Speech Watermarking for the VHF Radio Channel

Konrad Hofbauer[1,2], **Horst Hering**[2] **and Gernot Kubin**[1]

[1] Graz University of Technology
SPSC – Signal Processing
and Speech Communication Laboratory
8010 Graz, Austria
gernot.kubin@tugraz.at

[2] Eurocontrol Experimental Centre
INO – Innovative Research
91222 Brétigny-sur-Orge, France
konrad.hofbauer@eurocontrol.int
horst.hering@eurocontrol.int

***Abstract*— In air/ground voice radio communication for air traffic control, an automatic identification of the 'talking' aircraft could avoid call sign confusion and ambiguity. This automatic identification can be achieved with speech watermarking, with which it is possible to transmit the aircraft identification or other data over the existing VHF radio communication link. The digital data is embedded into the existing analogue voice transmission between air traffic controller and pilot.**

The constraints and requirements for this application yield to a watermarking system design which does not require a modification of the existing radio infrastructure. A number of speech watermarking techniques are suitable for this purpose, of which speech perception-based watermarking is a novel approach which promises high data rates along with low audibility. The watermarking system has to be robust against the noisy radio transmission fading channel, of which the properties will be determined in upcoming flight experiments. For measuring the time-variant impulse responses of the channel, the proposed baseband measurement method allows for an inexpensive setup.

I. Introduction

Air traffic control has relied on voice communication between aircraft pilots and air traffic control (ATC) operators since its beginning. The aeronautical radio channel is a so-called 'party-line' channel and is used by the air traffic controller and all aircraft in the corresponding flight sector. In order to establish meaningful communication, all pilots start their messages with the verbal call-sign to identify themselves to the controller. Vice-versa, the controller starts the message with the call-sign of the aircraft.

At present, the ATC computer systems have no knowledge about which aircraft is transmitting at a certain moment. This knowledge would however be very useful, as then the transmitting aircraft could be indicated to the controller, thus avoiding mis-identification and call-sign confusion.

By using so-called 'speech watermarking' techniques, it is possible to embed digital identification information, such as the aircraft identification address or tail number, into the analogue voice communication ([1], [2]). It is thus possible to transmit in conjunction with the voice transmission itself as well the identification of the aircraft. The transmitting aircraft can thus be unambiguously identified by the ATC ground systems.

Speech watermarking allows the transmission of a short digital message *within* the existing analogue VHF ('very high frequency') radio communication link. The data transmission and the analogue voice transmission can concurrently share the same legacy radio channel in a backward-compatible manner. The watermarking system does not require modification of airborne or ground radio equipment, only a small add-on is necessary. Speech watermarking systems are characterised by the trade-off between data rate, robustness and perceptual impairment.

The aim of this paper is to develop a system design for transmitting aircraft identification and other data over the VHF radio communication system, and to analyse the applicability and the capabilities of existing and new speech watermarking algorithms, which are the core of the system design.

The outline of this paper is as follows. In Section II the basic requirements and constraints given by the application are outlined. A system structure resulting from these constraints is shown in Section III. Section IV focuses on the algorithms for speech watermarking. The VHF radio channel is the critical touchstone of the system and analysed as outlined in Section V. Conclusions are drawn in Section VI.

II. System Requirements and Constraints

A number of requirements and constraints apply to an aircraft identification system for the aeronautical radio [3].

A. Deployment–driven constraints

As the basic approach is to improve the current communication system, a short time to operational use and a *simple deployment* of the system components is necessary. As a consequence, the system has to be *backward compatible* to the legacy radio system currently in use. Even though changing to a completely new radio technology would have many benefits, analogue radio will continue to be used worldwide for many years to come. Furthermore, the system should need only minimal modifications to existing aircraft equipment and should not occupy additional frequency bandwidth. These constraints make inapplicable most of the standard digital communication techniques widely used in other areas.

The system will consequently be a watermarking system, which transmits the data *within* the existing analogue voice channel.

B. *Operational requirements*

The operational requirements specify the necessary *performance* figures of the system.

