NOKIA

Rich Recording Technology
Technical overall description

Ari Koski
Nokia with Windows Phones
Product Engineering/Technology
Multimedia/Audio/Audio technology management

Nokia's Rich Recording technology summary

Nokia's rich recording technology in Nokia mobile phones records true stereo audio with video using microphones integrated into mobile phone device. It provides full audio bandwidth to match with human hearing capabilities (20 Hz ... 20 kHz) and extended dynamic range (107 dB) by tolerating high sound pressure levels with no perceptual distortion compared to commonly used digital microphones that typically operate on 87 dB dynamic range.

Nokia SW algorithms improve the capture results by providing means for wind and camera noise reduction and automatic level control.

First product that utilizes Nokia's rich recording technology is Nokia 808 PureView.

Capture challenges

A large difference between the quietest sounds and the loudest sounds (i.e. the dynamic range) to be captured sets a challenge to the small microphones used in mobile phones. As the diaphragm of the microphone is small, it needs a lot of amplification in order to be able to capture the quiet sounds. However, the high amplification in such recording events, where high average sound pressure levels are often exceeded with much higher amplitude transient sounds, will cause the recorded signal to be distorted.
A good example of such a recording event is a band practice.

In addition to large dynamic range to be captured, there are various noise sources that impact to the quality of the recordings. Wind noise is tubular and can easily mask the captured audio partially or totally. Disturbing handling noise by the user using the phone can be captured as well as potential noise coming from auto-focus function of the camera (mechanical sound arising when the objective of the camera is moved). User can also accidentally block one of the microphones, especially if their locations in the phone are not carefully considered.

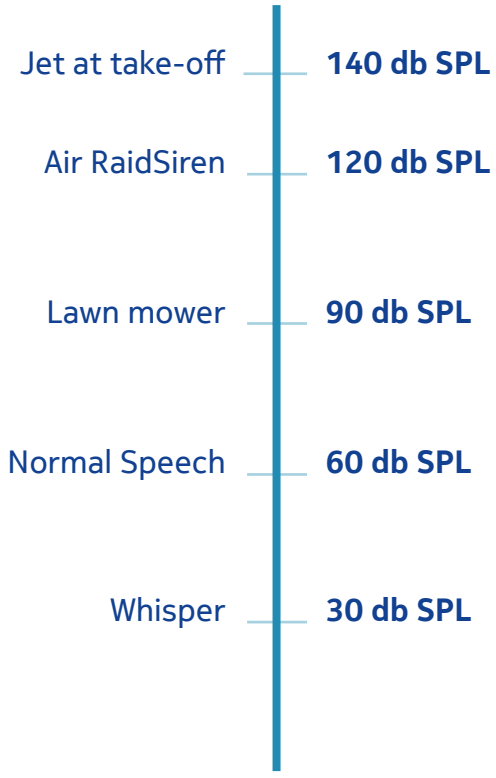

Figure 1. Sound pressure levels vs. typical sound sources. Dynamic range for Nokia Rich Recording technology is 107 dB (33 dBSPL to 140 dBSPL).

Nokia Rich Recording technology building blocks

User starts video recording from the phone's media recorder application (set into video mode) by pressing the camera key.

Audio capture consist of microphones picking up the sound, audio processing enhancing the sound, and audio encoding compressing the audio stream. Video capture comprises camera module with camera processing and video encoding. High quality audio recording is ensured with stereo AAC encoder at 48 kHz sampling rate with 192 kbit/s bitrate and high quality video recording using AVC (MPEG-4 Part10 AVC) encoder at 30 frames per second with 1980x1080p resolution and 20 MBit/s bitrate. Audio and video streams are combined into one file container (.mp4) and stored into a file.

