Output Gain Validation of the Sound xR Image Spatial Renderer

Experimental results on Speaker ID assignment and spatial control parameters

Taebin Yoo | 2026-09-26 | Revision 2.0

Abstract

This study investigates four operational observations in Sound xR Image: assignment-dependent output changes, Object Z invariance, low Size sensitivity, and unexpected position handling. Electrical gain measurements support an inverse-distance, power-normalized candidate with a maximum residual of 0.0046 dB over 12 internal-position conditions. Five Precision settings agree with 8-12 dB per distance doubling before normalization. A revised coupled Size model describes the recorded values from 13 recent XY ON and four XY OFF conditions within 0.0046 and 0.0051 dB, respectively. Speaker ID reassignment nevertheless changes distributions in an unchanged physical layout, and a fixed-segment prior prediction fails in a new five-speaker layout. Invariance in the Object Z sweep is confirmed, but missing mode metadata limit its interpretation. Distance Attenuation separates as a common gain component without supporting one universal reference distance. The report distinguishes measured observations, counterexamples, and post-hoc fits to identify a reproducible core and unresolved selection rules.

Keywords Sound xR Image · spatial audio · black box measurement · DBAP · system identification

1 Motivation and research questions

This study examines how Sound xR Image spatial controls translate into channel gains. Four observations motivated the experiments: changing the assignment of coordinates to Speaker IDs altered the gain distribution despite an unchanged physical layout; Object Z motion produced no gain change in a recorded sweep; Width and Height appeared less responsive at large values; and some positions produced unexpected output concentration or attenuation.

The experiments ask whether coordinates alone determine the output, how Z and Size transform the distance metric, and whether position handling can be separated from Distance Attenuation. The practical objective is a bounded computational model for a Max/MSP prototype. The study does not identify the proprietary implementation or establish a product-wide defect.

2 Experimental method and data treatment

Electrical renderer output gains were recorded on a DME7 testbed using Smaart transfer-function measurements. Speaker coordinates and ID assignments, Object coordinates, XY mode, Width, Height, Precision, and Distance Attenuation were varied, with condition blocks stored in Sheet1 of experiment.xlsx. The analysis concerns static channel gains rather than room response or listening quality. Narrative observations about phase and delay are not treated as verified results from the static table.

Workbook values and recorded settings are the primary evidence. The two DOCX reports and two RTF files document the questions and evolution of the hypotheses. Later counterexamples and explicitly documented settings take precedence over stronger early interpretations. The coordinate 8.66 is used as recorded. Blank cells are not filled by symmetry, and x is not converted to zero or negative infinity. Moderate Precision is an explicit analysis assumption for the early X, XY, and Width blocks where it is not recorded.

Residuals are measured minus predicted channel gains. The count n denotes finite recorded channel values, not independent layouts. Repeated captures and symmetric channels are not independent experiments. The 0.01 dB reporting increment is neither instrument accuracy nor a confidence interval. Firmware, repeated-measurement variance, and full path-calibration records are insufficient to infer measurement uncertainty from the model residuals.

$$e_j = L_j^{\text{meas}} - L_j^{\text{model}}, \qquad E_{\text{RMS}} = \sqrt{\frac{1}{n}\sum_{j=1}^{n} e_j^2} \tag{1}$$

3 Baseline model and Precision

The baseline candidate is inverse-distance weighting with power normalization, a distance-based amplitude panning (DBAP) model. D is the effective distance from a selected calculation position to each speaker, a is the distance exponent, and g is normalized amplitude gain. Equation (2) applies to positive distances; coincident-source handling and output floors are not identified by this expression.

For the triangle at (-5,0), (5,0), and (0,8.66), seven internal X conditions and five internal XY conditions are described within 0.0046 dB. At (0,2.89), the recorded gains are -4.78, -4.78, and -4.76 dB, close to the equal-distance three-output value of -4.7712 dB. The recorded coordinates are retained rather than replaced with an ideal equilateral construction.

The five Precision settings use b=8,9,10,11,12 dB in ascending order. This coefficient describes attenuation per distance doubling before normalization; it is not the slope of the final channel gain or the Distance Attenuation control. Across five conditions and 15 outputs, the maximum residual is 0.0048 dB. These results support the candidate core without validating the position-selection stage at every location.

$$w_i = D_i^{-a}, \qquad g_i = \frac{w_i}{\sqrt{\sum_{k=1}^{N} w_k^2}}, \qquad L_i = 20\log_{10} g_i \tag{2}$$

$$a = \frac{b}{20\log_{10} 2}, \qquad b \in \{8, 9, 10, 11, 12\} \tag{3}$$

Table 1 Recomputed baseline residuals

Condition / Sheet1	Records	Outputs n	RMS dB	Max dB
X · B2:D8	7	21	0.0031	0.0046
XY · B20:D24	5	15	0.0032	0.0045
Precision · B95:D99	5	15	0.0030	0.0048
Width · B46:D54	9	27	0.0029	0.0052

4 Size sensitivity and coupled controls

Width and Height initially appeared nearly inactive above approximately 10. In the planar triangle, however, Sp2 changed from -0.30 to -3.09 dB over Width 0→10 and from -3.09 to -4.75 dB over 10→100. The 50→100 change was only -0.08 dB. This supports progressively reduced sensitivity in that geometry rather than a universal cutoff. In a 3D layout with Height fixed at zero, Width=100 still produced approximately -3.03 dB at Sp1 and -26.51 dB at Sp3; increasing Width does not universally equalize outputs.