The required *real-time availability* of the identification implies a certain *data rate* for the digital transmission. It is indispensable to assure a *robust transmission* of the message and to be able to verify the *validity* of the received data.

From the user's point of view, the system should not degrade the perceptual quality of the voice transmission. Additionally, it should be autonomous and transparent to the user and not require changes to the well-established procedures in air traffic control and onboard the aircraft.

III. System Elements

Electronic watermarking in its most general definition is the technique of embedding some information, often called 'the watermark', into a host signal without noticeably modifying that signal. Figure 1 shows the general outline of a speech watermarking system.

a) Voice Signal: The voice signal is the host medium that carries the watermark or within which the watermark is hidden. In the aeronautical radio application the host signal is an electrical speech signal which is produced from the speaker's microphone or headset.

b) Watermark: The watermark itself is the information or data that is embedded into the voice signal. It could for example consist of the 24 bit aircraft identifier and—depending on the available data rate—auxiliary data such as the aircraft's position.

c) Watermark Encoder: The watermark encoder is an electronic device, which could be fitted into a a small adaptor box between the headset and the existing VHF radio[4].

It converts the analogue speech signal to the digital domain. An integrated digital signal processor embeds the watermark data according to embedding algorithms which are the subject of Section IV. The watermarked digital signal is then converted back to an analogue speech signal for transmission with the standard VHF radio.

d) Transmission Channel: The transmission channel consists of the airborne and ground-based radio transceivers, corresponding wiring, antenna systems, etc., and the VHF radio propagation channel. It has crucial influence on the performance of the system and is therefore the subject of investigation in Section V.

e) Received Signal: Although the signal contains a watermark, it is technically and perceptually still very similar to the original audio signal and can therefore be received and listened to with every standard VHF radio receiver without any modifications. This allows a stepwise deployment and parallel use with the current legacy system.

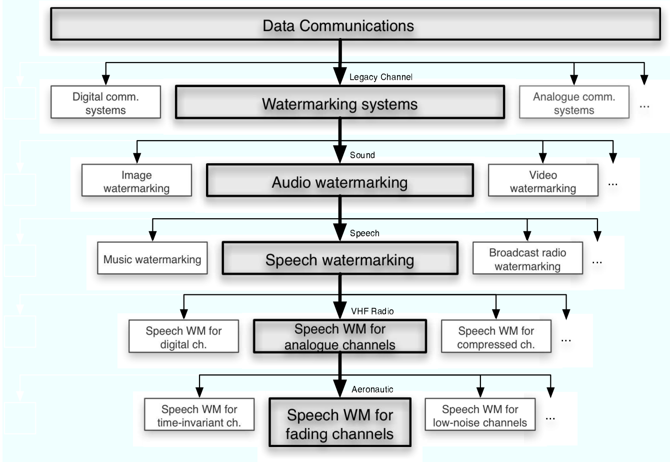

Fig. 2. Research focus within the field of data communications.

f) Watermark Decoder: The watermark decoder extracts the data from the received signal, assures the validity of the data and displays it to the user. The information can then become integrated into the ATC systems, e.g. by highlighting the radar screen label of the aircraft that is currently transmitting.

IV. Watermarking Methods

The current research focuses on algorithms for the watermark encoding and decoding. Figure 2 shows how the area of research narrows down by considering the constraints as outlined in Section II.

A. *Research focus*

Transmitting the identification tag is first and foremost a data communications problem. As the transmission should occur simultaneously within the legacy channel, a watermarking system has to be applied. The film and recording industries have shown increasing interest in watermarking for copyright protection of audio, video and multimedia content. A lot of of research has consequently been conducted in this area. The body of research concerning audio watermarking is small and even more so in the case of watermarking tailored to speech signals. (see e.g. [2], [5], [6]).

While most of the existing watermarking schemes are tailored to be robust against lossy perceptual coding, this requirement does not apply to watermarking for the aeronautical radio communication. The research ultimately focuses on speech watermarking algorithms for the noisy and time-variant analogue aeronautical radio channel.

B. *Related work*

The following section gives a short overview of some existing watermarking systems tailored to the aeronautical application.

