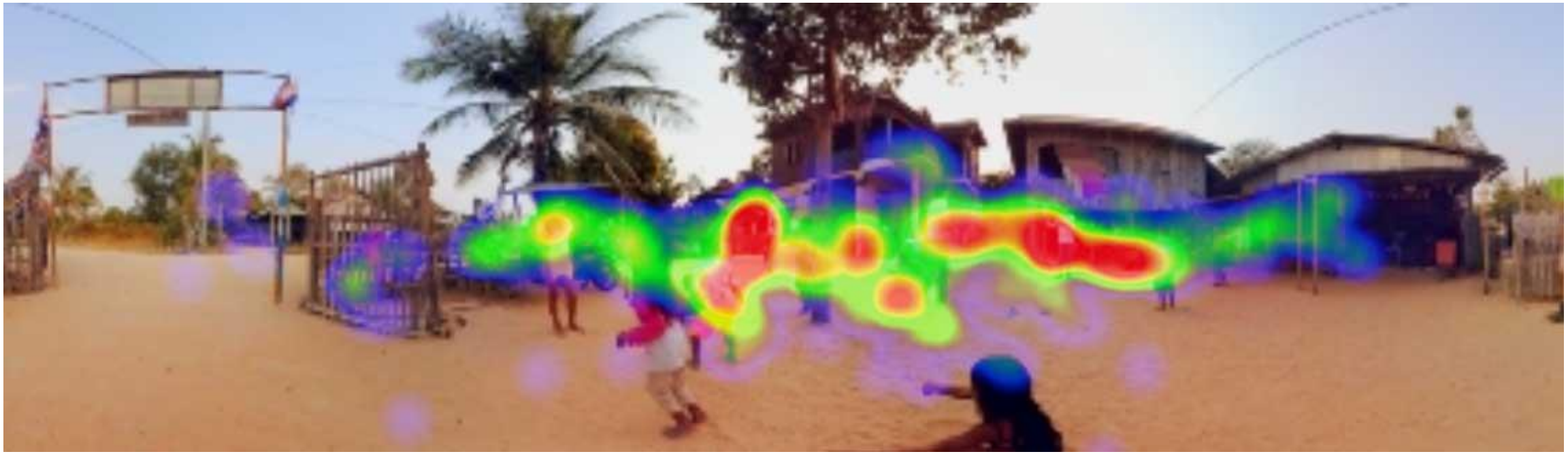


360° Heatmap Case-Study

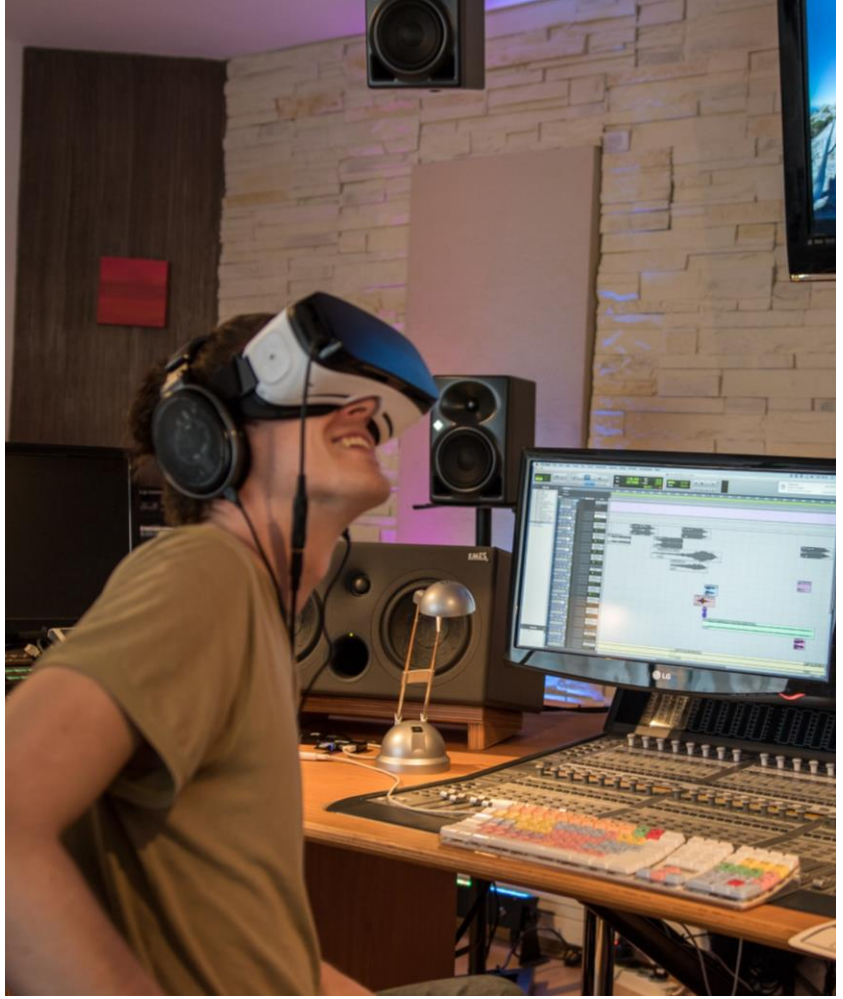
WHAT DIFFERENCES DOES SPATIAL AUDIO MAKE FOR THE VIEWER COMPARED TO A STEREO DOWNMIX?

