential phase matching.

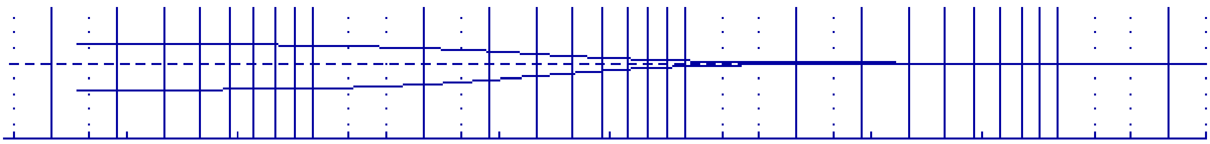


Fig 2.1 Shelf Filters with 0dB gain at HF for Rationalised Square & Rectangle Decoders [9]

Table 2.4 is the VST version of Wiggins' WAD, substantially the coefficients in [5] minus 2.11dB. Each system has an associated gain to equalise subjective loudness differences so these do not influence the results. This is done by normalising the sum of the "speaker signals squared" and validating with a listening test.

But decoders do not always have the same "gain" in all directions. A decoder can make pressure the same in all directions, by equalising the sum of the speaker signals, or it can make "energy" correct, by equalising the sum of the "speaker signals squared". The top diagrams in Fig 3.1 show pressure & "energy" gain with direction for our three Ambisonic systems, normalised to 1.0 maximum.

Subjective loudness is closely related to "energy". Only regular polygon decoders (eg Square) can have both pressure & "energy" gain equal in all directions. The "Vienna" decoders match pressure and use "dominance" to get equal loudness with direction.[7] The Wiggins ITU decoders match "energy". Sounds from the sides will appear louder on the Rectangle decoder. The Square & ITU decoders present sound from all directions with equal loudness. We choose to match frontal loudness with our mainly frontal material. Different material may require different criteria to be matched for equal loudness.

[1] There is much confusion over the exact value of decoder coefficients. For a Classic Decoder, once the speaker position factors are correct, there is only one other factor; the ratio of W vs XYZ. Anything else is simply a volume control. [9] shows only relevant factors with no extraneous scaling.

