AUDIO QUALITY ON THE AIR IN DAB DIGITAL RADIO IN NORWAY

SVERRE HOLM

Department of Informatics, University of Oslo, P O Box 1080, NO-0316 Oslo, Norway
`Sverre.Holm (at) ifi.uio.no`

The same program heard on DAB at 128 kbit/s and 192 kbit/s has been compared with a high quality FM stereo signal. The result from informal listening tests is that there are audible impairments in the stereo imaging of high-frequency instruments, and in particular hi-hats at 128 kbit/s. Some examples of songs where this effect can be heard are given. The effect can be explained by spectral analysis of the output of a DAB radio which shows that at 128 kbit/s, 3 kHz is used as the switch frequency for the intensity stereo coding of the MPEG-I layer 2 encoder, effectively removing most signal differences above this frequency. This degradation cannot be heard in a good quality FM stereo signal.

INTRODUCTION

DAB (digital audio broadcasting) radio is marketed as a leap forward in quality compared to today's radio and most people would also associate the term 'digital' with quality. Yet, many countries in Europe seem to be hesitant to promote DAB with the present audio encoder and today only a few countries seem to be pushing it actively as a replacement for FM. Most notably among them are the UK, Denmark and Norway.

The digital transmission of DAB offers many advantages over FM as the problems with multipath interference are virtually eliminated. Multipath reception may give garbled sound and also distort stereo imaging in FM. Therefore DAB often gives a perceived improvement of quality in particular with mobile reception in cars, but also with reception with stationary receivers where valleys and hills create obstacles to reception. The same applies to reception in a city as multipath problems due to surrounding buildings are eliminated. This is important in order for radio as a medium to catch up with its listeners who use their radios in a mobile environment to a much larger extent today than when FM was introduced in the 50's.

It is interesting to read papers that are more than a decade old concerning DAB as they often use most of the space for a discussion of the transmission advantages above, see e.g. [1]. Issues related to audio quality are not discussed much, and often the term 'near-CD' audio quality is used and more or less taken for granted.

Despite this, concerns have been raised over the audio quality of the codec of DAB. Terms such as 'DAB sounds worse than FM', 'old-fashioned codec', and 'low bitrate mp3-sound' are used. This criticism has been concerned with audio as coded with the original MPEG-1, layer II (mp2) codec of DAB. Recently, the aacPlus codec has been standardized for digital broadcasting [2]. It offers a more efficient coding, but this system, called DAB+, is not compatible with existing DAB receivers. At the time of writing, this codec is not yet used for regular broadcasts so it will not be discussed here.

Audio quality comparisons between FM and DAB are complicated by the fact that quality depends on two different dimensions, or varies along two orthogonal axes, in the two systems. FM's audio quality varies mainly with the radio transmission, i.e. with the signal strength, and in particular with the amount of multipath transmission. DAB's audio quality varies with the type of audio signal that is being transmitted, and what bit rate the broadcaster has allocated to that particular station or program, provided that the signal strength is large enough to avoid artefacts due to bit errors.

The bit rate allocation of DAB is one of DAB's strengths, but also an issue that makes audio quality comparisons harder to do. In DAB low bit rates are allocated to speech transmissions in mono and high bit rates are used for stereo transmission of music. Recently, one has also seen mono used for music transmissions in Norway, such as Moox Live at 80 kbit/s, which probably is the first mono music station since FM stereo was introduced in 1975. In the UK there are several such mono music stations. Bit rates also vary dynamically over the day, e.g. to accommodate regional transmissions during certain periods of the day. In the Norwegian DAB network, bit rates at present vary between 32-96 kbit/s for mono transmissions. For stereo transmissions, bit rates of 128, 160 and 192 kbit/s are regularly used.

The purpose of this report is to attempt a quantification of the audio quality of the MPEG-1 layer II coder of DAB as heard over the air. The following assumptions are made:

- Comparisons are made with a near-perfect FM signal, i.e. one that can be received over cable or from a good (outdoor) antenna.
- Audio quality is only discussed for music, not for transmission of speech.
- The focus is on describing the audio quality for the lowest rate presently used for stereo music, 128 kbit/s, as it seems to be used more and more for music transmission.