Hi!

- Martin Rieger, Munich
- Thanks for the invitation!
- Media Technology and Production (M.Eng.), Amberg (OTH-AW)
- Master Thesis 2016: “Sound on Vision of 360° Videos”
- Since then: VRTONUNG – 360° Sound Production

Ostbayerische Technische Hochschule
Amberg-Weiden

VRTONUNG



Field Recording

Post-Production

Talks

Why?

- after ≈20 VR/360° projects in two years – “wait, is it even worth what I’m doing?”
- There are so many things yet unexplored, first steps taken.
- VR/360° MeetUp Munich
- Ph.D. student Sylvia Rothe (LMU Munich) about VR

How?

- Content: 9-minute long episode of 360° documentary “Crossing Borders”
- Create a Unity-App with two versions for GearVR (stereo.wav + spatial.tbe)
- Find about 50 people, split in two groups, to watch it on a swivel chair in calm environment
- make a survey afterwards without telling the „Why?“
- collect data from smartphones and create two heatmaps

Survey

- age, date, gender etc.
- did the subject realize on its own, if the sound was spatial or not?
- how many VR-experiences has the subject already been watching?
- is the subject affiliated with visuals or sound?

Id (Datum/Uhrzeit)	Spatial (y/n)	Kam der Tor	Erstanwender (y/n-wenig/n-	Alter	Geschlecht	kennt Episode oder Cr	Bildmensch, Tonm	Bemerkung
20170412113933	yes	yes	no - viele Anwendungen	25	männlich	CPOC bekannt	Tonmensch	
20170412184308	yes	no	no > 3	25	weiblich	CB bekannt	Sonstiges	
20170412201956	yes	yes	yes	24	weiblich	unbekannt	Sonstiges	
20170413164520	yes	yes	no < 3	65	männlich	CB bekannt	Sonstiges	Hat während de
20170414095150	yes	yes	no < 3	58	weiblich	CB bekannt	Sonstiges	War kurz durch
20170415185933	yes	yes	no - viele Anwendungen	27	männlich	CPOC bekannt	Bildmensch	
20170416095023	yes	yes	no < 3	25	männlich	CB bekannt	Sonstiges	Hat bemerkt, "da
20170417121807	yes	yes	no < 3	28	weiblich	CB bekannt	Sonstiges	
20170419161051	yes	no	no - viele Anwendungen	24	männlich	CPOC bekannt	Bildmensch	
20170419171948	yes	yes	no < 3	28	männlich	CB bekannt	Sonstiges	"Bei den schreie
20170426173209	yes	yes	no < 3	43	männlich	unbekannt	Tonmensch	kann sein dass
20170429173517	yes	yes	no - viele Anwendungen	32	männlich	CPOC bekannt	Bildmensch	Schwein ist witz
20170429174121	yes	yes	no - viele Anwendungen	29	männlich	CPOC bekannt	Bildmensch	hat bemerkt, das
20170512210133	yes	yes	no < 3	25	männlich	CB bekannt	Sonstiges	
20170513195814	yes	yes	yes	23	weiblich	CB bekannt	Sonstiges	Immersion sehr
20170513200828	yes	yes	yes	23	weiblich	CB bekannt	Sonstiges	Zwilling von Pro
20170513203123	yes	no	no < 3	26	männlich	CB bekannt	Sonstiges	! Meinte dass de
20170514183032	yes	yes	no < 3	27	männlich	CB bekannt	Tonmensch	Ist beim Balkon
20170515185948	yes	no	yes	29	männlich	CB bekannt	Tonmensch	
20170520174131	yes	yes	no - viele Anwendungen	38	männlich	unbekannt	Tonmensch	im Stehen
20170520175158	yes	no	no - viele Anwendungen	29	männlich	unbekannt	Bildmensch	im Stehen, hat a
20170520205102	yes	yes	no < 3	27	weiblich	CB bekannt	Sonstiges	im Stehen
20170520210233	yes	yes	no < 3	29	weiblich	CB bekannt	Sonstiges	im Stehen
20170520210326	yes	no	no < 3	34	männlich	CB bekannt	Tonmensch	im Stehen, Ton

