Creation of “quad binaural” signals

Quad Binaural is a crappy format for spatial audio created by Chris Milk, and resembling the “quad stereoscopic” 3D panoramic video recording method (which is also crappy).

In practice, it I a set of 4 binaural recordings, made pinting the dummy head in four direction:

* 0° (Front)
* 90° (Right)
* 180° (Back)
* 270° (Left)

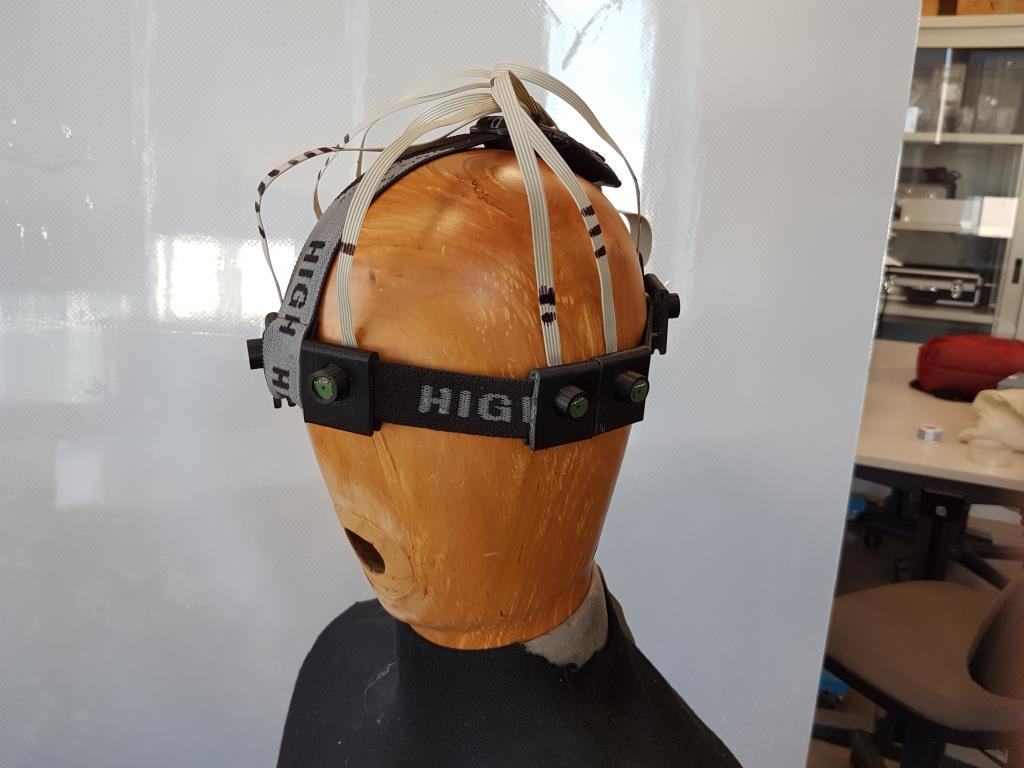
It must be noted that the angular system is reversed compared to the ISO standard, in which 90° is “left” and 270° is “right”.

The Quad Binaural signal can be synthesized from High Order Ambisonics, with poor results. Or it can be recorded directly, with even poorer results, employing a microphone system such as the following ones:



Quad Binaural microphones

We propose two other ways to record “Quad Binaural” signals, employing a wearable microphone array, made of 8 MEMS PCM digital microphones (Invensense) mounted on an elastic band surrounding the head, as shown here:



A wearable MEMS microphone array on a wooden dummy head.

The first method proposed is based just on directive shiedling effect of the head, so we select the “Proper” pair of MEMS microphone for feeding each of the binaural stereo pairs.

The microphones are numbered counter-clockwise starting from Front-Left (#1) and going to Front-Right (#8) as shown here:

8

1

7

2

6

3

4

5

Microphone positions

So the pairs selected for assembling the 4 stereo files are:

* Front (0°) 2 + 7
* Right (90°) 8 + 5
* Back (180°) 6 + 3
* Left (270°) 4 + 1

Of course this approach is suboptimal, as at frequencies below 1 kHz the shielding effect is modest, and the only directional cue is the time delay between the microphones. But also above 1 kHz the directional effect due to shielding is quite different from the directivity of human ears.

So we tested with a different approach. The microphones are grouped in pairs, and each pair is treated like a dipole.

So we can derive two cardioid microphones pointed one in direction opposite to the others. This is obtained employing Gary Elko’s “differential first order microphone” theory.

In practice , considering for example the pair of microphones 2 and 3 (placed on the left side of the head), a forward-pointing cardiod is obtained as:

LF(t) = x2(t)-x3(t-T)

Where **T** is the travel time for a sound wave coming from the Front direction, which equals **d/c** (if **d** is the distance between the two microphones).