We believe that it is of interest to the general public to present material on audio quality, not the least because several of the actors in the broadcast community have been very reluctant to give up on the original claims for 'near-CD' audio quality and admit that audio quality now may be considerably lower. Nowadays audio quality seems to be justified by referring to the average listener or targeted at obtaining a certain score in an investigation of user satisfaction.

A demonstration of listener's reactions to audio quality was when BBC last year reduced Radio 3 from 192 kbit/s to 160 kbit/s, and then received so many reactions from their listeners that they had to increase the bit rate again. Thus it is evident that despite refinements that have taken place in encoders, it is possible to hear imperfections in critical audio material even at 160 kbit/s [3] with the latest generation of encoders.

Therefore the emphasis in this paper is on audio quality comparison with the present FM stereo system. In addition to CD-sound, this is the main reference that a listener has. This may be more meaningful than relative comparisons between various codecs or implementations of codecs. But with all such testing that has taken place (e.g. [4] and [5]), the results of the present papers should not come as a surprise to the broadcasting community.

This paper is structured in the following way. First we give a discussion of subjective quality testing as found in the literature. Then some results from informal listening to music sent over DAB are given. The emphasis is on a comparison between 128 kbit/s, 192 kbit/s, and FM. This has been possible since NRK during certain periods of the day at present transmits the most popular channel in Norway, P1, in both multiplexes and at two different bit rates.

Third, some measurements of spectra of the mono component and the stereo component from the output of a DAB receiver are given. They are interpreted in light of the codec specifications and used to explain what was heard.

1 SUBJECTIVE AUDIO QUALITY TESTING

Formal audio quality testing of DAB encoders has been standardized in ITU standard BS.1116 (Methods for the subjective assessment of small impairments in audio systems including multichannel sound systems, last revised 10/1997) which *"… provides details regarding the selection of critical audio test materials, the acoustical characteristics of the listening environment, the performance of the playback system, and the training of listeners."*[6]. A related test was proposed in [6]. It is geared towards very low bitrate audio codecs and corresponding lower quality. It has the same requirement for selection of critical audio material.

There is not so much available material in the open literature on the results of these tests and most of it seems to be dated. The latest such paper that we have been able to find is from 1998 [4]. There is also a BBC Research & Development white paper which was published in 2003, but originally written in 1994, which recommends bit rates in the 192 – 256 kbit/s range and said that at 192 kbit/s, it is relatively easy to hear imperfections in critical audio material [7].

Newer material does not seem to exist outside of internal reports in the broadcasting industry and in the companies that make the coders, but this material is not publicly available. One such example is found in [5] from 2004 which documents an increase in the quality of the encoders. But the tests are not directly comparable [8] so we have not been able to find any measure of how large the improvement has been since the mid 90's. The audio material used in the test of [4] consisted of audio material that was designed to reveal weaknesses in the encoders, such as pure tones (pitch pipes), while the audio material of [5] consisted of audio samples that also were typical for broadcasting including male and female voices. Voice signals are particularly simple to encode. Still the test in 2004 is claimed to conform to BS.1116. This could indicate that the standard is much less restrictive with respect to selection of audio material than the above quote indicates.

It also seems to go against [9] which warns against changing the audio material: *"It would be futile to do subjective tests with other than critical materials. Picking, say, "average" materials, however this might be defined, would simply mean that differences are unlikely to be found in the subsequent formal tests, and* *it might then be falsely concluded that rival codecs are fully equivalent."*

It is rather surprising that the broadcast community now seems to be much more relaxed with respect to the requirement for using critical material. One cannot but speculate on why, but it certainly makes the latest subjective scores look better. It also makes it look as if there has been a larger quality improvement in the latest generation of encoders than there really is.

2 INFORMAL LISTENING TESTS

It was beyond the scope of this rather limited study to perform any formal, standardized tests. Therefore, this study started with informal listening sessions in the fall of 2006 where DAB at 128 and 192 kbit/s was compared to FM. The equipment used was a Matsui DA-1 tuner (dynamic range compression off), Onkyo TX-822 FM receiver, and Pluto active quasi omni-directional loudspeakers designed by S. Linkwitz (www.linkwitzlab.com). The tests took place in a living room where the loudspeakers and the listener were located for optimum stereo listening. The FM signal was received over cable and should presumably be free of multipath distortion. Care was taken to normalize the audio levels between FM and DAB.

