Considerations on Stereo and Surround recording, reproduction and perception

Helmut Wittek
SCHOEPS GmbH, Karlsruhe, Germany
Contents:

- An overview of psychoacoustics and applied microphone techniques
- Both theory and practical demonstration, by calculating and listening

Topics

• Source localisation / stereophonic localisation
• Theories of microphone design
• Microphone setups and their properties
• The microphone and its properties
Source localisation

- Where is the source?
- At which distance?
- In which room, what is the room like?
- What does the source radiate?
Stereophonic localisation

- $\geq 2$ radiating sources
- Only one perceived “phantom” source

Existing theory:

*Summing localisation = synthesis of the loudspeaker signals that leads to the original source signal* (Blumlein)
Real source, 15° on the right

Phantom Source, perceived ca. 15° on the right, \( \Delta L=7\text{dB}, \Delta t=0\text{ ms} \)

Phantom Source, perceived ca. 15° on the right, \( \Delta L=3.5\text{dB}, \Delta t=0.2\text{ ms} \)

Phantom Source, perceived ca. 15° on the right, \( \Delta L=0\text{dB}, \Delta t=0.4\text{ ms} \)
Summing localisation

**Interaural Cross Correlation**
( = Interaural Time Delay ITD vs. frequency)

**Real source**
Perceived Direction + 15°

**Phantom source**
Perceived Direction + 15°
Interchannel Level Difference $\Delta L = 7$ dB
Summing localisation

*Interaural Cross Correlation*  
(= *Interaural Time Delay ITD vs. frequency*)

**Virtual source**  
Perceived Direction + $15^\circ$

**Phantom source**  
Perceived Direction + $15^\circ$  
Interchannel Time Delay  
$\Delta t = 0,4$ ms
Other hypothesis

- **Problems of the summing localisation theory:**
  - Works only below ca. 1500 Hz
  - Works only for level panning
  - Works only in the sweet spot
  - It cannot explain the perceived sound colour

- Other hypothesis for stereophonic localisation/perception:
  Stereo signals can be perceived *separately*
  - „Binaural decolouration“ (Salomons, Brüggen)
  - „Association model“ (Theile, 1980)
    - After the separate localisation, the fusion of the coherent signals takes place → no physical superposition, no comb filtering
Consequences for the sound engineer

1. Do we aim at a „physical“ synthesis in stereo? (as do Ambisonics, Blumlein)
2. Can only coincident microphones create spatial sound?
3. Are time differences allowed? (see e.g. Lipshitz: „Are the purists wrong?“, JAES)
4. Can we use microphones at ear spacing? (like ORTF, SCHOEPS sphere microphone)
5. Is it wise to allow crosstalk in surround microphones? (see Lee/Rumsey, AES)
Source localisation

- Where is the source?
- At which distance?
- In which room, what is the room like?

Spatial impression

- Direction
- Distance, Depth, Spaciousness
- Envelopment
- Reverberance

© Theile
Source localisation

Discrete Signals:
- Correlated at both ears
- From discrete directions

Diffuse Signals:
- Decorrelated at both ears
- From all directions

Spatial impression

- Direction
- Distance, Depth, Spaciousness
- Envelopment
- Reverberance

© Theile
Directions and points in time

Direct Sound
Early Reflections
Reverb
1. Envelopment is hardly possible in 2-ch stereo
2. No real depth is possible in 2-ch stereo
3. Reflection density too high! → has to be reduced
4. Reverb should be reproduced as diffuse as possible → decorrelated!
Diffuse field correlation

Reproduction of the reverb tail:
- Should be diffuse, should be perceived from everywhere
- The more correlation, the narrower the image of the reverb tail

The diffuse sound should be reproduced decorrelated. The decisive measure is called:

*Diffuse field correlation* (DFC)
• Perceptual consequence of correlated diffuse field reproduction → Demo

• DFC of coincident microphone setups:

<table>
<thead>
<tr>
<th>Setup</th>
<th>XY, 90°, Cardioids</th>
<th>XY, 120°, Supercardioids</th>
<th>Blumlein, 90°, Figure-8</th>
<th>XY, 180°, Cardioids</th>
</tr>
</thead>
<tbody>
<tr>
<td>DFC</td>
<td>0.75</td>
<td>0.23</td>
<td>0</td>
<td>0.5</td>
</tr>
</tbody>
</table>

![Diffuse field correlation diagrams](Image)
a/b and ORTF correlation

• DFC of spaced microphones → Calculation

• Shuffling techniques could correct for that…
Directional Imaging

- Level and time differences govern the perceived phantom source direction
- \( \phi_L = f(\Delta L) \)
- \( \phi_t = f(\Delta t) \)
- \( \phi = \phi_L + \phi_t \)
- The necessary interchannel differences are rather similar to those known from natural hearing
Comparison of stereo setups