1) Multi-tone sequence: The most elementary technique to transmit the sender's digital identification is a multi-tone data sequence at the beginning of the transmission. Standard analogue modulation and demodulation schemes can be applied to create a very short high power data package which

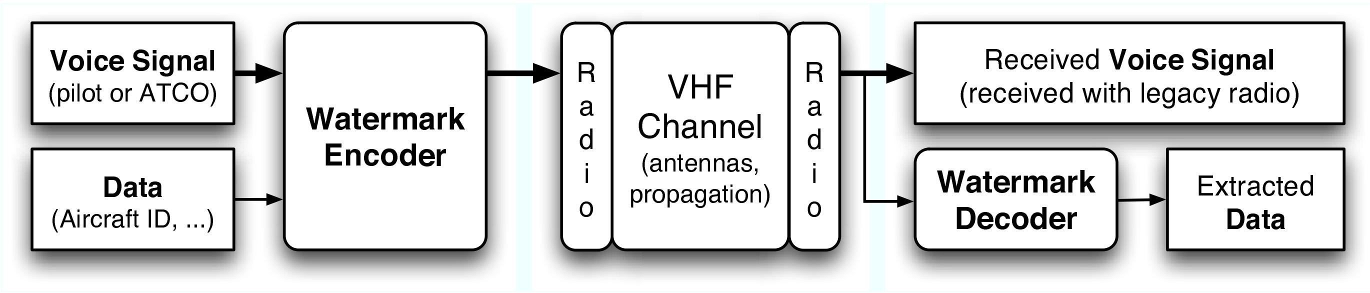

Fig. 1. Watermarking system for the aeronautic radio.

is transmitted before each message. This is a very simple and field-proven technology, which provides high robustness and data rates. Nevertheless, the transmission is clearly audible to the listener as a noise burst at the beginning of the message. The noise burst, which would hardly be accepted by the user base, could be removed by an adaptive filter. The fact that all of the receivers would have to be equipped with this technology renders the entire system undesirable.

2) Data-in-Voice (DiV): The 'Data-in-Voice' system presented in [7] tries to decrease the audibility of the multi-tone sequence by reducing its power and spectral bandwidth. Using a band-reject filter, the system removes some parts of the voice spectrum around 2 kHz as to operate an in-band modem. A standard 240 bit/s MSK (Minimum Shift Keying) modem technique is used, which occupies a frequency bandwidth of approximately 300 Hz. In order to fully suppress the audibility of the data sequence, the system also requires a filter on the side of the voice receivers.

3) Spread spectrum: The 'Aircraft Identification Tag' (AIT) system presented in [2] is based on direct sequence spread spectrum watermarking. This is a technique that embeds the watermark as additive pseudo-random white noise to the host signal.

The watermark encoder first adds redundancy to the data by an error control coding scheme. The coded data is spread over the available frequency bandwidth by a well-defined pseudo-noise sequence. The watermark signal is then spectrally shaped with a linear predictive coding filter and additively embedded into the digitised speech signal, thus exploiting frequency masking properties of the human perception.

The decoder relies on a whitening filter to compensate for the spectral shaping of the watermark signal produced by the encoder. A maximum-length pseudo random sequence (ML-PRS), which the decoder detects with a matched filter, is used to ensure synchronisation between encoder and decoder. The signal is then de-spread and the watermark data extracted. The voice signal acts as additive interference (additive noise) that impedes the decoder's ability to estimate the watermark. Therefore, even in the case of an ideal noiseless transmission channel, the data rate is inherently limited.

C. Informed embedding

Watermarking has a long history and is a well understood research area [8]. Based on so-called 'informed embedding' a new class of algorithms has developed within the last couple of years. The said approach promises an optimal trade-off between rate, distortion and robustness (see e.g. [9]).

The encoder makes use of the knowledge about the voice signal. The embedding is adaptive to the current input speech signal. The basic idea is that the speech signal is not a disturbance for the detection process anymore, but much more a carrier which is modulated by the watermark message. In order to achieve good results, the modulation has to occur in such a way that it does not significantly distort the speech signal and is robust against the transmission channel. To the authors' knowledge, no such watermarking system yet exists that would as well consider the peculiarities of the aeronautical application.

