
Peer reviewed version

Link to published version (if available):
10.1109/30.883429
10.1109/ICCE.2000.854643

Link to publication record in Explore Bristol Research
PDF-document

**University of Bristol - Explore Bristol Research**

**General rights**

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available: http://www.bristol.ac.uk/pure/user-guides/explore-bristol-research/ebr-terms/
ROBUST AND SCALABLE MATCHING PURSUITS VIDEO TRANSMISSION USING THE BLUETOOTH AIR INTERFACE STANDARD

M. Fahim Tariq, Przemysław Czerepiński, Andrew Nix, David Bull and Nishan Canagarajah
Centre for Communications Research, University of Bristol
Merchant Venturers Bldg., Woodland Rd., Bristol BS8 1UB, United Kingdom
E-mail: {M.F.Tariq, P.J.Czerepinski}@bristol.ac.uk

ABSTRACT

This paper introduces an error resilient implementation of the matching pursuits algorithm for video coding. The video bitstream is transmitted using a simulation of the Bluetooth air interface standard, which recommends ARQ as a means of overcoming channel errors in the data packets. This approach may be unsuitable for real time and broadcast applications. Therefore, a modified receiver is proposed in this paper, which does not request the retransmission of erroneous packets, but instead passes them to the video decoder to exploit error resilience. This strategy is shown to be superior to a standard compliant system if ARQ cannot be applied. The work confirms that wireless communication standards should support a transparent mode for video applications.

1. INTRODUCTION

Bluetooth is a short-range radio data communication standard released by ETSI (European Telecommunications Standards Institute) in July 1999 [1]. It can be used as a data access point, a cable replacement or to form an ad-hoc network between electronic devices. The aim of this paper is to investigate the implications of transmitting real time video using the Bluetooth standard.

Since the majority of video coding algorithms, including standards [2][3], employ a combination of predictive and variable length coding (VLC), they are known to fail catastrophically in the presence of channel errors. Techniques such as forward error correction (FEC) and retransmission (ARQ) are often employed in communication links to overcome this problem. However, the former can compromise the compression performance, especially if transmission across a time varying channel is required, while the latter is usually unsuitable for real time or broadcast applications. Error resilient techniques are attractive bearers of multimedia information across noisy environments. Such methods can tolerate a certain level of transmission error and provide acceptable quality of reconstruction without resorting to FEC or ARQ [4][5].

The contribution of this paper is twofold: Firstly, to introduce a scalable and error resilient implementation of a state-of-the-art video codec, known as the matching pursuits algorithm. Secondly, to characterize the transmission of the video bitstream over a wireless network, using the ETSI Bluetooth standard [1]. The standard recommends the low-level rejection of erroneous data packets. In simulcast mode, such packets are retransmitted, which increases the delay. In broadcast mode they are discarded at the receiver by means of the flushing procedure. In this paper, a modification is considered to make such packets available to the error resilient video decoder. When the ratio between the packet and bit error rate is high, this strategy is demonstrated to provide a better quality of service, compared to the mandatory discarding of the corrupted packets.

This paper is structured as follows: Section 2 describes the Bluetooth standard and the simulation model used in this study. Section 3 presents an evaluation of the Bluetooth modem performance in a worst-case Rayleigh fading channel. Section 4 introduces the error resilient and scalable implementation of the matching pursuits video codec. The results of video transmission over Bluetooth are presented in Section 5, and conclusions are drawn in Section 6.

2. BLUETOOTH SIMULATION MODEL

Bluetooth is a time division duplex (TDD) system that operates in the unlicensed ISM band at 2.4 GHz. Slow frequency hopping (at a rate of 1600 hops/s) is used to combat the effects of interference and multipath fading. The modem achieves a gross bit rate of 1 Mbps and each channel occupies a bandwidth of 1 MHz. GFSK (Gaussian Frequency Shift Keying) modulation is used with a BT (bandwidth symbol period) product equal to 0.5 and a modulation index (h) in the range 0.28-0.38 [1].