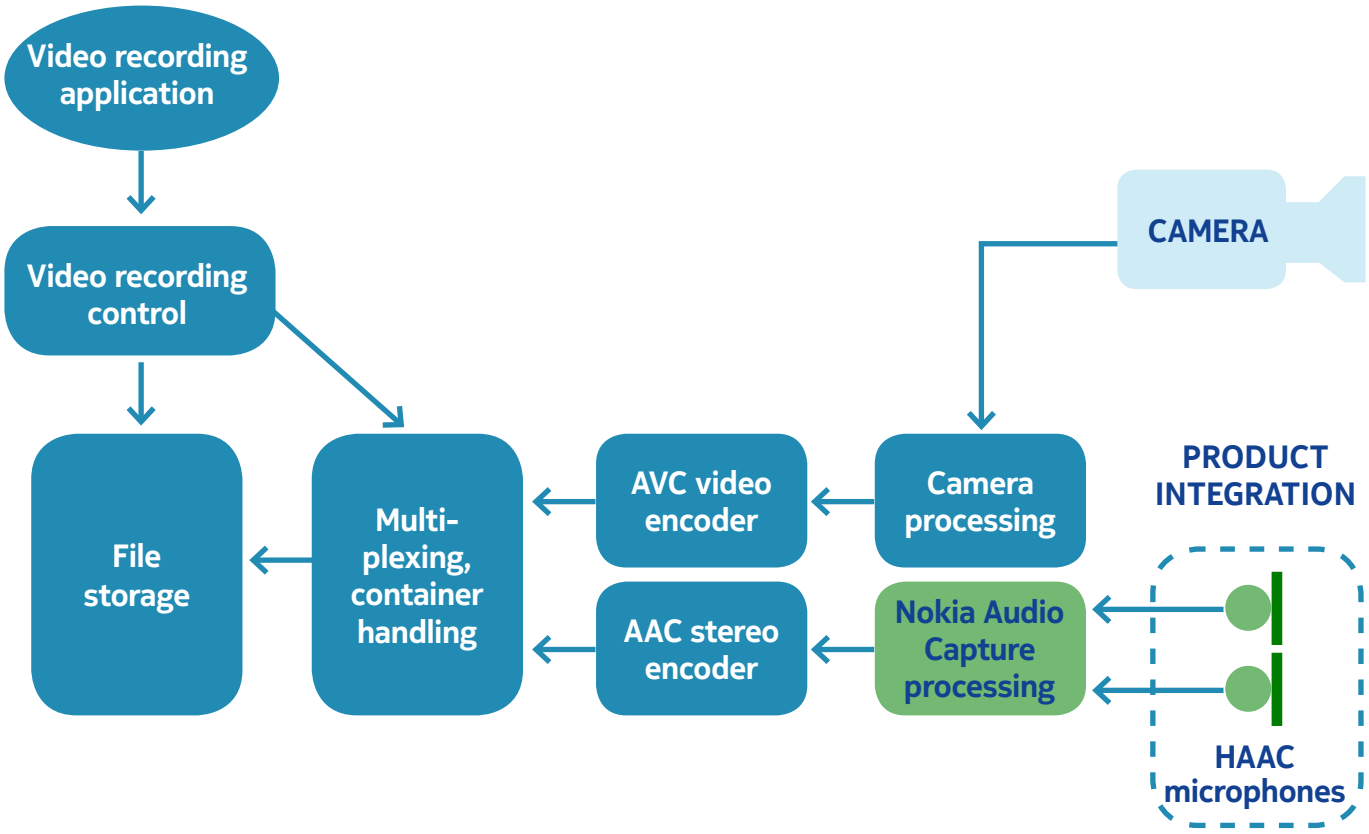

Figure 2. The building blocks for Nokia's Rich Recording are drawn with green objects.

High Amplitude Audio Capture microphone Solution

Conventional digital microphones in mobile phones start to introduce distortion and clipping when sound pressure level reaches ca.120 dBSPL (Sound Pressure Level). For example loud bass drum transients may cause such clipping. A common way to avoid this type of problem is to apply a high-pass filter, but it cuts of the low frequencies making the recorded sound thin.

To cope with high sound pressure levels, Nokia has developed together with microphone suppliers a new microphone concept (High Amplitude Audio Capture) that can capture high sound pressure levels with perceptually no distortion to the audio. It is capable of recording up to 140 dB and higher Sound Pressure Levels, while that commonly used digital microphones can only reach 120 dBSPL without heavy distortion. As high amplitude audio levels are tolerated, there is no need to limit the frequency response. Hence full audio bandwidth to match to human hearing capabilities can be provided when capturing with HAAC microphones.

The basic functionality of the HAAC microphone can be seen in Figure 3.

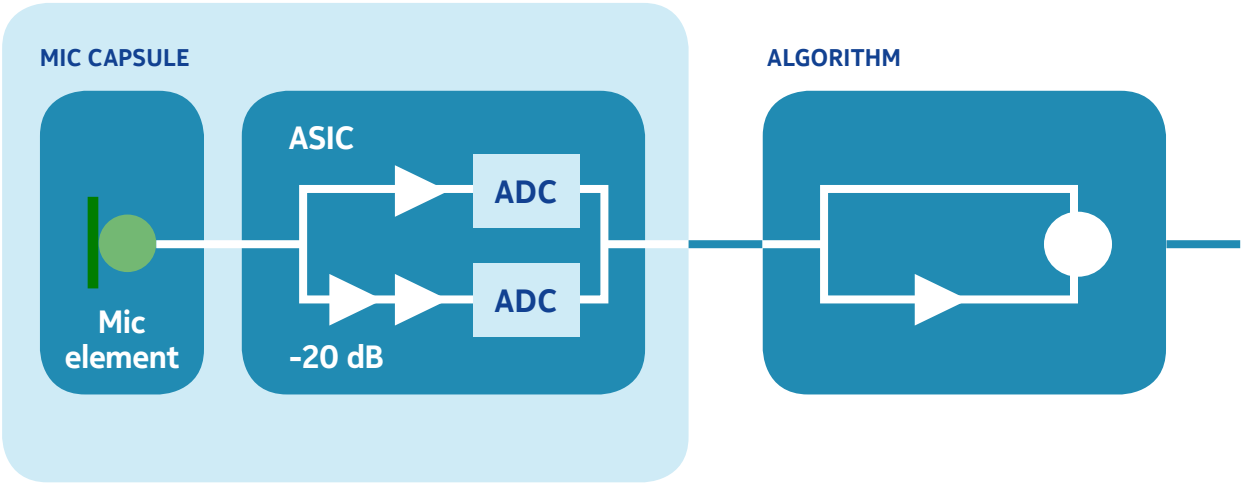

Figure 3. Basic functionality of the High Amplitude Audio Capture microphone.