Our first panel consisted of two people and it was discovered that something changed from time to time in the stereo image when listening to low-rate DAB compared to high-rate DAB and FM. We then did some measurements directly on the output of the DAB tuner (see next section). Informed by these measurements, we were able to pin-point the effect more exactly. Our subjects in the second panel (tests done in the spring of 2007) were therefore asked to listen for changes in the spatial image at high frequencies and in particular to focus their attention on the hi-hats and cymbals. These tests took place over several days and the panel consisted of 4-5 people.

2.1 Stereo Width

The effects heard were that if the hi-hat or cymbals originally were placed to the right or left in the mix (as heard on FM or DAB at 192 kbit/s), DAB at 128 kbit/s would consistently shift their location towards the centre. This can be heard on a number of songs and with only a short training period, our subjects could almost instantly detect whether they were listening to DAB at 128 kbit/s or FM based on this effect. Some tracks where this effect is audible are:

- Aha, The sun always shines on TV (Hunting High and Low, Warner Bros, 1990)
- Cardigans, Erase/Rewind (Gran Turismo, Island / Mercury, 1998)
- Salvador Santana, Evil Ways (Song for My Father, Target Spotlight Series, 2007)
- Louise Goffin, Up on the Roof (Song for My Father, Target Spotlight Series, 2007)

In one song in particular, this effect was more prominent than in any of the others. This was a modern jazz track:

- Silje Nergaard, When Judy Falls (Darkness Out of Blue, Universal, 2007).

Here the sound from the drum section including the hi-hats forms a wide, spacious sound stage that surrounds the other instruments and the vocals. The ambient sound stage from the hi-hat was reduced to a single centred source when sent over DAB at 128 kbit/s.

2.2 Stereo Depth

In our first session, it was heard from time to time that in particular for female vocals, a loss of depth perspective was detectable. In the second session, our most trained listener, a professional sound engineer, remarked that on a track such as Amy Winehouse, You know I'm no good (Island, 2007) such a loss of depth in the stereo image could be heard. This means that on the FM or high-rate DAB version a sense of distance between the singer and the instruments could be heard, while on the 128 kbit/s DAB version all sound seemed to originate on a straight line between the loudspeakers. This is a very subtle effect which seems to be hard to detect unless one is trained to listen for it.

2.3 Treble

All listeners, regardless of age, could instantly hear that there is a loss of clarity due to a lack of treble on a station such as Moox Live which is transmitted at 80 kbit/s. This station of course has neither stereo width nor depth since it is sent in mono. The veiled sound could be heard without any reference such as FM or high-rate DAB. From the measurements we found that the cut-off frequency is about 11 kHz.

The two youngest listeners, males in the early twenties, could also almost instantly detect a certain veiling of the treble when listening to either DAB at 128 kbit/s or FM compared to DAB at 192 kbit/s. This was not so evident for older listeners, even those in the mid and late twenties. The difference in cut-off frequency turns out to be 14.25 kHz vs. up to 18 kHz. It was hard for the younger listeners to hear much difference in treble between FM and low-rate DAB.

3 MEASUREMENTS

Due to the restriction of only having the output of the DAB radio available, it is quite limited what can be measured for characterizing the DAB sound. We have however found that the sum and difference of the left-hand and right-hand channels is what gives the most meaningful information, i.e. the mono or mid signal M=L+R and the stereo or side signal S=L-R.

The MPEG-1 layer II encoder samples the signal at either 48 or 24 kHz [10]. Half of this band is then divided into 32 sub-bands where bits are allocated according to perception criteria. At 48 kHz, each band will have a width of 24 kHz/32 = 750 Hz. The mid-signal's upper frequency will therefore be a multiple of 750 Hz. This value can be found by spectral analysis.

The joint stereo coding which is used at the lowest bit rates uses intensity stereo coding from either band 4, 8, 12 or 16 and up, i.e. from a switch-frequency of 3, 6, 9 or 12 kHz and up. For all sub-bands above this frequency only a common signal is transmitted instead of individual signals for the right-hand and left-hand channels. In addition pan-information is sent per band. In this way transmission capacity is saved, but information about signal differences is lost above the switch-frequency. Spectral analysis of the side-signal will reveal the switch frequency and thus in which mode the intensity stereo coding is operating.

The following measurements were done from the outputs of the Matsui DA-1 DAB tuner, and then repeated for a high-end Cambridge Audio Azur 640T DAB/FM Tuner in order to make sure that we measured the characteristics of the DAB encoding and not any oddities of the tuner.