Table 2.2 Rationalised Energy **Square** Decoder. Gain 10.1dB

	W	X	Y
LF	0.35355	0.25	0. 25
RF	0.35355	0. 25	-.25
LB	0.35355	-.25	0. 25
RB	0.35355	-.25	-.25

Table 2.3 Rationalised Energy 1.732:1 **Rectangle** Decoder. Gain 10.1dB

	W	X	Y
LF	0.35355	0.20412	0.35355
RF	0.35355	0.20412	-0.35355
LB	0.35355	-.20412	0.35355
RB	0.35355	-.20412	-0.35355

Table 2.4 Decoder for **ITU**-R speaker layout due to **Wiggins**. Gain 0dB

	W	X	Y
L	0.3928	0.4868	0.3450
R	0.3928	0.4868	-.3450
SL	0.6570	-.2890	0.4507
SR	0.6570	-.2890	-.4507

3 "OBJECTIVE" MEASURES

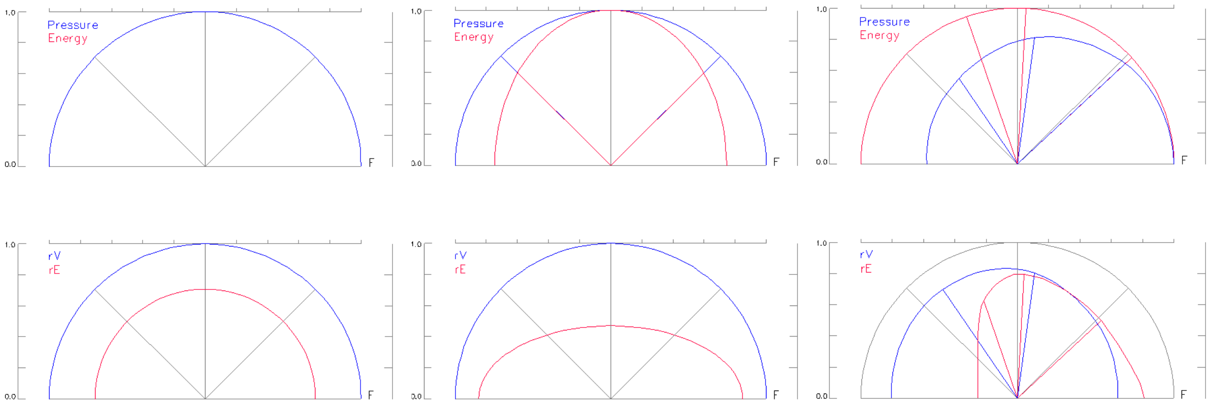


Fig 3.1 Pressure & Energy Gain, top, rV & rE, bottom, with direction for Square, Rectangle & ITU

The bottom diagrams show Gerzon's rV & rE factors [1][11]; measures of localisation performance according to his Velocity & Energy models of auditory localisation. Radial lines show the encoded (grey) and decoded (blue & red) directions for these models which *describe* (*give the same answers as*) all known theories of auditory localisation except for pinnae and high frequency ITD models.[11] Front is 'F' on the right side of the diagrams.

The "Classic" Square & Rectangle systems achieve optimum rV = 1.0 for all directions at low frequencies. The Rectangle has rE = 0.85 for front & back and 0.47 at the sides. It will perform better than the square for front & back sounds and poorer at the sides. The Wiggins ITU system works well in front and to the sides but is poor to the back. Its decoded directions are skewed to the front. Smaller rV values suggest frontal sounds might be more diffuse compared to the "Classic" systems.

These traditional measures of localisation performance imply the Square system is isotropic. But a Square system is corner sensitive. It's not quite a detent effect and Ambisonic systems have always sounded better than multiple pairwise panning. One of the listeners claims the Square has less *detent effect* (*speakers drawing attention to themselves*) than a 16-channel horizontal ring with pairwise panning. But localisation is definitely better in the direction of a speaker.

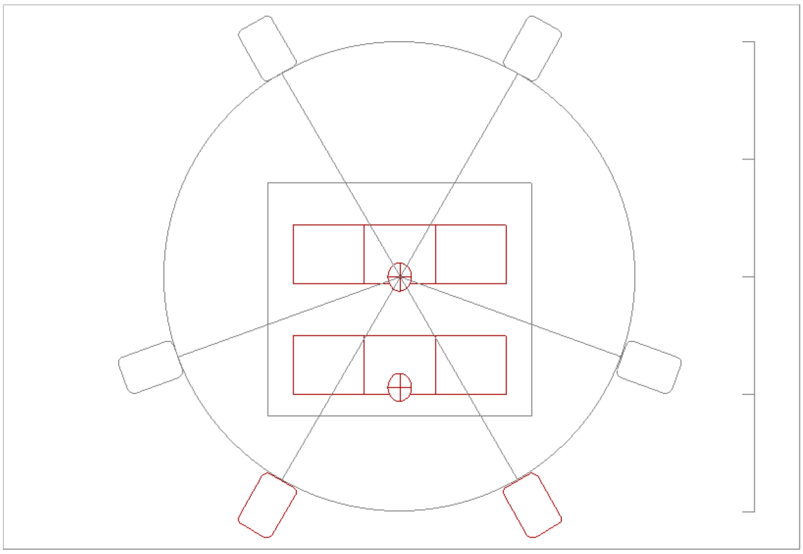


Fig 3.2 Two sofa positions in our 4m circle of speakers. Second row 0.941m behind. The animated pressure plots analyse the 2.24 x 1.98 m rectangle around the listening positions. Four speakers are an **ITU**_R array in the 4.64 x 6.75 x 2.64m room. Two extra speakers (red) form the 1.732:1 **Rect**angle array with the 2 front speakers to allow instant switching.

3.1 Animated pressure plots

Figs 3.3 - 3.8 are instantaneous pressure plots for the area around the two sofas in our circle of speakers at 400Hz. The grey scale is absolute pressure with white as 'zero'. A line from the central seat points towards the coded direction. 400Hz is the centre frequency for Shelf Filters in "Classic" decoders. Preliminary measurements suggest assuming point sources [12] is valid even for our large speakers in a room.[13]

True reproduction would show a plane wave from the coded direction. Alternatively, Fig 3.7 shows a real source at a finite distance. Instead we see an interference pattern characteristic of multiple correlated sources. For direct sounds, the pressure gradient, orthogonal to the pressure contours, aligns with the correct direction. When these plots are animated, some of these contour features are stationary while others move.

