Subjective assessments of spherical microphone arrays
Paired comparisons of two arrays designed using different microphone models

Vincent Koehl, Mathieu Paquier and Symeon Delikaris-Manias
Université de Bretagne Occidentale – UEB, Lab-STICC UMR 6285
6, avenue Victor Le Gorgeu, CS 93837
29238 Brest Cedex 3, France
vincent.koehl@univ-brest.fr

Abstract

Microphone arrays are commonly used to capture sound fields. As the number of sensors forming the array increases, the spatial sampling accuracy at high frequencies improves. Numerous prototypes of spherical arrays were developed over the last years. However, much less attention has been paid to the intrinsic performances of the sensors than to their number and arrangement. This study aims at evaluating the relative performances of two rigid spherical microphone arrays of the exact same size differing only in their capsules (pressure sensors). The two recording systems are based on higher order ambisonics and were used to acquire the exact same sound scene. Four short music excerpts were decoded as various types of audio content (mono, stereo and multichannel) and displayed through dedicated loudspeaker setups. The recordings issued by the two arrays were then to be compared by pairs, on a similarity basis and on a preference one, by twelve expert listeners (sound engineering students). The results showed that the perceived differences and preferences depended on the way stimuli were rendered. These assessments were consistent with those obtained from naive listeners in a previous study, although experts perceived significantly better the differences and reported more pronounced preferences.

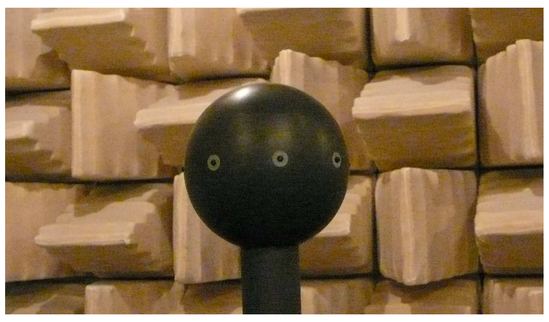
Aim of the study

Evaluate the perceptual benefit brought by high quality capsules to microphone arrays, by comparing:
• raw and decoded signals stemming from two arrays (differing only in the capsule type),
by pairs in terms of similarity and preference to determine whether:
• the two arrays are clearly perceived as different,
• the subjective differences could result in modifications of the audio quality,
according to various presentation methods (monophonic, stereophonic and multichannel).

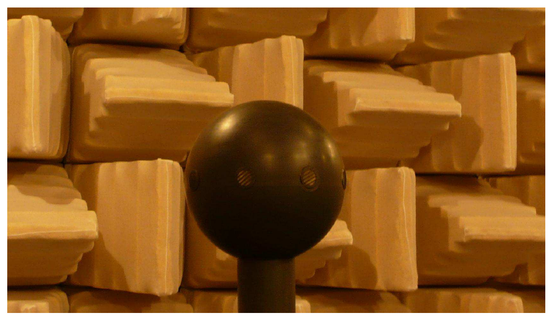
Microphone arrays

Two rigid spheres (radius 8 cm) fitted each with 8 omnidirectional capsules (pressure sensors).

Array *A* (miniature microphones)



Array *B* (small membrane microphones)



Objective differences between the two arrays were evaluated in a previous study by measuring the capsule responses and by calculating the spherical harmonics components:

[1] S. Delikaris-Manias, V. Koehl, M. Paquier, R. Nicol and J. Daniel, "Does capsule quality matter? A comparison study between spherical microphone arrays using different types of omnidirectional capsules", *2nd international symposium on ambisonics and spherical acoustics (Ambisonics Symposium 2010)*, Paris, France (2010)

Based on their frequency responses and signal-to-noise ratios, the small membrane microphones (array *B*) were of higher quality.

Recordings and listening tests

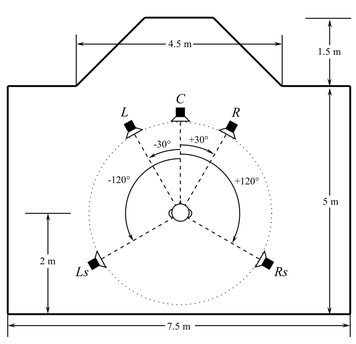
The two (assumed coincident) arrays were used to record 4 short music excerpts (approx. 5 s) in order to compare them. (bit depth of 16 bits, sample rate of 48 kHz)

For each excerpt, the recordings were used to generate 6 different stimulus types:

1. Raw monophonic signal (single capsule)
2. Monophonic decoding (*W* component)
3. Raw stereophonic signal (2 capsules)
4. Stereophonic decoding (1st order)
5. Multichannel decoding (3rd order)
6. Multichannel decoding (4th order)

For all excerpts and types, the stimuli issued by the two arrays had to be compared by pairs about similarity and preference by 12 expert listeners.

Paired comparisons were to be achieved by playing back the two stimuli using dedicated loudspeaker setups:
• Central loudspeaker (*C*) for monophonic stimuli,
• Stereo pair (*L*-*R*) for stereophonic stimuli,
• ITU setup (*L*-*C*-*R*-*Ls*-*Rs*) for multichannel stimuli,
in a listening room at a realistic level.