The spectral analysis was first performed with an 8-bit sampling oscilloscope, Agilent DSO6014A, using the built-in FFT function and with the persistence function on. The analysis is performed over 20-60 seconds of music in order to get a good visualization of the cut-off frequencies. Due to the 8 bit A/D-converter there is only about 50 dB dynamic range in the plots which creates a high noise floor at high frequencies. Therefore, the analysis has also been performed on data sampled with a high-resolution audio interface (E-Mu 1616 semi-professional external sound card sampling at 48 kHz) and analyzed with Adobe Audition ver. 1.5. The spectrum was found from 16384 samples, i.e. it covers 16384/48000 or about 1/3 sec.

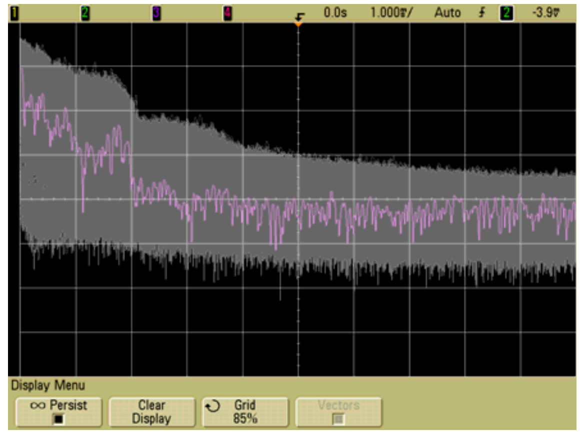

Figure 1: Long-term spectral analysis of mono-signal from Moox Live (80 kbit/s). x: 5 kHz/div., y: 20 dB/div.

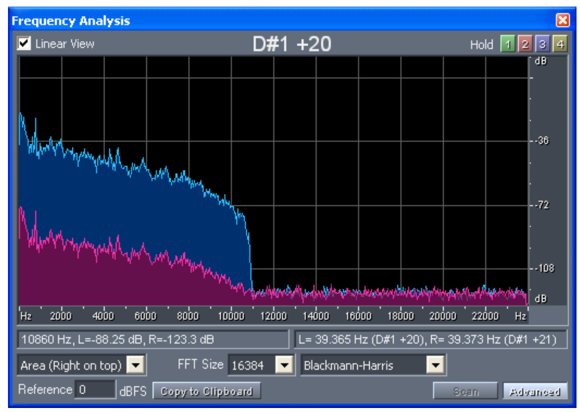

Figure 2: Short-term spectral analysis of mono-signal (blue, back) and stereo-signal (violet, front) from Moox Live (80 kbit/s, mono) using Adobe Audition.

Figs. 1 and 2 shows that Moox Live at 80 kbit/s mono, has an upper cut-off frequency between 10 and 11 kHz, and our estimate is that the sampling rate is 24 kHz and that maximum bands 0-28 have been included, i.e. upper cut-off is 10.875 kHz (varies with time).

Figs. 3 and 4 show NRK P1 at 128 kbit/s. The plots are typical for this bit rate when music which 'fills' the spectrum well is sent. It shows that there is a drop in level in the stereo-signal at 3 kHz of 20 dB or more. This shows that intensity stereo coding with the lowest switch-frequency is used. From Fig. 4 the estimated upper cut-off based on transmission of bands 0-18 is 14.25 kHz.

The same analysis of an FM signal would show almost overlapping spectra between the mono and the stereo-component. It is similar to Fig. 4 below 3 kHz, and has the same behavior all the way up to the treble cut-off around 15 kHz.

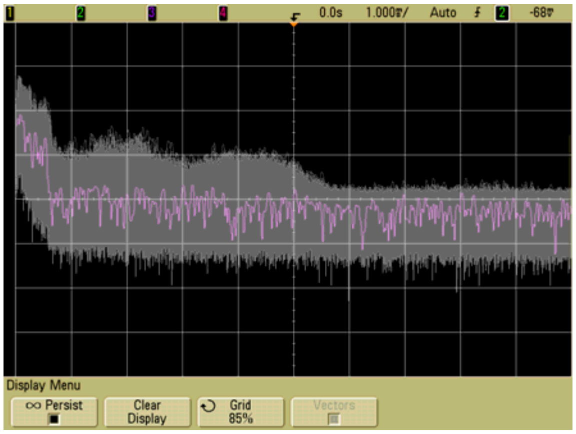

Figure 3: Long-term spectral analysis of stereo-signal from NRK P1 (128 kbit/s). x: 5 kHz/div., y: 20 dB/div.