D. Speech parameter-based watermarking

Based on the principle of informed embedding, watermarking by modification of speech production parameters is possible. The so-called source-filter model provides the most dominant parameters of speech production. The quasi-periodic segments of speech are characterised by the fundamental pitch of the vocal chord excitation and the frequency response of the vocal tract. The non-periodic segments are well described by the time-varying variance of the noise process that models the excitation.

A small segment of a speech waveform of a female speaker is shown in Figure 3. The depicted segment contains the beginning of the phoneme '*a*' in the word 'f*a*st'. It illustrates the periodic structure of voiced speech. The vertical markers indicate the beginning of a so-called pitch cycle, that is, one period of the periodic speech signal. The length of the pitch cycle—the pitch period—indicates the instantaneous fundamental frequency or pitch of the segment.

The watermark system presented in [6] modifies the *average* pitch period of every phoneme. This average pitch is quantised with a quantiser that depends on the watermark message. The sound is then resynthesised using PSOLA (Pitch Synchronous Overlap Add), an audio modification algorithm widely used for pitch shifting and time stretching. While providing high robustness, this method fails to show a good data rate, with only one bit being embedded per phoneme.

Accordingly, a new encoder was developed in order to overcome the data rate limitation. The new method quantises the duration of every *pitch period* with a quantiser that depends

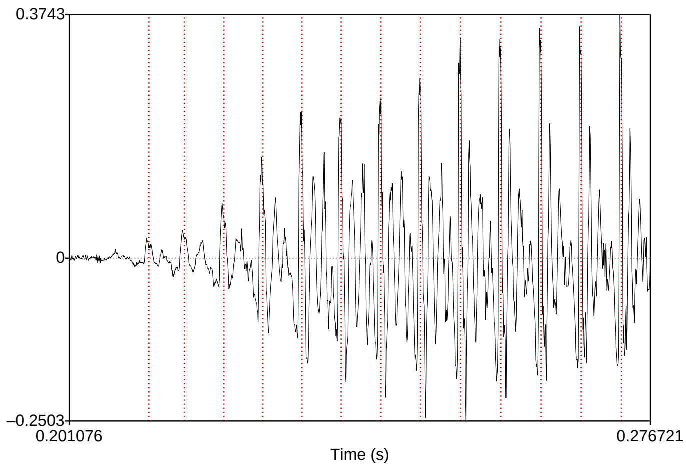

Fig. 3. Periodic structure and pitch cycle segments of the beginning of the phoneme *'a'* in the word 'f*a*st'.

on the watermark message. Again, the sound is resynthesised using PSOLA.

The method embeds one bit per pitch cycle, thus enabling high data rates. Nevertheless strong limitations were encountered. Only small modifications of the perceptually relevant parameter pitch cycle length are possible, as otherwise the embedding becomes audible and the sound of the speech unnatural. It is to conclude that trying to modify a perceptually relevant parameter is a fundamental limitation to the achievable robustness and data rate.

E. Speech perception based watermarking

The actual shape of the time domain representation of a speech signal is perceptually not very important. Similar to the effects exploited in speech coders, it is possible to modify the time domain speech signal without perceptual deterioration.

The first informal psychoacoustic experiments on the audibility of linear prediction residual and phase spectrum modifications have been conducted. With both techniques the data embedding is hardly audible. The data rate theoretically achievable with these methods is very promising and in the order of magnitude of 1 kBit/s. In other words, such methods allow a lot of information to be put into perceptually *irrelevant* parameters. In practise, the achievable data rate will certainly depend on the transmission channel and the encoder and decoder design, these being subject to further investigation.

This approach is seldom carried out in other watermarking domains, largely because of a preference for robustness against lossy perceptual audio or image coding. These lossy coders

remove the perceptually irrelevant parameters and would therefore render the watermark unreadable.

V. Aeronautical Radio Channel

The VHF radio transmission channel has a strong impact on the performance of the watermarking system in terms of data rate and robustness. It is important to know the degradation of the transmitted signal for the purpose of system design and simulation. It is therefore necessary to develop a model of the narrow-band aeronautic radio-channel, i.e., a mathematical description of the voice channel of a VHF radio.