Of course, a backward-pointing microphone is obtained as

LB(t) = x3(t)-x2(t-T)

However, the frequency response of these microphones is distorted, as the differentiator operator boosts high frequencies, with an approximate slope of 6 dB/octave. In reality, the boost is smaller above a certain frequency, when the distance between the capsules **d** starts to become significant compared with the wavelength. IN any case, some amount of low-pass equalization is required.

So we applied this processing to the 4 microphone pairs, taking into account that the distances d are not the same.

In practice, we did employ recordings of hand-clapping for measuring the delays T.

And for making this more accurate, we first oversampled the signal at Fs=192 kHz.

After measurements, the following delays were applied to each microphone pair:

* L (2+3) T = 19 samples
* F (1+8) T = 27 samples
* R (7+6) T = 19 samples
* B (4+5) T = 58 samples

So we computed these 8 “beamformed” signals:

LF(t) = F[x2(t)-x3(t-19)]

LB(t) = F[x3(t)-x2(t-19)]

FL(t) = F[x1(t)-x8(t-27)]

FR(t) = F(x8(t)-x2(t-27)]

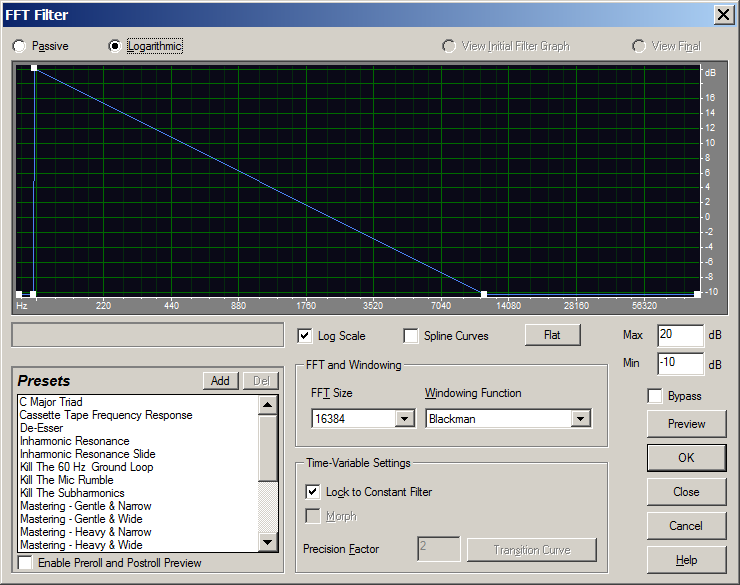
RF(t) = F[x7(t)-x6(t-19)]

RB(t) = F(x6(t)-x7(t-19)]

BL(t) = F[x4(t)-x5(t-58)]

BR(t) = F(x5(t)-x4(t-58)]

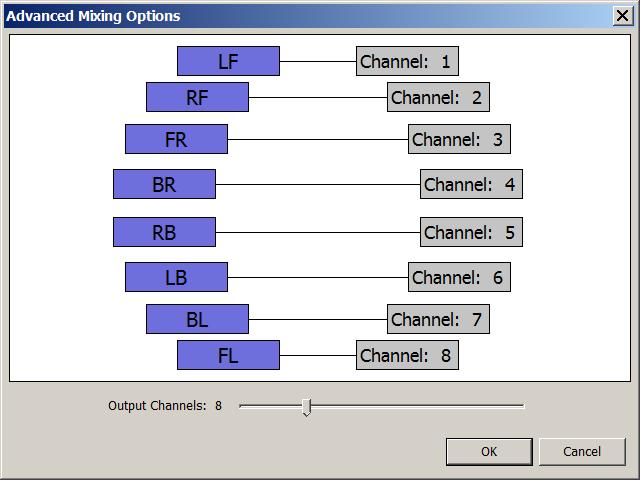
Where F[] is the equalizing function, which is shown here:



The equalizing filter F

The signals were downsampled to Fs=48 kHz, and finally reassembled in the Quad Binaural signals as follows:

* Front (0°) LF + RF
* Right (90°) FR + BR
* Back (180°) RB + LB
* Left (270°) BL + FL



Channel mapping for Quad Binaural with Cardioid virtual mikes

So we have got two “quad binaural” recordings coming from the MEMS wearable array (called QB-omni and QB-cardio), plus a third “reference” AMBIX recording (1st order horizontal) coming from a Zoom H2N mounted over the dummy head, as shown here:



The setup employed for recording

A shown above, the MEMS array was connected directly to the USB port of a Samsung S7 smartphone (running the excellent USB Audio Recorder app), whilst the panoramic video was recorded by a Samsung Gear 360 dual-camera, mounted above the Zoom recorder, and operating at a resolution of 3840x1920, 29.97 FPS.

The three resulting AV immersive recordings , in MKV format, can be seen and heard employing a Samsung Gear VR system, and employing the Oculus Video Player software (which currently is the only player supporting simultaneously Ambix and Quad-Binaural “spatial audio” soundtracks).



The Samsung Gear VR visor, with headphones