# A simple technical measure for the perceived source width

Matthias Frank, Georgios Marentakis, Alois Sontacchi
*Institut für Elektronische Musik und Akustik , Email: frank@iem.at*
*Universität für Musik und darstellende Kunst Graz, Austria*

## Introduction

The inverse relation between the Auditory Source Width (ASW) and the Inter-aural Cross Correlation (IACC) has been used to explain the broadening of instrument sounds encountered when listening in concert halls. Reflections from walls and the ceiling that arrive later than the direct sound decrease the IACC and therefore increase the ASW in proportion to their laterality [1, 2, 3, 4]. Although spatialization systems exhibit big differences in the directions loudspeaker signals emerge from, a fact that may influence the ASW, this possibility has not been evaluated much yet. Measurements presented here are a part of a bigger study on the perception of the ASW in different amplitude panning methods. The loudspeaker spacing and the location of the source within the speaker array are varied.

In this work, we investigate whether the magnitude of the so-called energy vector $\mathbf{r}_\mathrm{E}$ (see Equation 1) can be used to predict the ASW. The energy vector $\mathbf{r}_\mathrm{E}$ is computed analytically from the gains $g_l$ and the position vectors $\boldsymbol{\theta_l} = \{\cos(\phi_l), \sin(\phi_l)\}^\mathrm{T}$ of L loudspeakers:

$$\mathbf{r}_\mathrm{E} = \frac{\sum_{l=1}^{\mathrm{L}} g_l^2 \boldsymbol{\theta_l}}{\sum_{l=1}^{\mathrm{L}} g_l^2}, \qquad \mathbf{r}_\mathrm{E}^\mathrm{w} = \frac{\sum_{l=1}^{\mathrm{L}} (g_l w(\boldsymbol{\theta_l}))^2 \boldsymbol{\theta_l}}{\sum_{l=1}^{\mathrm{L}} (g_l w(\boldsymbol{\theta_l}))^2}. \quad (1)$$

Its direction and magnitude relate to the direction and spread of the acoustic energy. A magnitude value of 1 would indicate all acoustic energy comes from a single direction, while one of 0 would correspond to energy distributed in all directions. It is found that a psycho-acoustically informed direction dependent weight $w(\boldsymbol{\theta_l})$ of the gains $g_l$ needs to be considered (as in $\mathbf{r}_\mathrm{E}^\mathrm{w}$). Below, $\mathrm{r_E}$ and $\mathrm{r_E^w}$ denote the lengths $|\mathbf{r}_\mathrm{E}|$ and $|\mathbf{r}_\mathrm{E}^\mathrm{w}|$.

We present correlation results between subjective ASW ratings and different objective measures: $\mathrm{r_E}$, $\mathrm{r_E^w}$ (with different weightings) and $\mathrm{IACC_{E3}}$ obtained from binaural dummy head recordings (averaged over the first 80 ms at 500Hz, 1KHz and 2KHz octave bands [2]).

## Method

ASW pairwise comparisons were done for all stimuli combinations from Table 1. Stimuli was 1.5s of pink noise presented at 65dB(A) in front of the listeners who were seated facing forward. Genelec 8020 loudspeakers were equidistantly placed at ear height on a circle with 2.5m radius in the IEM CUBE (11m × 11m × 5m, $\mathrm{RT_{60}}$ = 470 ms, effective room critical distance 2.76m).

After a short training phase, all comparisons were repeated four times. Subjects responded which of the two sounds in the pair (A or B) sounded wider by pressing a button on a keyboard. They could only listen to each sound in the pair once. The experiment and stimuli were generated in real-time using pure-data [5]. 16 subjects (10 male, 6 female) aged from 21 to 38 (median = 28) years participated in this test. All were members of a trained listening panel [6, 7].

