Airborne ultrasound data communications: The core of an indoor positioning system

Sverre Holm
Department of Informatics, University of Oslo
P O Box 1080, Blindern, NO-0316 Oslo, Norway
Sverre.Holm@ifi.uio.no

Abstract — This paper gives an analysis of an ultrasonic indoor data communications system. It may be the core of an indoor positioning system which can rely on ultrasound alone, as no radio, no infrared channel nor any tether is required. The main hardware components are a tag (transmitter) and a detector (receiver). The detector uses digital signal processing to cope with the acoustic environment and its noise, reverberations and Doppler shift.

The attainable range is 10-20 meters and by making a comparison with the range for speech and for the Silbo whistling language, we find that the range can be predicted well. The channel efficiency of the system is found to approach 0.025 bits per second per Hz. This is somewhat less than speech which also has to deal with a similar environment. This comparison is done by using the Shannon channel capacity theorem. It indicates that there is a potential for improvement which will impact capacity in a positioning system as the channel is time-shared among the transmitters. The output energy of the system is compared to currently accepted exposure limits and found to be safe.

I. INTRODUCTION

Ultrasound communications is usually considered to be inferior to radio communications due to its short range and low data rate. Despite this, in nature bats are highly successful in utilizing ultrasound for ranging and navigation. Ultrasound with frequencies just above the audible range also shares most properties with human speech, the communications system which by far has the largest 'installed base' of all. This paper describes an airborne ultrasonic communications system in the 40 kHz range that was designed as the core of an indoor positioning system.

Indoor positioning is done by the principle of confinement, not by triangulation based on time-of-arrival or time-difference-of-arrival as in many combined RF/ultrasound systems [1], [2], [3]. Confinement positioning uses the insulating properties of walls, in the sense that if an ultrasonic tag is detected it has to be localized in that room. This interfaces well with humans as we find it more natural to position objects indoors by reference to rooms rather than x, y, z-coordinates. Besides, the positioning is virtually error free unlike RF systems which sometimes can estimate a tag's position to be in an adjacent room or floor. This is especially true for the important class of RF systems that use signal strength for positioning (aka finger printing) [4].

There are many similarities with underwater acoustic communications systems [5], [6]. Unlike in underwater acoustics, no analysis of the maximum possible range of airborne ultrasound systems exists, and neither does an analysis of the bit rate when background noise level, Doppler shifts, and reverberations are considered. This paper presents such results by comparison to the properties of speech communications.

The paper starts with the major challenges in designing a reliable ultrasonic communications system. The core of the paper is a discussion of attainable bit rates by comparison with the Shannon channel capacity theorem and a discussion of maximum range by comparison with spoken speech and the Silbo whistling language. The purpose of this analysis is to determine the potential for further refinement.

II. SYSTEM DESCRIPTION

The discussion in this paper is based on the development of a set of ultrasonic tags (transmitters) and detectors (receivers) that have been used for equipment and asset tracking. The hardware and software components of this system have been described elsewhere along with a preliminary performance analysis [7] which this paper will expand upon.

The system consists of battery-operated tags to be attached to equipment or be worn by humans. They will transmit their identity when triggered by a movement sensor or a timer. The ultrasound signals are received by detectors that consist of an ultrasonic microphone, A/D, and a digital signal processor. The output from the detector is sent to a central server in the form of an event consisting of the tag's ID, the detector's ID (equivalent to room number), time, and status information.

In an installation at a major hospital tens of detectors and hundreds of tags attached to various equipment, were used. This hospital had a central pool of equipment and the goal of the project was to provide a tool for locating the equipment throughout the hospital.

The experience after five months of use indicated that there was a substantial gain in effectiveness through saved time for searching as equipment could be located simply by consulting a user terminal. There was also a potential for reduced operational and investment costs as the equipment pool could be more precisely sized according to the documented usage. Finally such a system could improve quality assurance through better availability of vital equipment at the right time and place.

III. CHALLENGES IN ULTRASONIC COMMUNICATIONS

A. Doppler shift

The propagation speed in air of 340 m/s is almost 1:1,000,000 compared to radio. This implies that even a small velocity such as walking at 6 km/hr, will generate a relative Doppler shift of about two orders of magnitude greater than a radio system transmitting from a fast-moving spacecraft.