Creating the Heatmap

4 SPATIAL ANALYSIS AND STATISTIK

4.1 Space Time Cubes

A method of investigating spatiotemporal data is to analyze Space Time Cubes (STC). In these cubes two coordinates represent the space and the third one the time. With this technique the data can be visualized and explored in a comfortable way (fig.7).

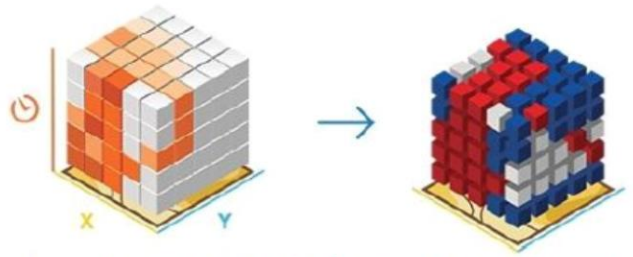


Figure 7: A space time cube has one time and two space coordinates, left the density of a value is shown for every space time section, right the result of a hotspot analysis (significant hotspots are red) [26]

For calculating the STC and applying spatial statistic methods, the GIS software ArcGIS Pro was used. We converted the vtt-file into an Excel-file which was imported into ArcGIS to employ the implemented STC method for analysing the data.

n first step we inspected space time cubes showing the counts idents – in our case "how often" users looked for a special certain area segment, as shown in fig. 7 (left) and fig. 8.

For finding statistical significant hotspots metho statistic can be applicated. These methods take into ac neighbor relations between the space-time segments in the The collected data are point incident data. Point incident d are points connected to an event – in our case the viewer looked to this point. We were interested in significant clusters. To find such clusters, we used the Getis-Ord Gi* statistic [8]. This statistical method requires values for the investigated points. In order to use this method, the incident data were aggregated and incident counts established. The incident counts - in our case the number of views in a segment - are the attribute values which are analyzed by the method.
The Getis-Ord Gi* statistic is given as:

$$G_i^* = \frac{\sum_{j=1}^{n} w_{i,j}x_j - \overline{X}\sum_{j=1}^{n} w_{i,j}}{S\sqrt{\frac{\left[n\sum_{j=1}^{n} w_{i,j}^2 - \left(\sum_{j=1}^{n} w_{i,j}\right)^2\right]}{n-1}}}$$

where x_j is the attribute value for point j, $w_{i,j}$ is the spatial weight between point i and j, n is equal to the total number of points and:

$$\overline{X} = \frac{\sum_{j=1}^{n} x_j}{n}$$

$$S = \sqrt{\frac{\sum_{j=1}^{n} x_j^2}{n} - \left(\overline{X}\right)^2}$$

Using the spatial statistic tools of the ArcGIS Pro Software again, we generated STCs displaying hotspots (fig. 9).

