

Coincident Microphone Techniques and Ambisonics

Yangang Cao

March 6, 2019

In coincident microphone technologies first- or higher-order microphones positioned ideally in the same position are used to capture sound for multichannel playback. The ambisonics technique is a form of this. The most common microphone for these applications is the first-order four-capsule B-format microphone, producing a signal $w(t)$ with omnidirectional characteristics, which has been scaled down by $\sqrt{2}$. The B-format microphone also outputs three signals $x(t)$, $y(t)$, and $z(t)$ with figure-of-eight characteristics pointing to corresponding Cartesian directions. The microphones are ideally in the same position. Higher-order microphones have also been proposed and are commercially available, which have much more capsules than the first-order microphones.

In most cases, the microphones are of first order. When the loudspeaker signals are created from the recorded signal, the channels are simply added together with different gains. Thus each loudspeaker signal can be considered as a virtual microphone signal having first-order directional characteristics. This is expressed as

$$s(t) = \frac{2 - \kappa}{2}w(t) + \frac{\kappa}{2\sqrt{2}}[\cos(\theta)\cos(\phi)x(t) + \sin(\theta)\cos(\phi)y(t) + \sin(\phi)z(t)],$$

where $s(t)$ is the produced virtual microphone signal having an orientation of azimuth θ and elevation ϕ . The parameter $\kappa \in [0, 2]$ defines the directional characteristics of the virtual microphone from omnidirectional to cardioid and dipole.

In multichannel reproduction of such first-order B-format recordings, a virtual microphone signal is computed for each loudspeaker. In practice such methods provide good quality only in certain loudspeaker configurations and at frequencies well below 1 kHz. At higher frequencies the high coherence between the loudspeaker signals, which is caused by the broad directional patterns of the microphones, leads to undesired effects such as coloration and loss of spaciousness. Using higher-order microphones these problems are less severe, though some problems with microphone self-noise may appear.

The decoding of first-order B-format microphone signals to a horizontal loudspeaker layout with directional characteristics controllable by κ is shown in the following example:

```
1 % Simple example of cardioid decoding of B-format signals
2
3 Fs=44100;
```

```

4 % mono signal
5 signal=(mod([0:Fs*2],220)/220);
6 % Simulated horizontal-only B-format recording of single
7 % sound source in direction of theta azimuth.
8 % This can be replaced with a real B-format recording.
9 theta=0;
10 w=signal/sqrt(2);
11 x=signal*cos(theta/180*pi);
12 y=signal*sin(theta/180*pi);
13 % Virtual microphone directions
14 % (In many cases the values equal to the directions of loudspeakers)
15 ls_dir=[30 90 150 -150 -90 -30]/180*pi;
16 ls_num=length(ls_dir);
17 % Compute virtual cardioids (kappa = 1) out of the B-format signal
18 kappa=1;
19 for i=1:ls_num
20     LSsignal(:,i)=(2-kappa)/2*w...
21     +kappa/(2*sqrt(2))*(cos(ls_dir(i))*x+sin(ls_dir(i))*y);
22 end
23 % File output
24 audiowrite(LSsignal,Fs,16,'firstorderB-formatexample.wav')

```

The previous example was of reproducing a recorded sound scenario. Such coincident recording can also be simulated to perform synthesis of spatial audio for 2-D or 3-D loudspeaker setups. In this case it is an amplitude-panning method in which a sound signal is applied to all loudspeakers placed evenly around the listener with gain factors

$$g_i = \frac{1}{N} \sum_{m=1}^M \{1 + 2p_m \cos(\alpha m_i)\},$$

where g_i is the gain of i th speaker, N is the number of loudspeakers, α is the angle between loudspeaker and panning direction, $\cos(\alpha m_i)$ represents a single spherical harmonic with order m , M is the order of Ambisonics, and p_m are the gains for each spherical harmonic. When the order is low, the sound signal emanates from all the loudspeakers, which causes some spatial artifacts due to unnatural behavior of binaural cues. In such cases, when listening outside the best listening position, the sound is also perceived at the nearest loudspeaker which produces the sound. This effect is more prominent with first-order ambisonics than with pair- or triplet-wise panning, since in ambisonics virtually all loudspeakers produce the same sound signal. The sound is also colored, for the same reason, i.e., multiple propagation paths of the same signal to ears produce comb-filter effects. Conventional microphones can be used to realize first-order ambisonics.

When the order is increased, the cross-talk between loudspeakers can be minimized by optimizing gains of spherical harmonics for each loudspeaker in a listening setup. Using higher-order spatial harmonics increases both directional and timbral virtual source quality, since the loudspeaker signals are less coherent. The physical wave field reconstruction is then more accurate, and different curvatures of wavefronts, as well as planar wavefronts can be produced, if a large enough number of loudspeakers is in use. The selection of the coefficients for different spherical harmonics has to be done carefully for each loudspeaker layout.

A simple implementation of computing second-order ambisonic gains for a hexagonal loudspeaker layout is shown here:

```

1  % Second-order harmonics to compute gains for loudspeakers
2  % to position virtual source to a loudspeaker setup
3
4  theta = 30; % Direction of virtual source
5  loudsp_dir = [30 -30 -90 -150 150 90]/180*pi; % loudspeaker setup
6  ls_num = length(loudsp_dir);
7  harmC = [1 2/3 1/6]; % Coeffs for harmonics "smooth solution", ...
   [MON00]
8  theta = theta/180*pi;
9  for i = 1:ls_num
10     g(i) = (harmC(1) + 2*cos(theta-loudsp_dir(i))*harmC(2) +...
11            2*cos(2*(theta-loudsp_dir(i)))*harmC(3));
12 end % use gains in g for amplitude panning

```