COMPATIBLE 2-CHANNEL ENCODING OF SURROUND SOUND

Indexing terms: Audio systems, Encoding

Several families of methods are discussed for encoding 2-channel surround sound, with a view to optimising mono and stereo compatibility. We discuss the mono balance, stereo image width and phasiness of various systems with symmetric circle, nonsymmetric circle and 'bent circle' loci on the representational sphere.

Two-channel systems of encoding horizontal surround sound must exhibit compatibility with a number of conflicting criteria. We state possible criteria and examine a number of encoding specifications from the viewpoint of compatibility.

We use the following notations for surround-sound signals prior to encoding: **1** represents an omnidirectional signal (consisting of sounds from all horizontal directions added together with unit gains), $\boldsymbol{x}$ represents a forward-pointing figure-of-eight pickup of signals with unit forward gain, and $\boldsymbol{y}$ a leftward-pointing figure-of-eight pickup with unit leftward gain. Thus signals from azimuth θ anticlockwise from forward are encoded in **1**, $\boldsymbol{x}$ and $\boldsymbol{y}$, with respective gains 1, $x = \cos\theta$, $y = \sin\theta$. We denote the four speaker signals of a pairwise mixed 4-channel encoding[1] by the usual symbols L_B, L_F, R_F, R_B, and we write $W' = \frac{1}{2}(L_B+L_F+R_F+R_B)$, $X = \frac{1}{2}k(-L_B+L_F+R_F-R_B)$, $Y = \frac{1}{2}k(L_B+L_F-R_F-R_B)$ and $Z = \frac{1}{2}k(-L_B+L_F-R_F+R_B)$, where k is a constant between $2^{-\frac{1}{2}}$ and 1.

Let L and R be the left and right stereo channel signals into which surround-sound information has been encoded. For 2-speaker stereo reproduction of a single sound, the stereo position is determined, according to all low-frequency interaural phase theories[2,3] of auditory localisation, by the quantity

$$P = \mathrm{Re}\,\{(L-R)/(L+R)\} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

where $P = 0$ for central images, and $P = 1$ and -1 for left and right speaker image positions, respectively. In Makita's theory of directional perception,[4] P is proportional to the distance from the central position of the apparent image along the line joining the two speakers. The degree of degradation produced by phase differences between the speakers is quite well described[3] by the 'phasiness'

$$Q = \mathrm{Im}\,\{(L-R)/(L+R)\} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

except for nearly antiphase signals.

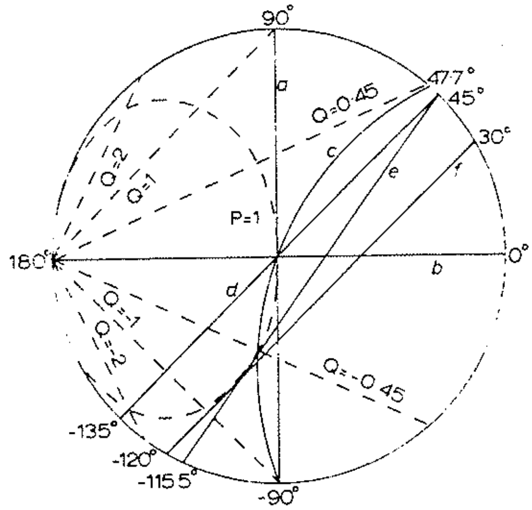

Fig. 1 *Side view of pan loci a to f of various systems of encoding*

Using these concepts, we may state the following criteria (i)–(vi) for mono and stereo compatibility of an encoding specification for 2-channel surround sound. The figures given may be subject to some degree of controversy, and must only be regarded as tentative guidelines:

(i) The total variation in gain for all sound positions for mono, i.e. $L+R$, reproduction should not exceed 6 dB, and back-quadrant sounds should not reproduce louder than front-quadrant sources.

(ii) In stereo, the gain of different sound positions should not vary more than 2·5 dB, and the back quadrant should not reproduce louder than the front quadrant.

(iii) The stereo position P should cover at least the range $-1 \leq P \leq 1$ as a sound rotates 360° around the horizontal stage.

(iv) In the front quadrant, the phasiness should satisfy $|Q| \leq 0{\cdot}45$. In the front three quadrants we should have $|Q| \leq 1$, and for all positions we should have $|Q| < 2$.

(v) The front quadrant should occupy about 0·6 to 0·75 of the stereo stage, i.e. $P \simeq \pm 0{\cdot}65$ for azimuths $\theta = \pm 45°$.

(vi) To ensure symmetry of the stereo image, the representational sphere locus[3,5] of the system should have left/right mirror symmetry.

In addition, the encoding specification must satisfy several criteria (vii)–(x) related to the quality and potential development of surround-sound reproduction:

(vii) The encoding must be such that it is possible to design good linear matrix 4-speaker decoders for rectangle speaker layouts with various aspect ratios and for noncentral listeners.

(viii) The system must be capable of accepting **1**, $\boldsymbol{x}$, $\boldsymbol{y}$ encoded input information, such as derives from coincident microphones, as well as pairwise mixed L_B, L_F, R_F, R_B signals.

(ix) No Z information should be used in the 2-channel encoding of pairwise mixed signals, so as to permit the addition of a third channel conveying, with the 2-channel baseband, **1**, $\boldsymbol{x}$, $\boldsymbol{y}$ or W', X, Y information uncontaminated by Z. Such 3-channel encoding is capable of good results,[1,6] which are likely to be degraded by crosstalk from Z.

(x) The 2-channel encoding must be such as to permit the addition of a bandlimited third channel for discs[1] in a manner permitting one to design good decoding circuits using three channels at lower frequencies, two at higher, while maintaining a substantially flat energy frequency response for all encoded directions.