If the pressure gradient aligns with the direction of movement over a large area, this is similar to real life; ie good localisation. If the gradient rotates in the course of a cycle, we expect confusing localisation or "phasiness"[11]. (*or a standing wave in real life)*. A gradient changing direction within a small area suggests critical sweet spot.

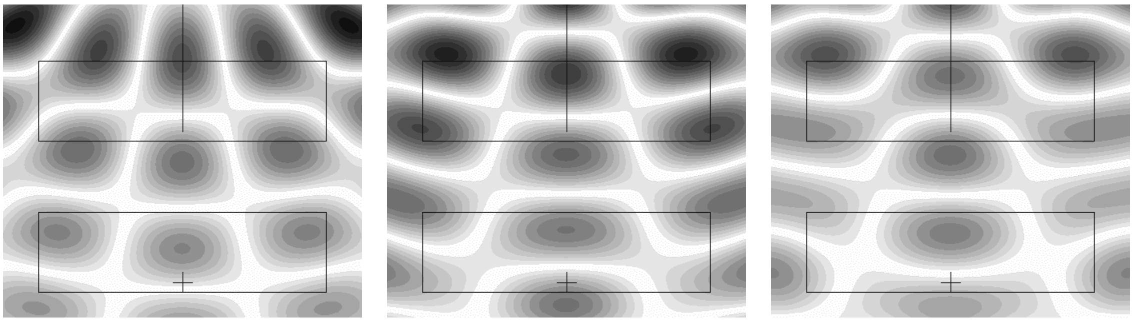
Fig 3.3a & 3.6a are sources at 0° & 45° for a Square. For 0°, two lines of stationary zero pressure narrowly enclose the area of good localisation. At 45°, the wavefront is much clearer and stationary lines are much further away.

In Fig 3.5b the Rectangle for a 90° source, though the wavefront is clearly coming from the left, the width of the area between the zero pressure lines indicate a listener facing left might have ears where the gradient is pointing at right angles to the encoded direction.

Fig 3.4 shows a case of very poor localisation with critical sweet spot; especially for the back row with wavefronts coming from both sides. Figs 3.5c & 3.8c show somewhat more skew in decoded directions on ITU than predicted by Fig 3.1, the rV & rE plots.

Pressure maps for speaker arrays not matched to the decodes show the speaker arrays mostly retain their character for each direction. The biggest change is for the ITU array fed by Square or Rectangle decodes of a source at 90°.

The animated pressure maps only show one frequency at a time. Things become more complicated at high frequencies while the approximation to a plane wave or single source is better at low frequencies. While difficult to interpret, they explain some features of Ambisonic localisation not shown by the traditional rV & rE factors.