Combined-control tests contradicted an independent additive treatment of Width and Height. The latest block uses speakers at (-5,0,0), (5,0,0), (0,5,10), and (0,5,-10), Object=(0,2,0), Moderate Precision, and DA=0. Equation (4) is evaluated for XY ON; equations (5) and (6) for XY OFF. Here r is XY distance and z is speaker height. The XY OFF candidate has only been evaluated at Object Z=0 and is not generalized to arbitrary Object heights.

Across 13 XY ON conditions, the 52 recorded values have a maximum residual of 0.0046 dB. The four XY OFF conditions contain only eight recorded Sp1/Sp3 values, with a maximum residual of 0.0051 dB; empty Sp2/Sp4 cells are excluded. The earlier unconditional minimum-1 hypothesis is contradicted at W=0, H=0.5. The revised equations account for small Size values and coupling, but their fit to model-development data is not independent validation across the full control range. Source: A293:O306.

$$W_e = \max(W, \min(H, 1)), \qquad D_i^2 = r_{xy,\,i}^2 + W_e^2 \quad (\mathrm{XY\ ON}) \tag{4}$$

$$u = \min(\max(W, H), 1), \qquad W_e = \max(W, u), \qquad H_e = \max(H, u) \tag{5}$$

$$D_i^2 = r_{xy,\,i}^2 + W_e^2 + \left(z_i \frac{W_e}{H_e}\right)^2 \quad (W + H > 0) \tag{6}$$

$$D_i^2 = r_{xy,\,i}^2 + z_i^2 \quad (W = H = 0,\ \mathrm{XY\ OFF}) \tag{7}$$

5 Scope of the Object Z observation

Nine Object Z conditions from -2 to +2 in 0.5 increments produced identical four-channel records: -5.48, -5.48, -6.63, and -6.63 dB. Speakers were at (-5,0,0), (5,0,0), (0,5,5), and (0,5,-5), with Object XY=(0,2) and W=H=0. This directly establishes invariance within the recorded sweep. XY mode, DA, and Precision are not specified in this block, so it cannot establish that Object Z is ignored in every XY OFF configuration.

In the latest Z=±10 layout at Object Z=0, Sp1/Sp3 changed from -12.03/-3.59 dB with XY ON to -3.47/-13.02 dB with XY OFF. This demonstrates a mode-dependent effect involving speaker height, but does not isolate Object Z handling. Narratives about larger Z excursions and visible interface movement remain supplementary observations. Source: A82:F91 and A293:O294.

6 Speaker ID reassignment and repeatability

Holding physical coordinates fixed while changing their ID assignments is the central control for a coordinate-only model. The six-speaker layout uses A=(-5,-2), B=(5,-2), C=(-5,2), D=(5,2), E=(0,5), and F=(0,8). At Object=(0,6), Sp1 measured -29.37 dB for ABCEDF and -32.81 dB for ABCDEF. Gains therefore changed even on a channel whose physical coordinate was unchanged, contradicting a mere relabeling of outputs.

During ABCEDF→ABCDEF→ABCEDF→ABCDEF round trips, each assignment returned to the same values at the recorded 0.01 dB precision. Two Object positions, (0,6) and (1,6), two assignments, and two visits yield eight captures but only four distinct settings. At (1,6), both assignments matched direct-position DBAP within approximately 0.012 dB. An ID sequence therefore cannot be classified as universally normal or abnormal. Source: A319:S323.

A new asymmetric five-speaker control swaps Sp3=(-2,4) and Sp4=(3,6) between assignments N and T. The other coordinates are Sp1=(-4,-3), Sp2=(5,-1), and Sp5=(0,1); Object=(1,3), XY ON, Moderate, DA=0, and W=H=0. Unchanged Sp2 and Sp5 shift by +0.66 and +0.43 dB. The physical point (-2,4) changes from -5.82 to -6.99 dB. Maximum direct-position DBAP residuals are 1.1728 dB for N and 0.0128 dB for T.

Table 2 Five-speaker assignment control in dB · A358:Q359

Assignment	Sp1	Sp2	Sp3	Sp4	Sp5
N	-20.03	-16.04	-5.82	-8.85	-2.41
T	-20.04	-15.38	-8.89	-6.99	-1.98

7 Projection candidates and a failed prediction

Some outputs are explained by orthogonally projecting the Object onto a candidate segment and then computing DBAP distances from that point to all speakers. In equation (8), A and B are the segment endpoints and clip limits the coefficient to [0,1]. This is not a model that activates only the two endpoint speakers. The unresolved issue is when projection occurs and which segment is selected.