In Fig. 1, we display the representational sphere loci[3,5] of various systems, viewing from the right side along the right/left axis. Thus angle round the circumference represents phase lead of the left channel over the right, and circular loci satisfying (vi) are viewed edge on and look like lines. The broken ellipse is a side view of points with stereo positions at one speaker or the other. According to (iii), all loci must touch or cross this. The broken lines (actually circles viewed edge on) delimit the regions of phasiness $Q = \pm 0{\cdot}45$, ± 1 and ± 2, as marked.

The marked loci include a, the BMX locus, which satisfies all criteria except (iv); b, the regular matrix, which fails (i) and (v), and includes positions near the antiphase points; c, the BBC Matrix H encoding system[7] for **1**, $\boldsymbol{x}$, $\boldsymbol{y}$ inputs presented in the form of 135°-null hypercardioid inputs, which fails (ix) and (x) and quite possibly (vii); d, the RM or BMX encoding systems with the 0° or 90° interchannel phase differences replaced by 45°—this fails (i), (iv) and (v). e and f are two possible circle loci that satisfy all criteria save (v); a 45° azimuth corresponds to $P = 0{\cdot}45$ and $P = 0{\cdot}41$, respectively, if it is assumed that the encoding of azimuth angle within the circle locus is symmetric.

These represent loci with **1**, $\boldsymbol{x}$, $\boldsymbol{y}$ inputs, but the encoding of pairwise mixed signals may be achieved by substituting these by W', X, Y, respectively, with $k = 1$ if centre phantom images are to be correctly encoded, and $k = 2^{-\frac{1}{2}}$ if corner images are to be correct.

Symmetric encoding within a circle locus means that the two transmission channels are complex linear combinations of **1** and $\boldsymbol{x} - j\boldsymbol{y}$ (or else $\boldsymbol{x} + j\boldsymbol{y}$ and **1**, respectively). We may produce asymmetric encoding of directionality within the same circle loci by replacing these two signals in the encoding equations by $\boldsymbol{1} + aj\boldsymbol{y}$ and $\boldsymbol{x} - (1+a^2)^{\frac{1}{2}} j\boldsymbol{y}$, respectively. For a typical value $a = 1/(2\sqrt{2})$, this does not cause stereo gains to vary by more than 0·5 dB with azimuth, but widens the front-quadrant stereo image width by moving points on the circle locus, as shown in Fig. 2. This enables loci e and f to satisfy condition (v), although there is a slight compromise with condition (iv).

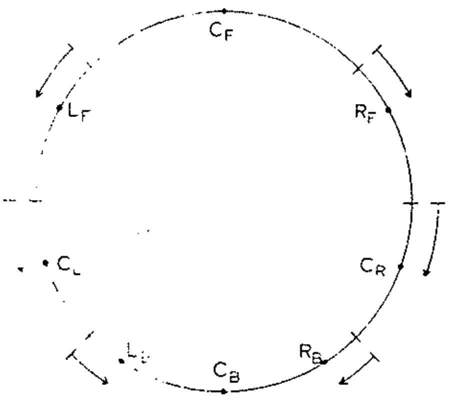

Fig. 2 *Positions of various azimuths on circle locus for 'boosted' or 'stereo-widened' encoding system*

We may produce 'bent circle loci' similar to locus c by replacing the signals $\mathbf{1}$ and $\mathbf{x}-j\mathbf{y}$ in the encoding equations for symmetric circle loci by signals of the form $\mathbf{1}$ and $b\mathbf{x}-j\mathbf{y}$ for constant $b < 1$ (typically, $b = 0{\cdot}7$). More generally, the slight loss of gain at the front relative to the sides of this system may be overcome by replacing these two signals by ones of the form $\mathbf{1}+\lambda\mathbf{x}+b^{-1}\lambda j\mathbf{y}$ and $\lambda b\mathbf{1}+b\mathbf{x}-j\mathbf{y}$, respectively, for some λ a little larger than 0; for example, $\lambda = 0{\cdot}1$ and $b = 0{\cdot}7$ are useful choices. However, we note that no 'bent-circle-locus' system can satisfy condition (x).

A previous letter[8] in this journal has proposed provisional 2-channel encoding standards within which one can devise systems satisfying or nearly satisfying all criteria (i)–(x) stated above.

M. A. GERZON

3rd November 1975

Mathematical Institute
24–29 St Giles, Oxford OX1 3LB, England

References

1 COPPER, D. H., and SHIGA, T.: 'Discrete-matrix multichannel stereo', *J. Audio Eng. Soc.*, 1972, **20**, pp. 346–360
2 GERZON, M. A.: 'Surround sound psychoacoustics', *Wireless World*, 1974, **80**, pp. 483–486
3 GERZON, M. A.: 'A geometric model for 2-channel 4-speaker matrix stereo systems', *J. Audio Eng. Soc.*, 1975, **23**, pp. 98–106
4 MAKITA, Y.: 'On the directional localization of sound in the stereophonic sound field', *EBU Rev.* 1962, (73), pp. 102–108
5 SCHEIBER, P.: 'Analyzing phase-amplitude matrices', *J. Audio Eng. Soc.*, 1971, **19**, pp. 835–839
6 GIBSON, J. J., CHRISTENSEN, R. M., and LIMBERG, A. L. R.: 'Compatible f.m. broadcasting of panoramic sound', *ibid.*, 1972, **20**, pp. 816–822
7 Based on work by CROMPTON, T. W. J.: 'The subjective performance of various quadraphonic matrix systems'. BBC Research Department Report BBC RD 1974/29, 1974
8 FELLGETT, P. B., and GERZON, M. A.: 'Encoding standards for 2-channel surround-sound', *Electron. Lett.*, 1975, **11**, 518–519