B. Reverberation

Reverberation is the lingering of sound in a room once the source of the sound has ceased producing. In building acoustics, the reverberation time is defined at the -60 dB point and typical values are in the range 0.3 – 3 seconds. In a communications system, -60 dB is a very conservative criterion for avoiding intersymbol interference. Typical design values are 50 – 300 ms. The larger value is valid for communications in e.g. a long corridor with concrete walls, floor, and ceiling. This ensures that the range will be limited by the background noise and not by the reverberations.

C. Noise

At ultrasound frequencies, the only background noise measurement that we are aware of is that by Bass and Bolen from 1985 [8]. They measured a level of 70–80 dB SPL (Sound Pressure Level in dB relative to 20 μPa) in the range 20–60 kHz in an industrial environment (3 kHz bandwidth), with grinding producing a level of 80 dB. Air tools are even worse and may produce levels up to 100 dB SPL 1.2 m from the source. We will use the 75 dB level later for comparison and the equivalent spectral density which is 75 dB-10 log(3000) = 40.2 dB/Hz.

The large variation in background noise level (about 50 dB) is one of the factors that distinguish ultrasonic systems from radio communications systems.

IV. COMPARISON WITH SPEECH COMMUNICATIONS

Speech is a communication form which operates in the same medium, therefore it is of interest to compare the two both with respect to achievable data rate and range. This will give an indication of the improvement potential.

A. Data rate

A typical data rate for raw speech is 50 bits/sec, and with prosodic information (e.g. intonation and mood) it increases to about 200 bits/sec [9]. Communications quality speech can be transferred in a bandwidth of about 300 – 2300 Hz. Therefore the efficiency in terms of bit rate per bandwidth is in the order of 0.025 – 0.1 bits/s/Hz. The efficiency, C/W, is a measure of how well the channel is utilized according to Shannon's information capacity theorem:

$$C = W \log_2 (1 + SNR) \tag{1}$$

Here C is the rate in bits/s, W is the bandwidth in Hz, and SNR is the signal to noise ratio. Speech communications is a means of communications that utilizes the channel much less efficiently than typical electromagnetic communications systems. This is due to the low quality of the acoustic channel and the redundancy built into speech that makes it robust to movement, reverberations and noise.

The expected data rate of an airborne ultrasonic communications system will now be estimated. We will assume that a transducer with f_0=40 kHz is used. Typical relative bandwidth (-6 dB) is 10% or W=4 kHz. A simple, but robust, modulation such as frequency shift keying (FSK) may be used, and the system should be designed to have a large tolerance to Doppler shift. If the maximum velocity is +/-v, the maximum Doppler shift is +/-f_0v/c and the minimum distance between each transmitted frequency is $2f_0v/c$. Thus the number of frequency pairs that can be used for FSK is:

$$M = \frac{W}{4 f_0 v / c} \tag{2}$$

Using c=340 m/s and v=6 km/h one can use only M = 5 pairs. Each frequency pair will have a data rate which is the inverse of the assumed reverberation time, T, and the total data rate will be:

$$C = M / T \tag{3}$$

For T = 50 – 300 ms, this gives a C=100 – 16.7 bits/sec and C/W = 0.025 – 0.0042 bps/Hz. The value of C/W depends on the maximum movement (M) and the maximum reverberation (T). The channel efficiency is lower than for speech, something which indicates that there is a potential for new modulation schemes that are better suited for the acoustic environment. Although beyond the scope of this analysis, this may make some indoor positioning applications feasible that are barely possible today due to too low capacity to handle many transmitters at almost the same time through time-multiplexing.

B. Range

The performance of communications systems can be analyzed using a calculation of the received signal to noise ratio and the passive sonar equation:

$$SL - PL - NL > DT \tag{4}$$

SL – Source level in dB at range R_0 (usually 1 m), PL – Propagation loss, due to spreading $6\log_2(R/R_0)$ and attenuation αR, NL – Noise level, DT – Detection threshold in dB.

The first example is a rather exotic form of human communications. There exist whistled languages among several peoples of the world, like the Silbo language of the island of La Gomera of the Canary Islands (Spanish silbar = to whistle). Silbo has been reported to reach 2-3 km in wooded, open valleys and 8-10 km in waveguide-like valleys with rock sides of low absorption [10] (p. 40). The normal language is translated into a system of tonal changes and stops that are whistled in the 1 – 4 kHz range. The translation is so precise that it is possible to communicate names that are unknown both to the whistler and listener a priori.