The tested conditions correspond to realistic scenarios occurring in spatialization systems using amplitude panning. The following systems with two loudspeaker spacings ($\Delta\phi_l = 45°$ corresponding to 8 and $\Delta\phi_l = 22.5°$ corresponding to 16 equally spaced speakers) have been tested:

1. Vector Base Amplitude Panning (VBAP) where a phantom source was created within a speaker pair by manipulation of the loudspeaker gains [8] (1 or 2 active speakers),
2. Multi Directional Amplitude Panning (MDAP) where 10 virtual sources within a spread equal to the loudspeaker spacing around the desired source direction were defined [8] (2 or 3 active speakers),
3. 3rd and 7th order Ambisonics systems implemented on 8 and 16 speaker arrays respectively, with two different order weightings applied to each: max $\mathrm{r_E}$ and basic [9] (2 or 8/16 active speakers).

**Table 1:** Description of test conditions: used system, loudspeaker spacing and source position relative to the loudspeakers. *Note on conditions 1, 5 and 7: Each of these conditions represents several systems, in which the loudspeaker spacing and source position yield the same loudspeaker signals.

| Cond. | System(s) | $\boldsymbol{\Delta\phi_l}$ | position |
|---|---|---|---|
| 1 * | real source,VBAP | - | on LS |
| 2 | MDAP | 22.5° | on LS |
| 3 | $\mathrm{Ambi^{max\,r_E}}$ | 22.5° | on LS |
| 4 | $\mathrm{Ambi^{basic}}$ | 22.5° | on LS |
| 5 * | VBAP,MDAP,$\mathrm{Ambi^{max\,r_E}}$ | 22.5° | half-way |
| 6 | $\mathrm{Ambi^{basic}}$ | 45° | on LS |
| 7 * | VBAP,MDAP,$\mathrm{Ambi^{max\,r_E}}$ | 45° | half-way |
| 8 | MDAP | 45° | on LS |
| 9 | $\mathrm{Ambi^{max\,r_E}}$ | 45° | on LS |

## Results

Within-subject repetitions were averaged and the pairwise comparison matrices were transformed into scale values, yielding 16 observations for each condition (see Figure 1). Outliers were removed (one for conditions 1, 4 and 9) with Grubbs test ($\alpha = 0.05$). There was a significant main effect of speaker spacing ($F(2, 45) = 155.35, p < 0.001$). The medians of the subjective data were regressed against: the $\mathrm{IACC_{E3}}$ (measured using B&K 4128C), the $\mathrm{r_E}$ and the $\mathrm{r_E^w}$ with 3 different weightings. Results are presented in Table 2.

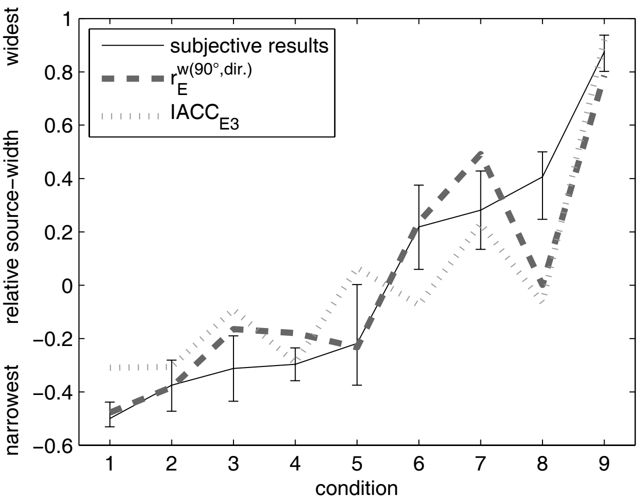


**Figure 1:** medians and corresponding 95% confidence intervals of subjective results and regression with objective measures.