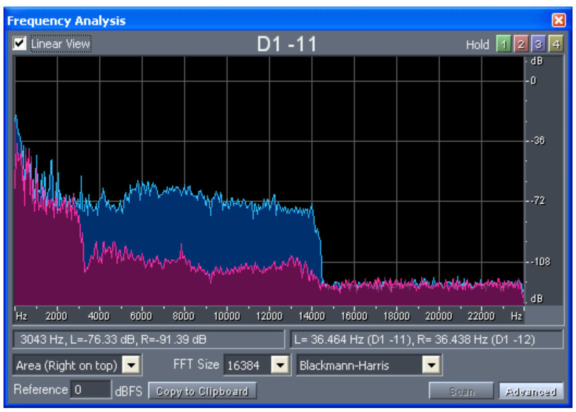

Figure 4: Short-term spectral analysis of mono-signal (blue, back) and stereo-signal (violet, front) from NRK P1 (128 kbit/s) using Adobe Audition.

We have not done a very comprehensive analysis of the switch frequency for the intensity stereo coding at the higher rates, as it varies more with the audio material. It seems often to vary between 3 and 6 kHz at 160 kbit/s, but it can also be higher. At 192 kbit/s we have found all values from 6 kHz and up in use.

The upper limit at the higher rates also varies with the contents, and we have found that the upper cut-off frequencies may be up to 18 kHz at 192 kbps and 15.75 kHz at 160 kbps.

4 DISCUSSION

4.1 Listening tests and measurements

NRK has confirmed that they are using the latest generation of encoders which came out best in the test in 2004 [5]. The results we found should therefore be typical for the state-of-the art today. We have no information about what the other content providers in the Norwegian multiplexes (P4 and Moox) are using.

The collapse of the stereo image for the hi-hat for the Silje Nergaard recording was so dramatic that contact was made with the sound engineer for that recording in order to investigate why this happened.

It turned out that the drum set was recorded with an overhead A-B stereo pair of microphones spaced 20-25 cm apart and hanging about 80 cm above the set. In addition there were individual close-in microphones on each drum including the hi-hats. The stereo pair was represented quite strongly in the final mix, in addition to the close-in microphones. But, the close-in microphones were a bit stronger on the particular song that we listened to compared to the other tracks on the CD as it is more pop-like than the rest [11].

The explanation for the changed stereo image is therefore the presence of the stereo pair in the mix. The intensity stereo coding cannot represent the signal differences above 3 kHz and therefore the track will sound as if recorded using panned microphones only. It is likely that this effect is present for all tracks for this particular CD, and it may actually be stronger on some of the other tracks. This is typical for many acoustic productions such as jazz and folk music and even some pop music. To be fair, it should be said that although this track was heard on NRK P1 at 128 kbit/s, NRK P2 (Culture channel) which probably carries more acoustic music than P1, transmits at a bit rate of 160 kbit/s.

The centering of the hi-hats for the other tracks, which are more typical pop and rock music recordings, is also an effect created by the stereo intensity coding. It gives audible differences compared to the corresponding FM stereo signal.

When it comes to the upper limit for treble, our rather limited testing seems to indicate that DAB at 192 kbit/s gives an audible increase in clarity for listeners below a certain age, which we estimate at being somewhere in the twenties for males. It could be higher for female listeners. Evidently the analogue, gradual cut-off of FM around 15 kHz sounds more or less like the digital steep cut-off at 14.25 kHz of DAB at 128 kbit/s in comparison to the 18 kHz of DAB at 192 kbit/s.

4.2 Why is audio quality not better?

Why are there audible impairments in low bit rate DAB compared to FM? In our original report [12], it was explained by the discrepancy between the present bit rates for music, 128-192 kbit/s, compared to the ones which were originally recommended, 192-256 kbit/s. Since then some of the results from testing in later years have been made available such as [5], so it is evident that improvement in encoders has taken place justifying a lower bit rate today than in the mid-90's. But as old and new test results are not directly comparable, it is hard to quantify exactly how large the improvement may be. If indeed 128 kbit/s today is equivalent in terms of audio quality to 192 kbit/s more than a decade ago as some would say, then perhaps DAB was designed with lower audio quality than FM from the very start.