Figure 1 shows the block diagram of the Bluetooth simulation model adopted in this paper. The simulation was performed at baseband taking 8 samples per symbol. A detailed description of the simulation procedure is given in the following section.
2.1 Bluetooth Transmitter

2.1.1 Packet Formation

The incoming binary data stream is prepended with a 1 or 2 byte payload header. The payload header contains information about the logical channel and the payload length. The payload can be protected using a rate 2/3 FEC code. The shortened (15,10) Hamming code is used for this purpose, with the capability of correcting one bit error and detecting two errors.

Bluetooth uses a time-slotted channel with a nominal slot length of 625 μs. The duty cycle of a standard packet is 366 μs, and a packet can be extended to occupy 3 or 5 slots. A summary of the various packet sizes and the corresponding gross symmetric and asymmetric data rates is listed in Table 1.

Each Bluetooth device has a 48-bit device address assigned to it by the manufacturer. This address is divided into three parts: a 24-bit Lower Address Part (LAP), a 16-bit non-significant address part (NAP) and an 8-bit Upper Address Part (UAP). The UAP is used to generate a 16-bit Cyclic Redundancy Check (CRC) on the payload, including the header. The CRC code is appended at the end of the payload. In this simulation, the CRC was generated using an arbitrary UAP.

The payload header is preceded with a 54-bit packet header, which contains link control information and active member addresses. The packet is then passed through a data whitening register which scrambles the packet using the 6 Least Significant Bits (LSB) of the Bluetooth master clock. Data whitening randomizes the data and minimizes the DC bias in the packet.

Whitened data is preceded with a 72-bit access code. This contains a synchronization sequence derived from the LAP of the Bluetooth device. Every device has a unique access code that distinguishes its packets from ones generated by other Bluetooth devices.

2.1.2 GFSK Modulation

After the entire packet has been formed, the binary data is mapped to antipodal pulses at 1 Mbps. They are then passed through a Gaussian filter with BT=0.5. The impulse response of the Gaussian filter is given as:

\[ h(t) = \frac{1}{\sqrt{2\pi sT}} \exp\left(-t^2/(2s^2T^2)\right) \]

where \( s = \sqrt{\ln(2)/2\pi BT} \). The convolution of the Gaussian filter and the rectangular pulses is defined as:

\[ g(t) = h(t) * \text{rect}\left(\frac{t}{T}\right) \]

where \( \text{rect}\left(\frac{t}{T}\right) = 1 \) for \( |t| < \frac{T}{2} \)

and zero otherwise. In this simulation, a 24-tap Gaussian filter was implemented that resulted in the Gaussian pulse shown in figure 2.
The filtered waveform is integrated by choosing a modulation index of 0.28. This corresponds to a maximum phase step of ±50.4 degrees per symbol period, as shown in figure 3. In practice, the GFSK signal can be digitally produced by modulating the integrated phase as I and Q signals onto the 2.4 GHz carrier. However, for simulation purposes, a baseband equivalent can be used based on the following mathematical relationship:

\[ S_r(t) = A_m e^{j\phi_m(t)} \]

where \( A_m \) represents the amplitude of the transmitted signal and \( \phi_m(t) \) the integrated phase, defined as:

\[ \phi_m(t) = \omega t \int \sum_{n} I_n e^{j(\omega t - nT)} d\tau + \phi_0 \]

where \( I_n \) is ±1 mapped according to the binary data and \( \phi_0 \) is the initial phase of the carrier. An eye diagram of the transmitted signal phase is shown in figure 4. A positive frequency deviation is interpreted as a +1 and a negative frequency deviation as a -1.

### 2.2 Radio Channels

Bluetooth is a frequency-hopped system. There are 79 channels available in Europe’s ISM band (except Spain and France where 23 channels are available). This corresponds to 79 hop frequencies. Each packet is transmitted on a different hop frequency.