**Table 2:** $R^2$ from linear regression of objective measures against medians of subjective results and correlation to $IACC_{E3}$

| **Objective Measure** | $\mathbf{R^2}$ medians | **Correlation with $\mathbf{IACC_{E3}}$** |
|---|---|---|
| $IACC_{E3}$ | 0.71 | 1 |
| $r_E$ | 0.30 | 0.35 |
| $r_E^{w(90°)}$ | 0.65 | 0.78 |
| $r_E^{w(dir.)}$ | 0.77 | 0.79 |
| $r_E^{w(90°,dir.)}$ | 0.84 | 0.88 |

## Discussion

The results indicate that variations in speaker spacing influence the perception of ASW, even when correlated signals are emitted simultaneously. This further extends similar findings for delayed signals used to simulate reflections [3].

It appears that the unweighted $r_E$ cannot sufficiently model the ASW. Improvement emerges, when the result of Morimoto in [10] is used. That states that sound from the rear does not contribute to ASW. This is done by windowing the loudspeaker signals (1 for all loudspeaker positions within $|\phi_l| \leq 90°$, linearly fading out for $90° \leq |\phi_l| \leq 135°$) as in $r_E^{w(90°)}$.

Incorporating the ear directivity as in $r_E^{w(dir.)}$, largely improves correlation, too (0dB in the front, +3dB at about ± 50°, 0dB on the sides, -6dB for rear positions [11]).

Combing both ($r_E^{w(90°,dir.)}$) leads to further improvement. The improvements in the regression against the subjective data accompany the increasing correlations with the $IACC_{E3}$.

## Conclusion

We presented an analytic measure for predicting the perceived ASW in amplitude panning methods. It seems that the weighting $w(\boldsymbol{\theta_l})$ of $r_E^w$ plays an important role and calls for further analysis. We are currently investigating ways to express the IACC in terms of speaker gains and positions for this purpose. Further investigations will also incorporate the effect of different head orientations.

## Acknowledgments

This project is funded by the Austrian Research Promotion Agency (FFG), the Styrian Government and the Styrian Business Promotion Agency (SFG) under the COMET programme. Thanks to all subjects for their participation and to the project partners for assistance.

## References

[1] M. Schroeder, D. Gottlob and K. Siebrasse: Comparative study of European concert halls: correlation of subjective preference with geometric and acoustic parameters. Journal of the Acoustical Society of America, 56, 4, pp. 1195-1201, 1974

[2] T. Hidaka, L. L. Beranek, and T. Okano: Interaural cross-correlation, lateral fraction, and low- and high-frequency sound levels as measures of acoustical quality in concert halls. Journal of the Acoustical Society of America, 98, 2, pp. 988-1007, 1995

[3] M. Morimoto and K. Iida: Appropriate frequency bandwidth in measuring interaural cross-correlation as a physical measure of auditory source width. Acoustical Science and Technology 26, 2, 2005

[4] J. Blauert and W. Lindemann: Spatial mapping of intracranial auditory events for various degrees of interaural coherence. Journal of the Acoustical Society of America, 79, pp. 806-813, 1986

[5] pure-data, URL: `http://puredata.info`

[6] A. Sontacchi, H. Pomberger, and R. Höldrich: Recruiting and Evaluation Process of an Expert Listening Panel. NAG/DAGA Rotterdam, 2009

[7] M. Frank, A. Sontacchi, and R. Höldrich: Training and Guidance Tool for Listening Panels. DAGA, Berlin 2010

[8] V. Pulkki: Spatial Sound Generation and Perception by Amplitude Panning Techniques. Dissertation, Helsinki University of Technology, 2001

[9] J. Daniel: Représentation de Champs Acoustiques, Application à la Transmission et à la Reproduction de Scènes Sonores Complexes dans un Contexte Multimédia. Dissertation, Université de Paris 6, 2000

[10] M. Morimoto: The Role of Rear Loudspeakers in Spatial Impression. AES 103rd Convention, Kobe, 1997

[11] G. Jahn: Zum Unterschied zwischen einohrigem und zweiohrigem Hören. Hochfrequenztechnik und Elektroakustik, 72, 1963