a) Square b) Rect c) ITU

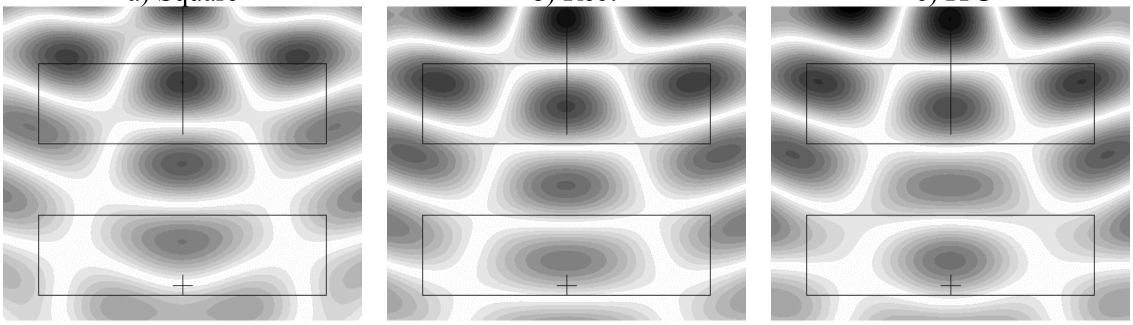


Fig 3.3 **0°** d) Square on Rect e) Square on ITU f) Rect on ITU

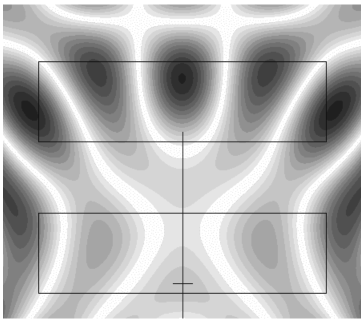


Fig 3.4 **180° ITU**

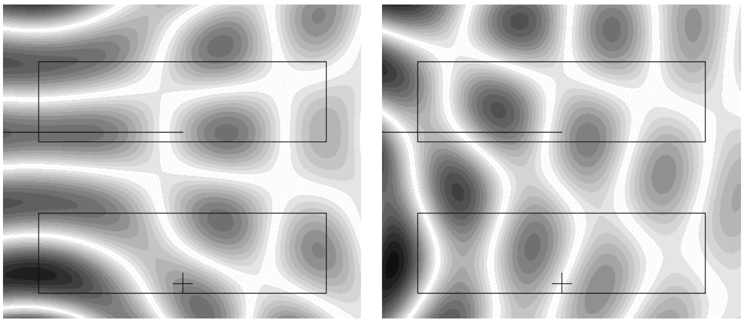


Fig 3.5 **90°** b) Rect c) ITU

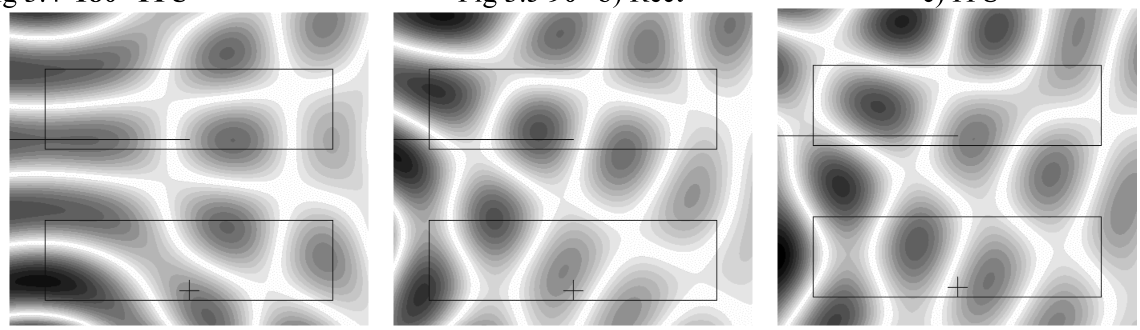
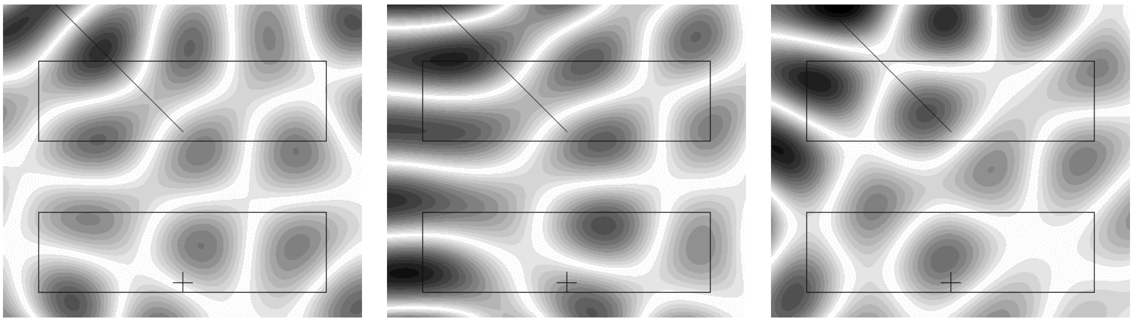


Fig 3.5 **90°** d) Square on Rect e) Square on ITU f) Rect on ITU

This may be the first attempt to show "large area" behaviour for simple 1st order systems. Ideally, a small set of numbers would sum up "large area" behaviour; perhaps based on Bamford's D-error.[15] Ward & Abhayapala [14] use a similar but static method to look at High Order Ambisonics but their elegant Truncation Error is not much use for crude 1st order systems. These "wavefront" measures ignore psychoacoustics.