In the case of combined packets (packets consisting of 3 or 5 slots), the hop frequency remains fixed until the end of the packet. In this simulation, an uncorrelated Rayleigh fading channel was assumed for every packet transmission. During the transmission of a single packet, it is likely that the connected Bluetooth terminals will remain stationary. Hence, a static fading channel with a Rayleigh distributed envelope was used. Rayleigh fading represents the worst possible scenario for the radio propagation channel. The effect of radio front-end noise was simulated by adding complex Additive White Gaussian Noise (AWGN) to the baseband modulated signal. An eye diagram of the received signal phase is shown in figure 5.

The distorted signal after passing through the channel can be described as:

\[ S_r(t) = A_c A_m \exp\{i[\phi_m(t) + \phi_c]\} + A_c \exp\{j\phi_c\} \]

where \( A_c \) and \( \phi_c \) represent the channel’s amplitude and phase, defined as:

\[ A_c = \sqrt{L_{ch}^2 + Q_{ch}^2} \]

where \( L_{ch} \) and \( Q_{ch} \) are Gaussian distributed random variables with zero mean and a variance of 0.5. They are generated by adding 12 uniformly distributed independent random variables \( U_i \) from 0 to 1, each representing a quadrature ray. They are defined mathematically as:

\[ I_{ch} = \frac{1}{\sqrt{2}} \left( \sum_{i=1}^{12} U_i - 6 \right) \quad Q_{ch} = \frac{1}{\sqrt{2}} \left( \sum_{i=1}^{12} U_i - 6 \right) \]
2.3 Bluetooth Receiver

The baseband simulation for the Bluetooth receiver starts with a complex low-pass filter that rejects all out-of-band noise from the received signal. In practice, this would be performed using an analogue filter implemented at RF. In this simulation, a baseband equivalent was implemented. The baseband signal bandwidth is 500 kHz, but due to the effects of local oscillator frequency offset (±75 kHz) and frequency drift (±40 kHz), a filter bandwidth of 615 kHz is required. Taking filter roll-off into account, a noise equivalent bandwidth of 750 kHz was assumed.

To simulate the effect of frequency drift, the received signal was multiplied by a carrier with a frequency equal to the offset frequency, \( f_o \). The received signal after the low-pass filter is given by the following equation:

\[
S_s(t) = A_s A_c A_o \exp \left( j \phi_0 + \phi_s(t) \right) + A_{no} \exp \left( j \phi_{no} \right)
\]

where \( A_s \) and \( \phi_s(t) \) represent the amplitude and phase of the demodulating carrier. The demodulating carrier phase is defined as:

\[
\phi_s(t) = 2\pi f_o t
\]

Since the noise is white, it is not affected by the offset modulation, however the receive filter rejects all out-of-band noise and thus modifies the noise amplitude and phase to \( A_{no} \) and \( \phi_{no} \) respectively.

After filtering, the received signal is passed through a phase detector. Ignoring the contribution of the AWGN, the detected phase can be written as:

\[
\phi_d(t) = \phi_n(t) + \phi_s(t)
\]

The detected phase is passed through a differentiator to recover the Gaussian filtered waveform. Due to the differentiation process, the constant phase of the channel is removed, leaving behind the amplitude noise and the effects of the frequency offset on the Gaussian signal. If the frequency offset is very small (i.e. the phase difference is small across a symbol period), the differentiation process will remove it. However, for fairly large offsets, an integrate and dump circuit was applied after differentiation. The differentiated signal also contains the effect of AWGN. An integrate and dump procedure helps to improve the signal to noise ratio prior to the decision device. The output signal from the integrate and dump filter is sampled at the symbol rate. A decision is made and data de-whitening is performed, followed by the CRC procedure to detect erroneous packets.

3. SIMULATION RESULTS

3.1 Performance in a Static Rayleigh-Fading Channel.

The evaluation of performance in a Rayleigh fading channel can be used to characterize the worst possible situation for a Bluetooth system. The bit error performance of DH and DM packets was evaluated in Rayleigh fading and the results are plotted in figure 6. It can be observed that for low \( E_b/N_0 \) values, there is no significant difference between the two curves. However as the \( E_b/N_0 \) increases, a coding gain starts to appear.