I find that hard to believe, so it is probably a better bet to say that the use of 128 kbit/s today for music transmission on DAB can be explained partly in terms of improved encoders, and partly due to the desire to include more stations in the multiplexes.

5 CONCLUSIONS

Informal listening tests of 128 kbit/s DAB give as a result that DAB changes the stereo image. High-frequency instruments and in particular hi-hats that in the original mix are located to the side in the stereo image will consistently be shifted towards the centre. This effect is particularly easy to notice if the drum set has been recorded with a stereo pair of microphones in order to create ambience. In this case it will change to one where the hi-hats seem to be recorded with a single microphone mixed to the centre. This effect is not present in FM stereo and has not been heard on DAB at 192 kbit/s.

Spectral analysis of the mono (mid) and stereo (side) signal has been used to find the upper cut-off frequency and the switch frequency in the intensity stereo coding of the MPEG-I layer 2 encoder. Music at 128 kbit/s will most of the time use the lowest switch frequency at 3 kHz. This removes signal differences above this frequency and explains why high-frequency instruments lose stereo imaging.

The listening tests have only been done on 128 and 192 kbit/s DAB. It was beyond the scope of this investigation to test 160 kbit/s. Based on that the conclusion is that DAB at 128 kbit/s has audible impairments that are not present in a good quality FM stereo signal.

ACKNOWLEDGEMENT

I would like thank Rune Skramstad, Paragon Arrays, for taking part in the informal listening sessions, and Prof. Morten Dæhlen and Steve Green, MSc, for valuable comments.

REFERENCES

[1] F. Kozamernik, Digital Audio Broadcasting – radio now and for the future, EBU Technical Review Autumn 1995, pp. 2-27.

[2] ETSI TS 102 563 V1.1.1 (2007-02) European Telecommunications Standards Institute (ETSI), Technical Standard, Digital Audio Broadcasting (DAB); Transport of Advanced Audio Coding (AAC) audio.

[3] B. Aarseth, technical advisor to NRK in an interview on NRK P2 Kurer, 17. Feb. 2007.

[4] G. A. Soulodre, T. Grusec, M. Lavoie, L.Dear Profs. Ed Thibault, "Subjective evaluation of state-of-the-art two-channel audio codecs," J. Audio Eng. Soc. 1998, vol. 46, no3, pp. 164-177.

[5] G. A. Soulodre, "Subjective Evaluation of Stereo Audio Coders for Digital Radio" presentation at the WorldDAB General Assembly Barcelona, Oct. 2004.

[6] G. Soulodre, M. Lavoie, "Subjective evaluation of large and small impairments in audio codecs," Audio Eng. Soc. Int. Conf. 17, Italy, 1999, pp. 329-336

[7] C. Gandy, "DAB: an introduction to the Eureka DAB System and a guide to how it works," BBC Research & Development White Paper, WHP 061, June 2003.

[8] Personal communications, G. A. Soulodre, Camden Labs, Canada, March 2007.

[9] T Grusec, L Thibault, G Soulodre, "Subjective Evaluation of High-Quality Audio Coding Systems: Methods and Results in the Two-Channel Case," Proc. 99th Int. Conf. Audio Eng. Soc, Sept. 1995.

[10] ETSI EN 300 401 V1.4.1 (2006-01), European Standard (Telecommunications series), Radio Broadcasting Systems; Digital Audio Broadcasting (DAB) to mobile, portable and fixed receive.

[11] Personal communications, A. Suurna, All Arts Recording AB, Stockholm, Sweden, April 2007.

[12] S. Holm, Lydkvalitet i DAB digitalradio (Audio quality in DAB digital radio), Research report 350, Department of Informatics, University of Oslo, 30. January 2007.

CORRECTION

The phrase "phase information" has been replaced with "signal differences" in four places: Abstract, 3 Measurements, 4 Discussion, and 5 Conclusion.

It was pointed out after the presentation of this work, that the ear is not so sensitive to phase information above a certain frequency and cannot use it for determining direction. This is the basis for the stereo intensity coding model.

However, the high frequency ambience which stereo intensity coding has problems in reproducing is due to something which is stronger, i.e. an actual difference between the signals of the left and right channels.