The source level of the whistlers can reach SL = 110-120 dB SPL at R_0 = 1 meter [10] (p. 37). This is in the order of 30 dB louder than the loudest speech. In addition to the increased source level, it is reported in [10] (p. 36) that the outdoors background is very quiet in areas where whistled languages are used. In particular there is no man-made noise. The background level was found to be in the order of 20 - 40 dB SPL, and could reach 50 dB SPL with 5 m/s wind. It is not clear as to how the noise level was measured, but from the look of the curves in Fig. 19 of [10] we assume that it was measured with octave band filters. Such a filter centered on 2000 Hz gives a noise spectral density of 10 log(1414) lower than the sound pressure level or NS = -11.5 to 8.5 dB-Hz.

The threshold for detection can be taken from a somewhat similar form of communications, decoding of Morse code by ear. The study in [11] reports a range of variation of the threshold of about 3 dB from 20% to 80% discrimination in a noise bandwidth of 230 Hz. The absolute value of the threshold varies with the telegraphy speed. For average speeds of about 60 characters per minute, it is about 3 dB (80% discrimination). This corresponds to a threshold of DT = 3 + 10 log(230) = 25.6 dB-Hz. We will assume a threshold DT = 25 dB-Hz here.

In the most favorable case for Silbo, there is no attenuation, only loss due to spherical spreading. In that case, SL = 120 dB and the lowest noise level, NS = -11.5 dB-Hz, in eq. (6) gives:

$$120 - 6\log_2 R - (-11.5 + 10\log(3000)) > 25 \qquad (5)$$

This equation is illustrated in Fig. 1. The range can be solved and is shown as the right-hand circle in Fig. 1 at R ≈ 4 km, which is close enough to the reported range in the 2-10 km range. Had Silbo been used in an industrial noise environment with a background noise of 70 dB SPL (up by 50 dB) or 35.2 dB/Hz, and ideal conditions with respect to attenuation, the range would only be about 18 meters as shown by the left-hand circle in Fig. 1. Thus both the high source level of whistling and the very low noise level, explain the long range properties of the whistled languages. Low noise level alone is not enough to get long range. An important conclusion is also that it is possible to predict the range approximately using standard attenuation and signal to noise considerations.

The second example is speech and shouting. There is no absolute value for the maximum sound pressure level (SPL) for shouted speech. We have found values such as 1 Pa (94 dB), 90 dB SPL at 0.3 m [9] and 88 dB SPL at 1 m. We will assume a threshold for detection in the range assumed in optimal detection theory of 15-20 dB-Hz [13]. Here we will use 20 dB-Hz. Note that this is 5 dB lower than the threshold value previously assumed, building on an assumption that our hearing system is better adapted to speech than to Morse code and whistling. An analysis of speech for SL = 88 dB and a noise level, NS = 30... 40 dB in eq. (1) gives:

$$88 - 6\log_2 R - NS > 20 \qquad (6)$$

The maximum range is about 25 m for 40 dB noise level and about 80 m for 30 dB noise level, which both seem to agree with common sense. Thus we are confident that the equation can be used for ball park predictions of range. This equation has been illustrated in Fig. 2 with the two solutions indicated by circles.

A similar range calculation for a 40 kHz communications system is now done based on a source level *SL*=115 dB SPL at 1 meter. The propagation loss consists of spherical spreading and attenuation which increases with relative humidity from α=0.27 dB/m (0% RH) to a maximum of 1.25 dB/m for 40% RH at 20 ºC and 1 atmosphere pressure [14]. The processing bandwidth *BW*=25 Hz and the noise level is 75 dB SPL measured in 3000 Hz or 40.2 dB/Hz.