Results

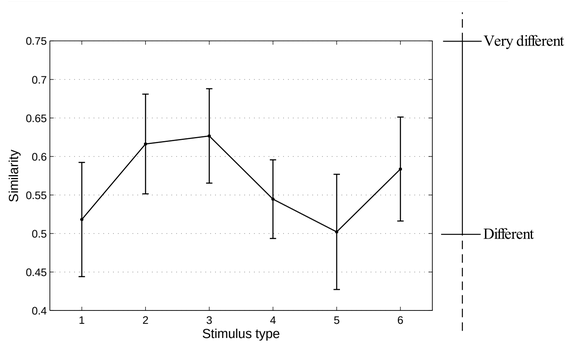
Analyses of variance were carried out to examine the factor effects on similarity and preference ratings:

Similarity

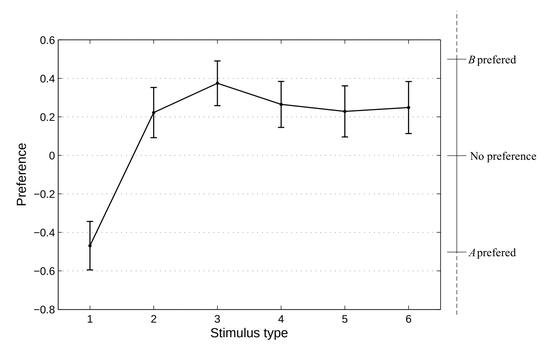
Source	SS	DF	MS	F	p
Stimulus type	0.64	5	0.13	2.43	<.05*
Excerpt	0.07	3	0.02	0.46	0.71
S*E	0.59	15	0.04	0.75	0.73
Error	13.92	264	0.05		
Total	13.22	287			

Preference

Source	SS	DF	MS	F	p
Stimulus type	22.44	5	4.49	23.56	<.001***
Excerpt	0.58	3	0.19	1.02	0.38
S*E	2.82	15	0.19	0.98	0.47
Error	50.29	264	0.19		
Total	76.13	287			



Differences were well perceived for all stimulus types but could significantly differ from one to another.
As an example, differences between the two arrays were significantly better perceived with stimulus type 2 (mono – omnidirectional component) than with type 1 (mono – frontal capsule).



The preference towards one of the two arrays depends on the type of stimulus used for the comparison.
Array *B* (high quality sensors) is preferred for all stimulus types except type 1 (mono – frontal capsule) where *A* is surprisingly preferred (could be due to differences in the capsule positions).

Comparison to subjective assessments obtained from naive listeners in a previous study:

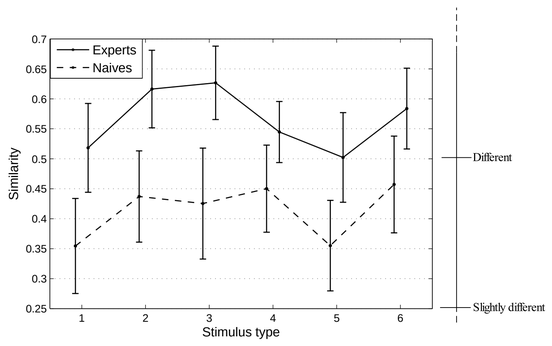
[2] S. Delikaris-Manias, V. Koehl, M. Paquier, R. Nicol and J. Daniel, "A comparative study of spherical microphone arrays based on subjective assessment of recordings reproduced over different audio systems", *Forum Acusticum*, Aalborg, Denmark (2011)

Similarity

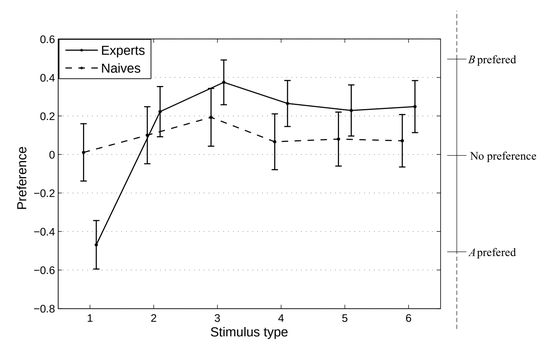
Source	SS	DF	MS	F	p
Group	3.46	1	3.46	56.64	<.001***
Stimulus type	1.02	5	0.20	3.11	<.01**
Excerpt	0.33	3	0.11	1.67	0.17
G*S	0.18	5	0.03	0.55	0.74
G*E	0.23	3	0.08	1.18	0.32
S*E	2.04	15	0.14	2.07	<.05*
G*S*E	0.67	15	0.04	0.68	0.8
Error	36.33	552	0.07		
Total	44.34	599			

Preference

Source	SS	DF	MS	F	p
Group	0.51	1	0.50	2.19	0.14
Stimulus type	15.47	5	3.09	13.36	<.001***
Excerpt	1.53	3	0.51	2.21	0.09
G*S	8.76	5	1.75	7.56	<.001***
G*E	0.38	3	0.13	0.55	0.65
S*E	2.93	15	0.19	0.84	0.63
G*S*E	3.56	15	0.24	1.03	0.43
Error	127.88	552	0.23		
Total	160.22	599			



Similarity ratings were significantly higher for expert listeners. The differences between the two arrays were better perceived by experts than by naives.



Preference ratings (towards array *A* or *B*) were significantly higher for the expert group than for the naïve one (for which the stimulus type had no effect).

Discussion and conclusion

The use of different capsule models in the two microphone arrays resulted in significant perceptual differences, whatever the restitution method.

As a rule, with the exception of monophonic signals issued by the frontal capsule, and as expected, the array made of high quality microphones has been systematically preferred.

Therefore, the use of higher quality capsules in microphone arrays may enhance the perception of sound renderings.

This benefit should be more significant for expert listeners. Their assessments were consistent to naive ones, but they perceived significantly better the difference between the two arrays and attributed a higher mark to the preferred one.

ACOUSTICS 2012