A. Channel Modelling

1) Radio transmitter and receiver: The VHF radio transceivers represent one of the two general blocks in a VHF communication system. The low frequency (LF) audio part, consisting of preamplifiers, filters, automatic gain control, etc., are modelled separately from the high frequency (HF) part with modulation, demodulation, power amplifier, etc. The basic characteristics are usually available through data sheets or can be determined with standard HF measurement equipment.

2) Transmission channel: Detailed theoretical models of the transmission channel exist, from which the effects occurring on the channel can be derived and simulated. A detailed up-to-date literature review on applicable channel models is provided in [10].

The dominating effects are multipath propagation through reflection and scattering as well as Doppler shifts due to the fast movement of the aircraft. The conditions vary widely due to different positions and velocities of the aircraft during a flight (e.g. en route, approach, parking, . . .).

Many of the existing channel models contain a large number of parameters and it is therefore difficult to obtain realistic simulation scenarios. Unfortunately, very few measurements which could support and quantify the theoretical models are available to the public.

B. Channel Sounding

Measurements of time-variant impulse responses of the aeronautic VHF radio channel between a fixed or moving aircraft and a ground-based transceiver station are planned for beginning of 2006. Test signals transmitted over the standard radio channel during different phases of a flight will be recorded in these experiments.

The chosen method of measurement is as follows: Narrowband measurements (within one 8.33 kHz channel) from baseband LF audio to baseband LF audio will be pursued. Precomputed measurement signals (maximum length sequence, chirp, . . .) will be transmitted half-duplex in both directions (up- and downlink) and recorded on the receiver side. The measurements will be post-synchronised through parallel recording of the transmitted and received signal and a highly accurate 1 PPS (one pulse per second) GPS synchronisation signal. Figure 4 shows the basic setup for the measurements.

Even though this approach will not deliver results as accurate as wideband channel sounding could, it is the method of choice. Wideband channel sounding would otherwise require a setup that is large and complex, that is very expensive and that occupies a frequency bandwidth which is virtually unfeasible in the VHF band.

VI. Conclusion

Speech watermarking is the key technology for embedding an aircraft identification tag into the existing air/ground voice