The Sp2-Sp3 projection candidate fixed before the five-speaker N/T measurements failed, with maximum errors of 9.2876 dB for N and 8.6677 dB for T. After measurement, projecting N onto Sp1-Sp4 yields P=(0.792308,3.161538) and a maximum residual of 0.0064 dB. This is a post-hoc explanation, not a successful prior prediction. The same physical segment exists in T, yet T matches the original Object position; segment existence alone cannot determine the branch.

For the six-speaker ABCEDF layout at Y=6, X=-0.1, 0, and +0.1 are described by one Sp5-Sp6 candidate within 0.0145 dB. A phenomenon confined to the single point X=0 is therefore insufficient. In an irregular quadrilateral, even the interior Object=(0,1) matches an internal-segment projection, limiting a universal nearest-convex-hull explanation. These segments were selected during analysis; an actual ID-dependent mesh or topology implementation remains unconfirmed. Source: A334:H341, A347:F351, and A358:Q359.

$$t = \mathrm{clip}_{[0,\,1]}\left(\frac{(O-A)\cdot(B-A)}{\|B-A\|^2}\right), \qquad P = A + t(B-A) \tag{8}$$

8 Common gain from Distance Attenuation

Matched DA=0 measurements were subtracted from DA=-3 and -6 at the same positions to separate distribution from common gain. The latest triangle block uses XY ON, Moderate, and W=H=0. Four positions have three DA settings and two positions have two, giving 16 output vectors. DA=-6 changes agree across finite channels within approximately 0.01 dB at each position. At the four positions with DA=-3 pairs, the -6 change is approximately twice the -3 change within rounding.

Equation (9) is a candidate for the common component. DA denotes the control setting; d and R depend on the distance definition and an effective reference scale. The finite output at (7,0) changes by -2.45 dB, and outputs at (0,-2) by approximately -2.44 dB. At (5,4), also approximately two coordinate units from the boundary under Euclidean projection, the change is -3.29 dB. A common distance definition with a global R=6 therefore does not explain all directions.

Inverting the paired measurements under the Euclidean boundary-distance assumption gives approximately R=6.12 toward a vertex, 6.13 below the array, and 4.33 obliquely. The earlier R≈5 interpretation depended on assumptions for an initial oblique block without matched DA=0 records; the recent paired comparison takes precedence. These R values are effective fits, not identified internal constants. An alternative distance metric remains indistinguishable from a direction-dependent reference scale. Source: A309:U315.

$$\Delta L = \mathrm{DA}\log_2\left(1 + \frac{d}{R}\right), \qquad G_i = g_i\, 10^{\Delta L/20} \tag{9}$$

9 Discussion and model limits

The results support a working structure comprising distance weighting and power normalization, a Size-dependent effective distance, position selection dependent on assignment and location, and a common DA gain. The core in normal regions and the tested Size combinations yield small residuals, whereas position selection has not been recovered as a general rule. Choosing the best-fitting segment after inspecting each measurement is not a predictive algorithm for new layouts.

Gain distributions may also fail to identify an internal position uniquely. With speakers on a circle centered at the origin, distances from two particular positions can differ only by a common factor, which power normalization removes. A single inverse-position estimate or a small residual therefore does not establish the internal coordinate or mesh.

Early blocks contain missing settings and conflicting records under apparently identical settings, so all data are not pooled into one accuracy statistic. Further work should target a version- and mode-documented Object Z control, prospective tests of a fixed segment-selection rule on new layouts, and separation of the DA distance metric from its reference scale. Completed assignment round trips and ±0.1 perturbations are not presented as outstanding experiments.

10 Conclusion

The study evaluated four operational concerns through output-gain measurements. In normal regions, the DBAP candidate and Precision coefficients reproduce the recorded values. Reduced Size sensitivity depends on geometry and Width-Height coupling rather than a fixed threshold. Object Z invariance is established in the recorded sweep but cannot be generalized because mode metadata are missing. Speaker ID reassignment produces repeatable changes with unchanged physical layouts, and the fixed-pair projection prediction failed. DA separates as common gain at the tested positions, while a universal reference distance is unsupported.

The contribution is a numerical boundary between a reproducible core and unresolved selection rules, rather than a complete replica of the product. A Max/MSP implementation should state its validated conditions and retain unresolved position handling and the DA distance metric as separate candidate modules.

Sources and reproducibility

[1] experiment.xlsx, Sheet1, supplied workbook with 366 rows and 45 columns. Table 1 and the latest Size and assignment comparisons were recomputed from the recorded values. A companion data preview preserves raw x markers and missing entries.

[2] AFC Image Technical Validation Report (Korean DOCX). Initial observations, baseline model, and early experiment summary.

[3] Sound xR Image Technical Review and Detailed Questions (revised Korean DOCX). Later counterexamples, assignment round trips, the new five-speaker prediction, and revised interpretations.

[4] testing result.rtf. Evolution of the initial equations, Width-Height coupling, and DA interpretation. [5] final result.rtf. Recent model comparisons and subsequent interpretations. These files document analysis history and are not counted as independent measurements.

This research record was rewritten from the supplied materials; it is not a peer-reviewed article or an official manufacturer specification. The revision date is not a measurement date. Equations, tables, and prose share the same source across the Korean and English web previews and PDFs.