At a bit error rate of 1 in 1000, DM packets (coded) have an approximate 2 dB gain over DH packets (uncoded). One of the reasons why DM packets are not superior to DH packets at low \( E_b/N_0 \) values is the fact that the (15,10) shortened Hamming code can only correct a single error and detect two errors in a block of 15 bits. If more than two errors occur in a block, it is likely that the decoder will introduce more errors than originally present by inverting correct bits.

Although the length of the packet does not change the bit error rate performance, it does have a considerable affect on the packet error probability. Longer packets have less chance of passing the CRC check as compared to shorter packets. This can be confirmed from figure 7, which plots the probability of packet errors against \( E_b/N_0 \). The packet error probability of DH1, DM1, DH5 (asymmetric) and DM5 (asymmetric) has been evaluated in a Rayleigh fading channel. This choice of packet formats reflects the upper
and lower bound of packet error (DM1 & DH5) for all the combinations defined by the standard. The packet error results indicate that there is a difference of 4 dB between DH1 and DH5 and 2.5 dB between DM1 and DM5 at a packet error probability of 0.1. At the same packet error probability, the difference between the DM1 and DH5 is approximately 8 dB.

4. THE VIDEO CODEC

4.1 Matching Pursuits

The video codec employed in this paper is based on matching pursuits—a method for decomposing a signal over an overcomplete basis set. The matching pursuits codec developed by Neff et al., which decomposes the displaced frame difference (DFD) signal over a basis of Gabor functions, was reported consistently to outperform standard discrete cosine transform (DCT) methods for low bit rate video coding [6][7].

An example matching pursuits decomposition is illustrated in figure 8. The aim is to characterize the coded DFD signal, shown in figure 8a, in terms of the dictionary of 256 separable functions, shown in figure 8b. The decomposition is iterative: the function that best matches the signal is determined by an exhaustive search. The contribution of this function is then subtracted from the signal, and the search is repeated on the residual. This process continues until a pre-determined quality or bit rate criterion is met. Figures 8c and 8d show how a matching pursuit gradually approximates the DFD signal after 10 and 100 iterations of the search process.

An elementary data structure that fully describes a single step of the decomposition consists of three parameters: (1) coefficient position in the DFD, (2) an index into the dictionary set and (3) coefficient magnitude. This structure is referred to as an atom. Contrary to the computationally expensive decomposition process, a matching pursuits reconstruction simply involves computing a linear combination of atoms.

The codec employed in this paper shares a lot of similarity with the codec described in [6]. However, the 256 basis functions are factorized into short-kernel convolutions. This enables a 20-fold reduction in the computational cost [8][9] without any loss of the reconstruction quality.

4.2 Error Resilience

Since the majority of video coding algorithms employ a combination of predictive and variable length coding (VLC), transmission errors usually lead to a significant degradation of the reconstructed signal, due to the following two interrelated mechanisms:

1) Temporal and spatial error propagation due to the predictive codec structure.
2) Synchronization loss due to decoding an incorrect variable length codeword.

The effects of synchronization loss can potentially be catastrophic, i.e. the decoder may fail to recover any meaningful information from the bit stream.

The codec presented in this paper addresses the second degradation mechanism outlined above. This is accomplished by abandoning any variable length entropy coding altogether. Instead, a combination of fixed length coding (FLC) and error resilient entropy coding (ERPC) is employed.

FLC represents the simplest and the least compact form of coding. However, it is naturally resilient, with a single bit error affecting only a single codeword. FLC is used for coding two of the three atom parameters: 8 bits are allocated to the dictionary index (to select one out of the possible 256 functions) and 4 bits are allocated to the product value.

The ERPC was developed by Cheng and Kingsbury [10] as an efficient, yet resilient method for positional coding of sparse data. A description of the algorithm falls beyond the scope of this paper; it is shown in [10] that the compression performance of ERPC is close to the first order source entropy, while a single error in the coded positional information affects an average of 2.7 coefficients. In this paper, the ERPC is employed to code the motion field and atom positions.