HAAC microphone has one diaphragm, but it has two internal pre-amplifiers and ADCs in the ASIC inside the microphone. The normal microphone path saturates with high sound pressure levels (> 120 dB SPL). The other microphone path has 20 dB attenuation before the pre-amp. As being 20 dB less sensitive, this path can tolerate sound pressure levels of up to140 dB SPL without perceived distortion.

Outputs from the amplifiers are delivered via the digi-mic interface to the host processor in mobile phone. A SW algorithm inside the mobile phone selects the output path based on the sound pressure level and ensures a smooth transition between the signals. The two signals received by the phone SW are combined so that the sound input signals under 120 dBSPL are used from the 120 dB path and the inputs exceeding 120 dBSPL (and thus distorting in the 120 dB path) are used from the 140 dB path. The switching between paths is done so quickly that e.g. in a case of a pumping drum bass beat, when the bass drum exceeds 120 dB only those fractions of a second are taken from the 140 channel and the rest of the time the 120 dB path is used.

SW algorithms

Nokia has developed SW algorithms package to improve audio recording quality. It provides wind noise reduction, microphone self-noise reduction, microphone equalizers (to fit product acoustics to frequency response requirements), camera autofocus noise reduction and automatic gain control. Part of the SW package is also the algorithm needed by the HAAC microphone.

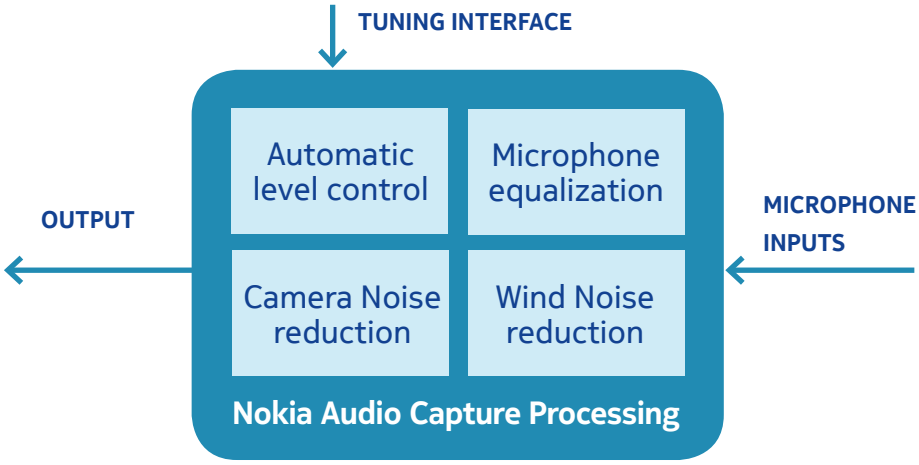

Figure 4.Nokia audio capture processing functionalities.

Audio capture SW processes microphone signals using 32-bit arithmetic preserving the high dynamic range. However, as the captured dynamic range is higher than what is meaningful to re-create in the playback (especially when the playback occurs into a normal living room), the automatic gain control compresses the dynamic range to fit the playback dynamic range into 16-bit arithmetic, corresponding to 96 dB dynamic range, which is more suitable for a listening room. Still, the Nokia Rich Recording technology ensures that the loudest sounds are preserved with no perceptual distortion and that the silent sounds are well audible.

Product integration

Location of the microphones in the product has a significant impact to the achievable stereo image and susceptibility to the handling noise. The recommended location for high quality stereo recording is to locate the microphones as far as possible from each other into the opposite phone ends. This assumes that phone is held as landscape position during video shooting. Symmetrical location, the best one being in the middle of the phone ends, is preferred to avoid handiness or that a hand would be blocking one of the microphones.

Short and straightforward acoustic portings are be provided in order to create a frequency response that provides the full audio bandwidth with a simple equalization scheme. Proper microphone sealing from mechanically vibrating sources and audio sources within the mechanics prevent unnecessary sound leakages to enter into the microphones and to reduce disturbances caused by handling noise.

Nokia SW package for audio capture supports tuning interface and tools for easy tuning of the microphone frequency response in the product acoustics to match the defined performance targets.

Comparison to other technologies

Due to the extended dynamic range of the used new HAAC (High Amplitude Audio Capture) microphone technology, it is possible to record sound pressure levels up to 140 dBSPL with no perceptual distortion. This means Nokia Rich Recording Technology offers recording of four times louder sound pressure levels than with traditionally used digital microphones.

As there is no need to modify the frequency response on purpose for distortion prevention, Nokia Rich Recording Technology can provide high audio quality stereo capture over the whole audible frequency range.