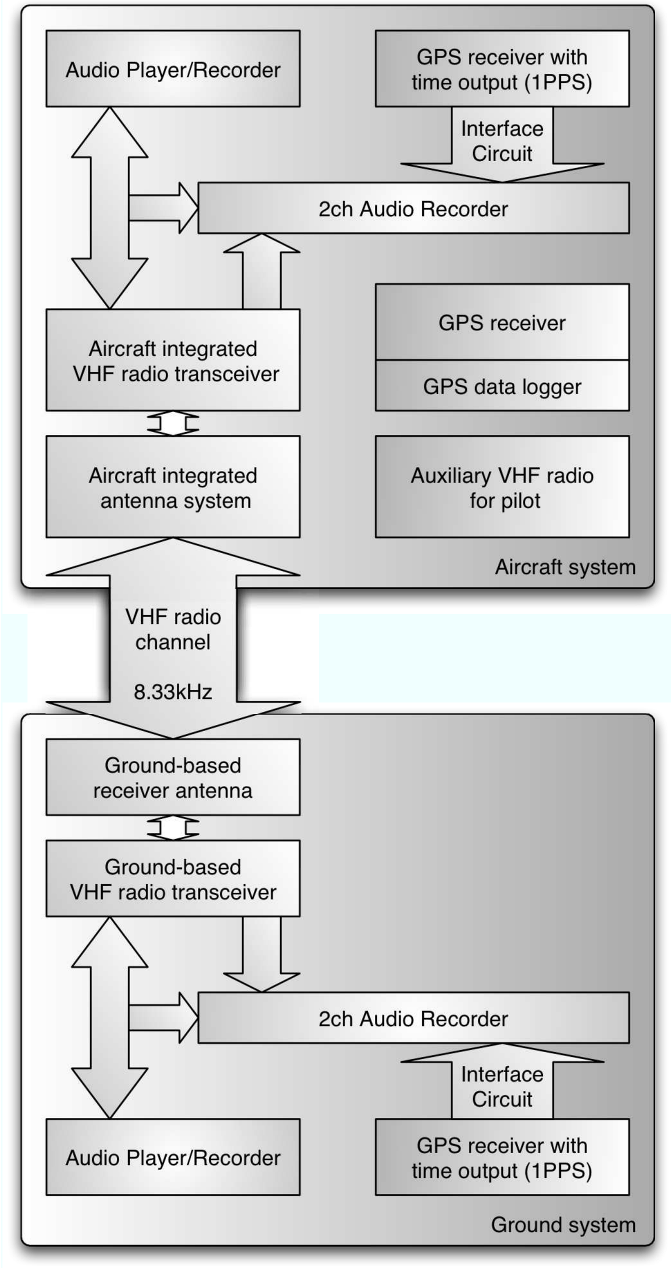

Fig. 4. A setup for narrowband aeronautical radio channel sounding.

communication link. It allows for automatic detection of the transmitting aircraft and thus helps avoiding call sign confusion and ambiguity.

Future research shall focus on the possibility of extending the capabilities of the existing systems by developing specialised watermarking algorithms. Even though a lot more research is required, the proposed speech perception-based watermarking algorithms promise data rates which could open a door to applications far ahead of sole identification. The additional capacity could be used for example for transmitting GPS positions for general aviation, security keys for authentication, or even side information about the original speech signal for receiver-side speech quality enhancements [11].

The planned channel measurements serve two purposes. The insight gained into the peculiarities of the VHF channel will support the development of powerful watermarking algorithms. The measurement results will, on the other hand, facilitate a realistic simulation of the channel for evaluation of candidate systems.

REFERENCES

[1] H. Hering, M. Hagmüller, and G. Kubin, "Safety and security increase for air traffic management through unnoticeable watermark aircraft identification tag transmitted with the VHF voice communication," in *Proceedings of the 22nd Digital Avionics Systems Conference (DASC 2003)*, Indianapolis, USA, 2003.

[2] M. Hagmüller and G. Kubin, "Speech watermarking for air traffic control," Eurocontrol Experimental Centre, EEC Note 05/05, 2005.

[3] K. Hofbauer and H. Hering, "Digital signatures for the analogue radio," in *Proceedings of the 5th NASA Integrated Communications Navigation and Surveillance Conference and Workshop (ICNS)*, Fairfax, USA, May 2005.

[4] H. Hering and K. Hofbauer, "System architecture of the onboard aircraft identification tag (AIT) system," Eurocontrol Experimental Centre, EEC Note 04/05, 2005.

[5] Q. Cheng and J. Sorenson, "Spread spectrum signaling for speech watermarking," in *Proceedings of the 2001 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2001)*, 2001.

[6] M. Celik, G. Sharma, and A. M. Tekalp, "Pitch and duration modification for speech watermarking," in *Proceedings of the 2005 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2005)*, 2005.

[7] M. Sajatovic, J. Prinz, and A. Kroepfl, "Increasing the safety of the ATC voice communications by using in-band messaging," in *Proceedings of the 22nd IEEE Digital Avionics Systems Conference (DASC '03)*, 2003.

[8] I. J. Cox and M. L. Miller, "The first 50 years of electronic watermarking," *EURASIP Journal of Applied Signal Processing*, vol. 2, 2002.

[9] I. J. Cox, M. L. Miller, and J. A. Bloom, *Digital Watermarking*. Morgan Kaufmann Publishers, 2001.

[10] BAE Systems Operations Ltd, "Literature review on terrestrial broadband VHF radio channel models," B-VHF Broadband VHF aeronautical communication system based on MC-CDMA, Report D-15, 2005. [Online]. Available: http:///www.B-VHF.org

[11] B. Geiser, P. Jax, and P. Vary, "Artificial bandwidth extension of speech supported by watermark-transmitted side information," in *Proceedings of the 9th European Conference on Speech Communication and Technology INTERSPEECH 2005-EUROSPEECH*. International Speech Communication Association (ISCA), 2005.