Finally…. The Heatmap

Demo-Videos

Results: Similarities

- humans draw attention.
- non-humans don‘t (buildings, structural objects) and actually guide away from themselves.
- mainly straight forward view
  - Hardly up and down more that than 15° elevation
  - Effortless head movement is around 30° azimuth
- If you show a scene for a second time, people automatically feel comfortable looking away
- some sounds don‘t matter if they are spatial or static

Results: Differences

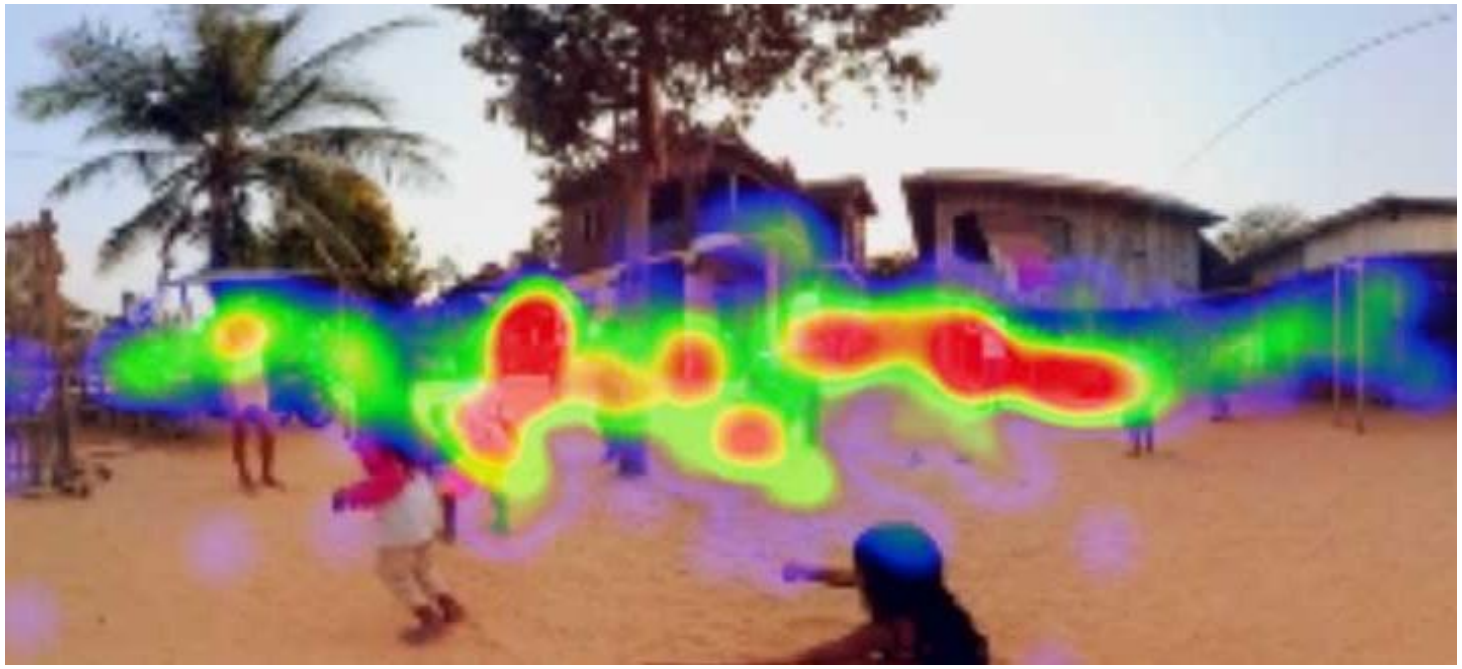
- viewers with static sound turned away from the interviewee after around 8 seconds.
  With spatial sound, people stayed around 1 to 3 seconds longer, mostly until the end of the scene.
- also applicable for details like animals, moving objects (trees, motorbikes etc.) because people wandered more around there.
- spatial sound group found interviewee faster and were confusing it less with a voice-over
- since people weren‘t told about the „why?“, even some sound guys couldn‘t recall if it was spatial or static sound -> small gap between feeling natural and not making any difference.

Learnings

- HMDs overheated & have no eye-tracking
- content
  - shouldn’t be less focused at 0° azimuth.
  - content could be more extreme to force people turning around.
- it gets very difficult to find the causes of the results, roles may play
  - the content
  - HRTFs
  - Subject’s experience with VR
  - Psychoacoustics

Conclusion

- there are differences, yet not as many as expected
- but: a binaural stereo downmix is not the same thing as a traditional stereo mix.
- Early study, so feel free to extend
- immersive audio is nonetheless great and needs more attention in public.
- look-up VR-Sound Blog @ www.vrtonung.de/en



Thanks!

Martin Rieger

info@vrtonung.de

linkedin.com/in/martin-rieger/

VRTONUNG