Some caveats. Millward [18] shows early reflections are important for localisation. Simulations of stereo localisation are very far from real life unless some early reflections are included. Our assumption of point sources needs further checking and it may be that for large speakers, we should assume line sources. Lastly, we assume that the listener interacts with a soundfield created by several sources in the same manner as for a single source. An alternative approach using HRTFs, eg Wiggins [6], would avoid this possible inaccuracy.



a) Square b) Rect c) ITU

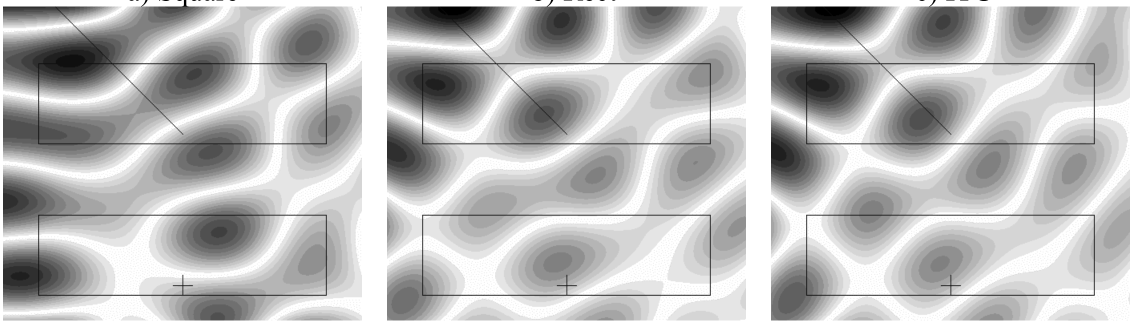
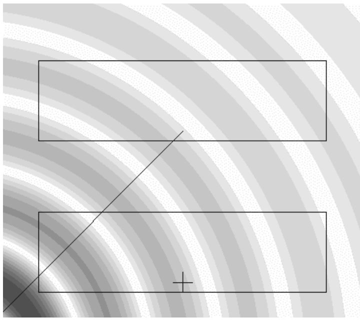