The clean channel coding performance of the proposed error resilient codec is shown in figure 9 for two example test sequences. For higher bit rates, it matches the performance of the VLC algorithm. For lower bit rates, its performance is slightly inferior, mainly due to the fact that the VLC algorithm codes the motion field more compactly.

It should be noted that a small portion of the data that comprises frame header information, such as the frame
type, the temporal reference and the quantization parameters, are highly protected against channel errors by means of forward error correction. These data fields are crucial for any meaningful decoding and constitute a small enough fraction of the bitstream to not compromise the compression performance.

4.3 Scalability

A two-layer SNR scalable codec was implemented. The base layer encoder operates in a standard motion compensated configuration. The enhancement layer adopts the reconstructed base layer signal as a prediction. This strategy prevents any temporal error propagation in the enhancement layer. The base layer can be decoded independently of the enhancement layer to provide a lower quality reconstructed signal. Decoding both layers provides a higher quality reconstructed signal.

The scalable bit stream adds flexibility to the system, which can be exploited in a number of circumstances, for example:

- Varying computational power at the decoder. If the receiving unit only has a limited amount of computational power available, it will focus on decoding the base layer bit stream and ignore the enhancement layer bit stream.

- Varying medium availability. A transmitting or transcoding unit may select between sending the high priority base layer data packets or both the base and enhancement layer packets depending on the amount of network traffic.

- Varying channel quality. If the data loss rate is likely to be high due to poor channel conditions, more protection can be added to the base layer data at the expense of ignoring the enhancement layer data.

5. TRANSMISSION RESULTS

The performance of the proposed video codec was tested in the presence of channel errors. Four packet configurations were investigated for transmitting base and enhancement layer data, namely:

1) DH1 (base) and DH1 (enhancement);
2) DM1 (base) and DH1 (enhancement);
3) DH3 (base) and DH3 (enhancement);
4) DM3 (base) and DH3 (enhancement).

In addition, two types of receiver were investigated. Firstly, a standard-compliant receiver that ignores packets that are received in error. According to the Bluetooth protocol, such packets would be retransmitted.

However retransmission is not considered here due to real-time constraints; it also cannot be realized when a broadcasting mode is employed. Secondly, a modified receiver is investigated, that makes erroneous packets available to the error resilient decoder. This is justified by the fact that the bit error rate figure is significantly lower than the packet error rate figure. Often, a single bit error disqualifies the whole data packet.

All tests were performed using the 300 frame CIF resolution sequence ‘Silent Voice’. Figure 10 shows the obtained PSNR results for various channel conditions. Every data point was obtained by averaging ten transmission trials. Figure 11 shows example reconstructed frames. The following observations can be made from figures 10 and 11:

- Both error resilient decoder configurations satisfactorily coped with bit errors and did not lose synchronization.
With the exception of one packet configuration (DM1—DH1), the policy of accepting all data packets surpassed the mandatory rejection of corrupted packets. The success of accepting all data packets can be related to the ratio between the packet and bit error rate figures. This is approximately equal to 10 for packet configuration (2), 20 for packet configurations (1) and (4) and 50 for packet configuration (3).

Decoding the enhancement layer improves the reconstruction quality only in good channel conditions.

Owing to the FEC protected base layer, packet configurations (2) and (4) enable a more graceful quality decay, compared to unprotected packet configurations (1) and (3). However, since the latter provide a better clean channel signal quality, the signal degradation in poor channel conditions is roughly the same whether or not a protected base layer is used.

The proposed approach, which allows the error resilient decoder to utilize all (including erroneous) packets, can outperform the standard-compliant system. The standard system rejects corrupted packets and these results provide a strong case for Bluetooth, and other wireless data standards, to support a transparent mode for video applications.

