

Characterization of Audio Items

For the characterization of most of the audio samples listening tests were conducted. The test method used is according to ITU-R BS.[6/106], nick named MUSHRA (Multiple Stimulus with Hidden Reference and Anchors). This method is also into consideration for within AES SC (project X74)

MUSHRA

Several instances produced from the same original audio item are presented to a subject in random order. Among these instances there are the predefined conditions "low pass at 7 kHz" and "low pass at 3.5 kHz" as hidden anchors, and the unprocessed original (hidden reference). The subject are allowed to switch between all instances and the known unprocessed original in any order. The listeners is asked to compare each instance to this unprocessed original and to each other, and has to rank them using the 5 interval quality scale (see Figure 1). Note, that the hidden reference is scored by the subjects, too. The subjects are told to score the instance which is most similar to the known reference at top of the scale (=100). However, if there are one or more other instances which are very similar to the unprocessed original, the average score obtained for the hidden reference sometimes drop below 100.

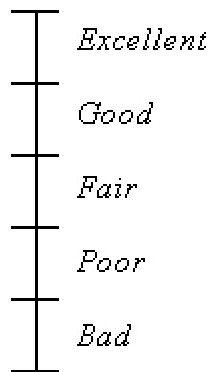


Figure1: 5 interval quality scale

References:

- [1] ITU-R SG6: Draft New Recommendation ITU-R BS.[Doc. 6/106]; "Method for the subjective assessment of intermediate audio quality

Test Setup

The test panel consisted of 9 subjects, 5 female and 4 male, aged 21 to 39 years (average 26). The audio items were played back from a PC using a digital I/O card with external DA converters and using STAX Lambda Pro headphones.

Note1: As a rough estimation of the severeness of the distortions only the average of the scores, rounded to full numbers, are indicated. A formal listening test usually would require a careful statistical evaluation including an indication of the size of the confidence intervals.

Test Results

Track	MUSRA Score
Preecho	
1	100
2	35
Aliasing	
3	NA
4	NA
5	86
6	21
7	81
Birdies/ Band limitation	
8	89
9	85
10	61
11	21
12	97
13	76
14	43
15	20
16	NA
17	NA
18	NA
19	100
20	48
21	36
22	23
23	69
24	62
25	48
26	34

Track	MUSRA Score
Speech Reverberation	
27	95
28	29
29	40
30	57
31	64
32	54
33	74
34	81
35	89
36	64
37	86
38	88
39	98
40	74
41	89
42	90
43	95
44	NA
45	NA
BMLD	
46	NA
47	NA
48	NA
49	NA
Stereo Imaging	
50	95
51	88
52	90
53	85
54	83

Track	MUSRA Score
Tandem Coding	
55	99
56	91
57	87
58	77
59	66
60	65
61	58
62	55
63	49
64	42
65	NA
Speech Coding	
66	52
67	42
68	44
69	25
70	67
71	24
72	46
73	12
74	44
75	18
76	98
77	86
78	99
79	73
80	100
81	29
82	100
83	33
84	100
85	40

Track	MUSRA Score
MPEG Filtering	
86	NA
87	NA
88	NA
89	NA
90	NA
91	NA
92	NA
93	NA
94	NA
95	NA
96	NA
97	NA

Notes:

NA

not available (no listening test done)

Speech Reverberation:

scores of Track27 are context dependent:
Context Track28 to Track31: Track27: 100
Context Track32 to Track35: Track27: 92
Context Track36 to Track39: Track27: 90
Context Track40 to Track43: Track27: 96

Stereo Imaging:

very large deviations between scores of different subjects

Speech Coding:

16kHz used as reference, 8kHz used as hidden anchors,
no additional hidden anchor