In the worst case (40% RH), the range can be found from

$$115 - 6\log_2 R - 1.25R - (40.2 + 10\log(25)) > 20 \qquad (7)$$

This equation has been illustrated in Fig. 3 showing derating of output power with range for three cases: only spherical spreading, 0% relative humidity and the worst case for 40 % humidity. The most pessimistic case gives a range of 14.3 meters (left-hand circle in Fig. 3) which is close to the

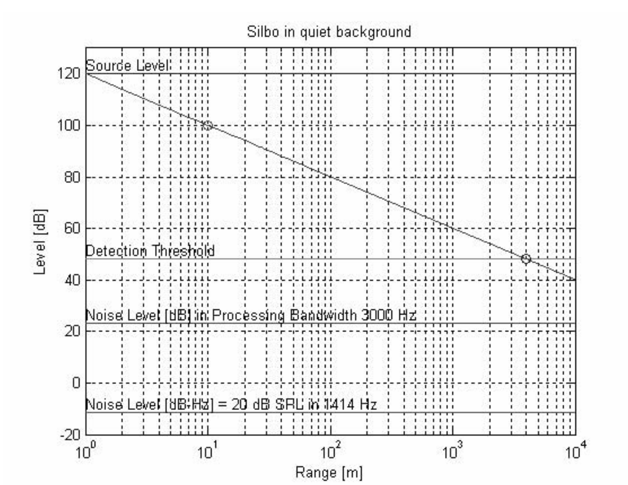

Fig. 1. Received level vs. range for Silbo whistling language. The circles show predicted range for a quiet background (4 km) and in industrial noise (18m).

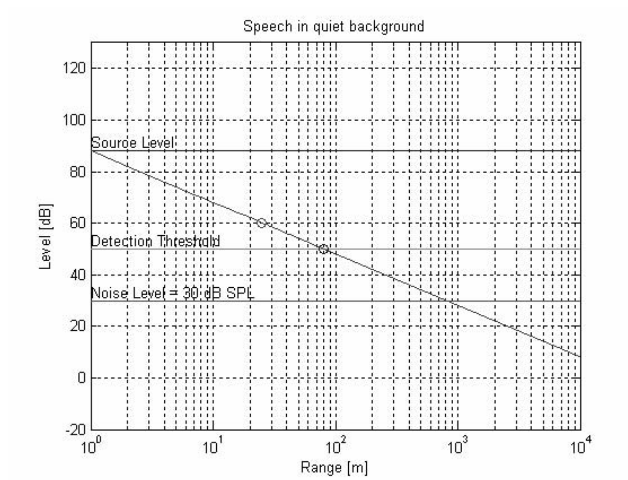

Fig 2. Received level vs. range for shouted speech in a quiet background. The circles show predicted range of 80 m for noise level 30 dB SPL and 25 m for noise level 40 dB SPL.

minimum practical range for the system developed. In 0% relative humidity, the range would increase to 36 meters, and if also the noise reduces by 10 dB, the range will be almost 60 meters.

V. SAFETY

The safety regulation of [15] uses ultrasonic welders and cleaners as examples. Unlike [16], it stresses the importance of avoiding direct body contact with the source as is the case for the present system. [15] then allows for a 30 dB increase of the threshold. The 1/3 octave band safety threshold for a band centered on 40 kHz is therefore 115 dB SPL + 30 dB = 145 dB SPL. This value represents "*conditions under which it is believed that nearly all workers may be repeatedly exposed without adverse effect on their ability to hear and understand normal speech*". The worst case scenario is a tag in a shirt pocket with the transducer pointing upwards (115 dB SPL @ 1 meter or about 125 dB @ 0.3 m). This gives an exposure level which is 20 dB below that of the safety regulation.

VI. CONCLUSION

The performance of an ultrasonic indoors communications system has been estimated by comparing with speech and the Silbo whistling language both for bit rate calculation and range estimation. Although there are many uncertainties in these first attempts at estimating performance parameters, the system has been found to be consistent with these predictions.

The system assumes a fixed detection threshold, but one of the results is that it is desirable with a system which adapts to the environment. Such a system could achieve a much higher range under favorable conditions, while maintaining robustness under less favorable conditions. The bit rate comparison with speech also indicates that speech communications utilizes the

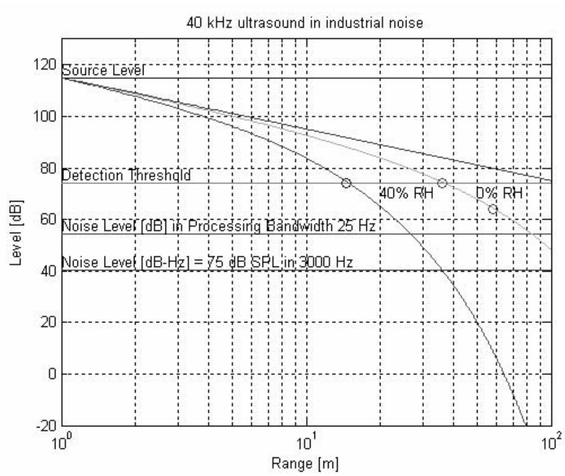

Fig 3. Received level vs. range for 40 kHz ultrasound industrial noise. The circles show predicted range of 14.3 m for noise level 75 dB SPL and worst case attenuation (40 % RH), 18 m for 0 % RH, and almost 60 m for 0% RH and 65 dB SPL noise level.