- Recording angle
- Recording angle 75%
- Image Assistant on www.hauptmikrofon.de
- Comparison of different stereo setups sharing the same recording angle ➔ Demo Showroom

www.schoeps.de/showroom/
2-ch Microphone Design

• *Directional Image*: can be calculated, differences between stereo techniques regarding the image

• *Diffuse Field Correlation DFC*: should be minimal: not possible with small a/b or normal cardioid XY

• *Further points of interest*:
  – Reality is not always ideal!
  – Type and properties of single capsules (pressure/pressure gradient transducer)
  – Direct/Reverb ratio $\rightarrow$ controls reflection level
  – Loudness balance
  – Size of the setup
  – Sensitivity to wind noise
MS

CMC 64 + CCM 8

RCY

SGMSC
(CCM4 + CCM8)

CMIT MS

UMS 20

WSR MS

MS-BLM

SCHOEPS Mikrofone
ORTF and AB Stereo

Decca
3 * CCM 2H

UMS 20

KFM 6

MSTC 64 U, STC

M100C

MAB1000

ORTF Outdoor Set
5.1 Reflection pattern

Direct Sound
Early Reflections
Reverb
5.1 Spatial reproduction

- Imaging area
- Early reflections
- Reverb
Surround Microphone Design

• Directional Image:
  – Image Assistant
  – Crosstalk: no 3 correlated signals!

MK 41
MK 41V

L, R: Super-Cardioid
C: Cardioid
b: 40 .... 100 cm
h: 8 cm

OCT
Stability in 5.1

- Localisation curve on off-centre position

![Graph showing perceived source direction vs sound source direction]

- Perceived source direction $\theta$
- Sound source direction $\Omega$
- Data points for OCT 70, INA 3, and Quasi-ORTF
OCT Surround

Early reflections
Reverb

L

C

h

R

b

40 cm

b + 20 cm

LS

RS
OCT Surround Setup

OCT Surround (using MAB1000 and CCM microphones)

Front + 20cm

40 cm

8 cm

40 - 100 cm
OCT + Hamasaki square

L

h

b

C

R

Early reflections

Reverb

L + R > 2 m

L + S > 2 m

R + S > 2 m

OCT + Hamasaki

SCHOEPS Mikrofone
IRT cross

- IRT cross for Surround atmos: 20-25cm
Surround – non-coincident

**OCT Surround**

**Decca**
3 * CCM 2H

**KFM 360**

**IRT cross**

**KFM 360**
mit 3 Kondensatormikrofonen CCM 8ig (Acht)

DSP-4 KFM 360
Steuerung mit integrierten A/D- und
D/A-Wandlern
Double MS: The M/S principle

\[ M = L + R \]
\[ S = L - R \]

\[ L = \frac{1}{2} \times (M + S) \]
\[ R = \frac{1}{2} \times (M - S) \]
The Double M/S idea

- Front M/S pair
- Rear M/S pair
- Combined Double M/S triplet
The Double M/S decoding

Double M/S enables 2 different decoding methods:

a) each pair is decoded separately
b) decoding utilizes the third microphone as well

\[
\begin{align*}
\text{L} & \quad \text{R} \\
\text{a)} \\
\end{align*}
\]

\[
\begin{align*}
\text{L} & \quad \text{R} \\
\text{b)} \\
\end{align*}
\]
Crosstalk

- Crosstalk in Double M/S

→ Avoid crosstalk by optimal decoding
**Tools for Decoding**

Decoding variants:

- **2 M/S Matrices**
- **Hardware (MDMS U)**

- **Software (VST PlugIn)**

- **Try by yourself, it’s free!**

[www.schoeps.de/dmsplugin.htm](http://www.schoeps.de/dmsplugin.htm)
Film recording

- Double M/S with shotgun
Surround - coincident

CMIT-Double M/S

Double M/S
Higher Order Techniques

• Which higher order technique can be used in practice?

for reproduction:
• Wavefield synthesis (WFS) and Higher Order Ambisonics (HOA) can recreate sound fields
• Stereo can (only) generate the same perception in the sweet spot

for recording:
• Higher order microphone techniques are either noisy and/or can not yet deliver a fully satisfying and stable timbral quality, but proposals will come for special applications
WFS or Stereo?

Wittek, Rumsey, Theile, Journal of the AES, Vol.55/9, 2007:

*Colouration of WFS and Stereo:*

*Is stereo just a poor 2-ch wavefield synthesis?*
WFS or Stereo?

Wittek, Rumsey, Theile, Journal of the AES, Vol.55/9, 2007:

*Colouration of WFS and Stereo*

**Perceived Colouration: Means**

Error Bars show 95.0% CI of Mean
Bars show Means
What does the microphone need?

• The microphone is the basis for all.
• Well-known parameters govern the quality:
  – Good frequency response
  – Smooth polar pattern
  – Low noise floor
  – Good reliability
  – And the sound…

Thank you.

wittek@schoeps.de