Fig 3.6 **45°** d) Square on Rect e) Square on ITU f) Rect on ITU



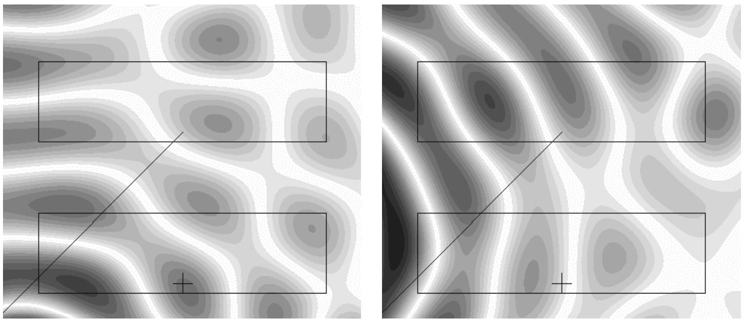


Fig 3.7 **Real source** @ 2m Fig 3.8 b) Rect c) ITU

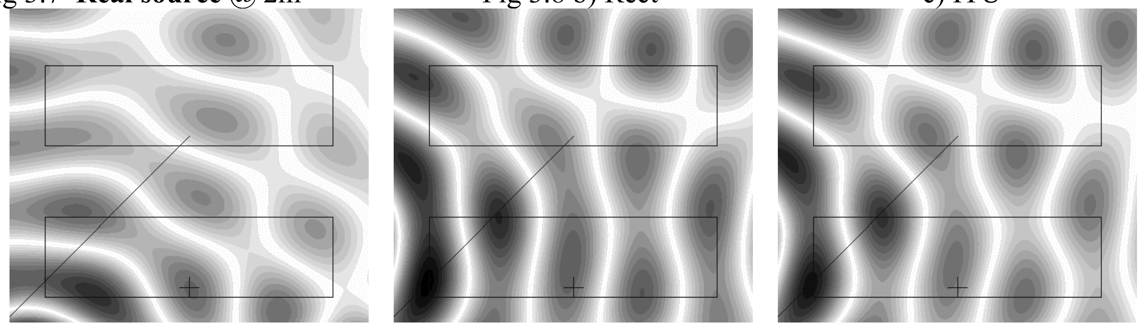


Fig 3.8 **135°** d) Square on Rect e) Square on ITU f) Rect on ITU

4 LISTENING TEST MATERIAL

- Ambisonically panned alto voice announcing the eight cardinal & diagonal directions
- bandlimited noise Ambisonically panned in a circle
- Pulcinella : chamber orchestra in 1200 seat hall with very good acoustics. Good applause
- Aran music : open air recording of folk trio in the harbour area of Inishmore, Ireland.
- 2nd movement from Britten's Simple Symphony : chamber orchestra in church

Pulcinella is a recording made by one of the authors with a carefully aligned Mk4 Soundfield microphone. Aran music and the Britten were recorded with ST250 Soundfields. These three recordings are from a public archive of B-format Ambisonic material.[16]

5 LISTENING TEST RESULTS

Primary source for these comments was the panned voice. The music recordings were used to corroborate.

Front & rear are positioned correctly and compactly in all combinations. Rear sounds were closer and duller over the Rectangle. This is not affected by front/back movement. With left/right movement, ITU front remains centered, whereas back collapses to the nearer speaker. Rectangle front & back move a bit but remain centered with movement.

The edges of the front sound stage, ± 45°, range from about 10° inside the speakers, to 20° outside. The Rectangle system correctly reproduces the edges of the frontal soundstage at ± 45° as expected. The ITU system and Square on Rectangle, both show a narrow ± 20° soundstage. ITU was more spread out at the center and compressed at the limits; sound especially transients are drawn to the speakers. With Aran Music and to a lesser extent the Britten, there was a tendency for some frontal sounds to come from the rear. Square on ITU yields a soundstage of ± 30°. Rectangle on ITU produced a very wide soundstage, but was vague and phasey [11] especially with small head movements. All except the last provide acceptable sound.

An obvious tonal difference between the ITU system and Square or Rectangle decode on the Rectangle array, is apparent only when there are significant sources to the rear. In Pulcinella, the brass reverberates from the back and applause comes from all around. These differences cause the sound *during these passages* to have emphasized midrange on the Rectangle array and less midrange on ITU. Other passages are not affected in this way. The tonal differences may be explained by different rear speaker HRTFs for the two arrays.[17]

For left & right, ± 90°, the Rectangle system reproduced sources in the correct position, but slightly diffuse. Previous tests show this is the most difficult direction for 4-speaker arrays to get right. Six speaker arrays do a much better job. Rectangle on ITU was extremely diffuse. On the other combinations, the sound was drawn to the rear speaker.

Back left & right, ± 135°, were drawn to the nearest speaker on all combinations.

5.1 Moving back to the second row ...

Front localization stayed the same for all decodes. The ITU array was unstable for rear sources; phasey artifacts with small head movements on both the Square and its matching decoder; and to an unacceptable degree with the Rectangle decoder. A listener is unlikely to move so far back on an ITU array.

For sounds at the edge of the front soundstage, the Square decoder on both ITU and Rectangle arrays remained stable and compact. The Rectangle system was wider but became diffuse. The ITU system changed from being the narrowest to the widest and became very diffuse. At this position, the listener is almost inline with the rear ITU speakers.

Sources from ± 90° as well as ± 135° appeared from direct left & right, ± 90° with slight differences among the different combinations.

The Square decode was the most "robust" when moving back. The ITU array appeared generally more critical of listening position then the Rectangle.

6 DISCUSSION

We stress the value of this work is in the listening test results. The theory and simulations are only useful to the extent that they predict these results.

The animated pressure plots explain (with judicious hindsight) many, but not all, of these results. eg. no amount of hindsight allows us to say the pressure plots predict the poor subjective performance of the Rectangle decode on the ITU array !

However, the animated plots show sufficient promise to be worth developing further to include the effects of Distance Compensation and the points at the end of section 3.

An important system not tested was a Square speaker array. This is a common domestic arrangement with the listener back from centre; the "Real World System" in [3]. The matching decoder is complex, requiring compensation for exact speaker distances.