6. CONCLUSIONS

This paper has developed a baseband simulation of the Bluetooth radio standard. BER and PER results were presented for various packet types under worst case Rayleigh fading assumptions. The packet error results at a 0.1 probability showed a difference of 4 dB between DH1 and DH5 and 2.5 dB between DM1 and DM5. At the same packet error probability, the difference between the DM1 and DH5 was approximately 8 dB.

A scalable and error-resilient implementation of the matching pursuits video codec has also been developed. The implementation uses the error resilient positional code to encode the motion field and the atoms. The proposed codec was tested in the noisy Bluetooth radio environment, and shown to tolerate the fading conditions without recourse to forward error correction.
In addition, the results presented in Section 5 demonstrate that, depending on the relationship between the packet and bit error rate, it may be beneficial to pass all (including corrupted) data packets to the error resilient decoder. For the simulated Bluetooth environment, a ratio of PER to BER equal to 10 favours the standard-compliant rejection of erroneous packets, however for larger values of PER/BER, accepting erroneous packets offers a clear performance improvement. This forms a strong case for Bluetooth, and all other emerging wireless standards, to support a transparent mode for video applications.

ACKNOWLEDGEMENT

The authors are indebted to the EPSRC for funding this work. M. Fahim Tariq wishes to acknowledge Yusuf Baltaci and Ismail Kaya (Karadeniz Technical University, Turkey) for their help in the simulation of the Bluetooth receiver.

REFERENCES


M. Fahim Tariq obtained a B.Sc. in Electrical and Electronic Engineering in 1994 from the Eastern Mediterranean University in Turkey. He joined the Centre for Communications Research, University of Bristol, in 1995. He is currently a Research Associate working in the digital communications group. His research interests include advanced modulation techniques, adaptive equalisation, channel modelling and capacity evaluation techniques. He is a member of IEEE.

Przemysław Czerepiński (S'95–M'98) received the M.Sc. degree in electronics and telecommunications from the Wrocław University of Technology, Poland, in 1994. In 1999 he was awarded the Ph.D. degree by the University of Bristol, UK, where he submitted a thesis on displaced frame difference coding. Currently, he is a Research Associate at the Centre for Communications Research, University of Bristol. His research interests include all aspects of image and video coding technology.

Andrew Nix was awarded B.Eng. and Ph.D. degrees in Electrical and Electronic Engineering from the University of Bristol in 1989 and 1993 respectively. He is currently a Reader in Digital Communications at the University of Bristol. His research interests include radio-wave propagation modelling, advanced digital modulation techniques and high-speed wireless LAN design. He has published over 90 papers in these areas. He is an active member of ETSI BRAN and has contributed to the European HIPERLAN, LAURA, AWACS, WINHOME and SATURN projects. He is a member of the IEEE.

David Bull is currently Professor of Signal Processing at the University of Bristol. Prior to joining Bristol he was an Electronic Systems Engineer at Rolls Royce and subsequently a lecturer at the University of Wales (Cardiff). He leads the Image Communications Group in the Centre for Communications Research at Bristol and has worked widely in the fields of 1 and 2-D signal processing. He has published well over 150 papers and a book in these areas. His recent research has focused on the problems of image and video communications, in particular error resilient source coding, linear and non-linear filterbanks, scalable methods, content based coding and architectural optimisation. He is a member of the EPSRC Communications College and the programme management committee for the DTI/EPSRC LINK programme in Broadcast Technology. He is also a director of the Virtual Centre of Excellence in Digital Broadcasting and Multimedia Technology and a member of the UK Foresight ITEC panel.

Nishan Canagarajah is currently a Senior Lecturer in Signal Processing at the University of Bristol. He has B.A. (Hons) and a Ph.D. in DSP Techniques for Speech Enhancement, both from the University of Cambridge. He is a committee member of the IEE Professional Group E5, member of the Virtual Centre of Excellence in Digital Broadcasting and Multimedia Technology and an associate editor of the IEE Electronics and Communication Journal. He is also an editor of a book on Mobile Multimedia Technology. His research interests include image and video coding, non-linear filtering techniques and the application of signal processing to audio and medical electronics.