acoustic channel more efficiently than FSK-systems. A challenge is therefore to come up with better modulation schemes that are better suited to the acoustic environment. There is a potential for an increase in data rate by a factor between 4 and 20. This could directly impact on the ability of a positioning system to locate many objects at the same time, or to get a faster update rate, as the ultrasonic channel has to be time-shared between the various transmitters in use.

ACKNOWLEDGMENT

Thanks to Prof. O. Kr. Ø. Pettersen, Norwegian University of Science and Technology/SINTEF for help in interpreting the acoustic measurements of [10] and to Sonitor Technologies AS for collaboration on the hardware used for testing.

REFERENCES

[1] A. Ward, A. Jones, and A. Hopper, "A New Location Technique for the Active Office," IEEE Personal Communications, Vol. 4, No. 5, October 1997, pp. 42-47.

[2] N. B. Priyantha, A. Chakraborty, H. Balakrishnan, "The Cricket Location-Support system," Proc. 6th ACM MOBICOM, Boston, MA, August 2000.

[3] Y. Fukuju, M. Minami, H. Morikawa, T. Aoyama, "DOLPHIN: An Autonomous Indoor Positioning System in Ubiquitous Computing Environment," in Proc. I EEE Workshop on Software Technologies for Future Embedded Systems, Hakodate, Japan, pp. 53-56, May 2003.

[4] Eiman Elnahraway, Xiaoyan Li, and Richard P. Martin, "The limits of localization using signal strength: A comparative study," in Proc. of the First IEEE Int. Conf. on Sensor and Ad hoc Communications and Networks , Santa Clara, CA, (2004).

[5] A Baggeroer, Acoustic telemetry - an overview, IEEE Trans. Ocean. Eng., OE-9, 1-10 (1984).

[6] D B Kilfoyle and A. Baggeroer, The state of the art in underwater acoustic telemetry, IEEE Trans. Ocean. Eng., OE-25, 1-1111 (2000).

[7] S. Holm, O. B. Hovind, S. Rostad, and R. Holm, "Indoors data communications using airborne ultrasound," in Proc. IEEE Int. Conf. Acoustics, Speech, Sign. Proc., Philadelphia, PA, (2005).

[8] H. E. Bass and L. N. Bolen, "Ultrasonic background noise in industrial environments," Journ. Acoust. Soc. Am., vol 78, No. 6, pp. 2013 – 2016, Dec. 1985.

[9] L. Rabiner and B.-H. Juang, "*Fundamentals of Speech Recognition*," Chapter 2, Prentice-Hall, 1993

[10] R. G. Busnel and A. Classe, *Whistled Languages*, Springer (1976).

[11] P. Montnemery, B. Almquist, S. Harris, "Recognition of telegraphy disturbed by noise at different S/N-ratios and different telegraphy speeds," Scand. Audiol., Vol. 20, pp. 33-39, 1991.

[12] B. Gold and N. Morgan, *Speech and audio signal processing*, J. Wiley & Sons (2000)

[13] S. Holm, "Optimum FFT-based frequency acquisition with application to COSPAS-SARSAT," IEEE Aerospace and Electronic Systems, pp. 464-475 April 1993.

[14] H. E. Bass, L. C. Sutherland, A. J. Zuckerwar, D. T. Blackstock, and D. M. Hester, "Athmospheric absorption of sound: Further developments," Journ. Acoust. Soc. Am., vol 97, No. 1, pp. 680 – 683, Jan. 1995.

[15] U.S. Department of Labor, Occupational Safety & Health Administration, Section III: Chapter 5.V. Ultrasonics, 2000.

[16] International Non-Ionizing Radiation Committee of the International Radiation Protection Association, "Interim guidelines on limits of human exposure to airborne ultrasound," Health Physics, Vol. 46, No. 4, pp 969-974, April 1984 (see www.icnirp.de).