Radial speaker distance is more critical than angular direction. We measured time of arrival with an impulse to each speaker and a microphone at the centre to confirm distance within 2cm. Today, even inexpensive AV receivers [19] have facilities to do this type of alignment. Radial displacement of a speaker by 8 cm. noticeably degraded *spatial and tonal* aspects of reproduction. Moving back by 100cm, still gave good reproduction, suggesting lateral symmetry is more important than front/back.

We are disappointed to find our previous paper is the first published record of listening tests on "Classic" 1st order Ambisonic systems. These may be the first formal listening tests on Ambisonic decoders since the original Ambisonic team in the late 70s. Only Wiggins has conducted and reported similar tests (on ITU systems). We urge those working on decoders to conduct and report listening tests in comparison with "Classic" systems, which in our opinion, still set the standard for Ambisonic surround.

7 CONCLUSIONS

We re-affirm our support for the "Classic" Ambisonic Rectangle Decoder and Speaker Array especially for recordings with mainly frontal sound. While impressed by the Wiggins decoder, we find the ITU layout sub-optimal, especially for back and side sources, even with its matched decoder. It is critical of listener position and domestically unfriendly.

We cautiously endorse the Nimbus recommendation [2] to use "Classic" Ambisonic Square decode with Shelf Filters to produce recordings for general home theatre release. This gives good results with a wide range of domestic speaker arrays.

FURTHER WORK & DEDICATION

More simulations and listening test results will be in an expanded preprint on www.ai.sri.com/ajh/ambisonics. We will develop the animated simulations further and do more listening tests on "Real World Systems", Distance Compensation, different Shelf filters and lossy compression.

This work is dedicated to Eric who inspired much of it including the animated plots.

REFERENCES

[1] Benjamin, Lee & Heller; *Localization in Horizontal-Only Ambisonic Systems*, 121st AES convention San Francisco oct06 expanded preprint at www.ai.sri.com/ajh/ambisonics

[2] *Nimbus DVD Music* www.wyastone.co.uk/nrl/dvd_intro.html

[3] Lee, R; *Ambisonic Surround Decoder* www.ambisonia.com/Members/ricardo

[4] Gerzon, M; *Practical Periphony* preprint 1571, 65th AES convention, London feb80 AES-EL3794.PDF.

[5] ITU-R speaker decoder gains by Wiggins, B www.e-wig.co.uk/sursound/gformat_ac3.jpg

[6] Wiggins, B; *An Investigation into the Real-Time Manipulation and Control of Three Dimensional Sound Fields* PhD. Derby University 2004 sparg.derby.ac.uk/SPARG/PDFs/BWPhDThesis.pdf

[7] Gerzon & Barton; *Ambisonic Decoders for HDTV* preprint 3345 92nd AES convention, Vienna mar92

[8] Adriaensen, F. Private communication on the design of Ambisonic decoders for ITU layouts

[9] Lee, R; *Shelf Filters for Ambisonic Decoders* www.ambisonia.com/Members/ricardo

[10] Craven, P; *Continuous Surround Panning,* 24th AES conference on Multichannel Audio, Banff 2003

[11] Gerzon, M; *General Metatheory of Auditory Localization*, preprint 3306, 92nd AES convention, mar92

[12] Beranek, L; *Acoustic Measurements* John Wiley, New York (1949) pp 56 - 64

[13] Meyer Sound Labs Inc; *Line Arrays: Theory, Fact & Myth*, 2002

[14] Ward & Abhayapala; *Reproduction of a Plane Wave Soundfield Using an Array of Loudspeakers,* IEEE Trans. on Speech & Audio Processing **9** no. 6 sep01

[15] Bamford, J; *An Analysis of Ambisonic Sound Systems of 1st and 2nd order*, MSc. Waterloo University, Ontario 1995

[16] Public domain library of Ambisonic B format material www.ambisonia.com

[17] Benjamin & Brown; *The effect of head diffraction on stereo localization in the mid frequency range* 122nd AES convention, Vienna may07

[18] Millward, G; *Loudspeakers and the Stereo Image*, Hi-Fi News **29**, nov84

[19] Multi Channel Acoustic Calibration System : Pioneer VSX516 www.pioneer.com.au/home_entertainment