

SVD-aided Unequal-Protection Spatial Multiplexing for Wireless Video Telephony

Du Yang, Nasruminallah, Lie-Liang Yang and Lajos Hanzo
School of ECS, University of Southampton, SO17 1BJ, UK

Tel: +44-23-8059 3364; Fax: +44-23-8059 4508

E-mail: dy05r,n06r,lly,lh@ecs.soton.ac.uk; <http://www-mobile.ecs.soton.ac.uk>

Abstract—¹ The proposed Singular Value Decomposition (SVD) aided Spatial Multiplexing (SM) based Multiple-Input-Multiple-Output (MIMO) system is capable of achieving a high bandwidth efficiency. The SVD operation effectively splits the MIMO channel into orthogonal eigen-beams having unequal Bit-error-ratios (BERs). This property may be exploited for the transmission of speech, audio, video and other multimedia source signals, where the source-coded bits have different error sensitivity.

Index Terms—SVD, MIMO, H.264, UEP

I. INTRODUCTION

Video communication over wireless channel has attract substantial research interest in recent years because of the popularity of diverse video applications, such as for example video-phones, as well as YouTube- and BBC iplayer-style video streaming. Considering the typically higher video data transmission rate compared to other sources, such as for example audio [1] and text, as well as the error-prone and bandwidth-limited properties of wireless channels, video transmission schemes require both a high robustness and a high spectrum efficiency [2]. To achieve these ambitious goals, a variety of novel schemes designed for video telephony are proposed. A widely used technique is constituted by Unequal Error Protection (UEP), where the video data bits are classified into different groups according to their error-sensitivities and the more vulnerable bits are protected by stronger Forward Error Correction (FEC) codes, so as to increase their robustness to channel impairments. Another advanced near-capacity transmission scheme was proposed in [3], where the joint-source and channel-coding algorithm employed Variable Length Error Correction Coding (VLEC). It also employed an irregular FEC design combined with iterative decoding schemes for the sake of approaching the capacity, which is the upper-bound of the achievable bandwidth efficiency. Recently, employing multiple antennas for video transmissions was advocated as an attractive solution [2], since Multiple-Input-Multiple-Output (MIMO) systems have the promise of an increased bandwidth efficiency. The scheme proposed in [4] employed Space Time Trellis Coding (STTC) to exploit the spatial diversity provided by multiple antennas, additionally

using Trellis Coded Modulation (TCM) for error protection without extending the bandwidth requirements, which was also combined with Multi-level Coding (MLC) to provide UEP, for the sake of improving the MPEG-4 video scheme's performance.

In this paper, we propose a novel Singular Value Decomposition (SVD) aided Spatial Multiplexing (SM) based MIMO transmission scheme for video communications, which decomposes the MIMO channel into orthogonal eigen-beams exhibiting unequal channel gains characterized by their corresponding eigen-values. Consequently, the data transmitted via the different eigen-beams exhibits a different Bit-error-ratio (BER), which provides a novel form of UEP instead of using different-rate FEC schemes. Another advantage of the proposed scheme is its ability to perform well even when using a low-complexity detector. Since the eigen-beams are orthogonal to each other, a group of simple single-stream matched-filters can be used instead of complex multi-stream detectors. Naturally, by employing multiple transmit and receive antennas, the achievable bandwidth efficiency becomes higher than that of a single-antenna aided transceiver.

We investigate an interactive video-phone scenario using the H.264 [2] source codec. The channel is assumed to be a both spatially and temporally independent Rayleigh fading medium, which is perfectly known by both the transmitter and the receiver. *The novel contribution of this paper is the design of a powerful SVD-aided video-phone scheme, including the guidelines determining the number of activated eigen-beams, the modulation scheme and the power allocated to each beam in order to improve the average video performance.*

This paper is organized as follows. In Section II and Section III, the H.264 codec [2] and the proposed SVD-aided MIMO technique are introduced, respectively. In Section IV, our simulation results are illustrated. Finally, our conclusions and suggestions for future work are provided in Section V.

II. THE VIDEO CODEC H.264

The H.264/AVC codec constitutes an attractive candidate for wireless video streaming applications not only because of its high compression ratio, but more importantly, because of its error resilience. The H.264/AVC codec employs various Variable Length Coding (VLC) and predictive coding techniques in order to achieve a high compression efficiency. However, these entropy-coded techniques also make the compressed bit-stream

¹The work reported in this paper has formed part of the Core 4 Research Programme of the Virtual Center of Excellence in Mobile and Personal Communications, Mobile VCE, www.mobilevce.com, whose funding support, including that of EPSRC, is gratefully acknowledged. Fully detailed technical reports on this research are available to Industrial Members of Mobile VCE.

vulnerable to transmission errors. This is because a single bit error in the video stream may render the correct decoding of future codewords impossible. Furthermore, owing to predictive coding, the effects of channel errors are likely to be propagated to the neighboring video blocks. Therefore, source codecs designed for the highest possible compression efficiency are typically unsuitable for error-prone wireless transmission. The H.264/AVC codec employs a range of techniques in order to enhance the error resilience of the video data bits and to recover the flawless original video frame. As a result, the H.264/AVC codec has been widely used in wireless video communication and it is also employed in our study.

One of the error resilient techniques used in the H.264/AVC codec and also employed in our proposed scheme is Data Partitioning (DP), which constitutes the foundation of implementing UEP. To be more specific, in contrast to encoding all video parameters of a Macro-Block (MB) into a single bit-string representing a single video slice, we may generate several bit-strings per slice referred to as partitions, which represent different-sensitivity protection classes. This allows us to protect the different-sensitivity bits based on their relative importance. In the context of the H.264/AVC scheme we may create three different partitions, namely Partition A, B, and C.

- 1) Partition A contains the most vulnerable bits representing the slice header information, MB types, quantisation parameters and motion vectors. When Partition A is corrupted, the bits of an entire video slice, including Partitions B and C are dropped. In this scenario the entire slice is marked as corrupted and the decoder will apply error concealment techniques using the corresponding video segment of the previously decoded frame.
- 2) Partition B typically contains intra-frame coded MB coefficients and intra-MB Coded Block Pattern (CBP) bits, indicating which blocks within a MB contain non-zero transform coded coefficients. Recovery from error propagation can only be achieved after refreshing the corrupted image regions in the intra-frame mode by switching off inter-frame prediction for certain MBs, when no feedback channel is available. Typically a certain limited fraction of the MBs is encoded in the high-rate, but error-resilient intra-frame coding mode, hence Partition B typically hosts the lowest number of bits in an encoded slice.
- 3) Partition C carries inter-frame CBP information and inter-frame Motion Compensated Error Residual (MCER) bits for the particular MBs that were encoded using motion compensated prediction. By contrast, when in the intra-frame mode, Partition C carries the intra-frame CBP and intra-frame MCER bits for the MBs coded using the H.264-specific intra-frame prediction mode.

Therefore, in the H.264 video stream, Partition A is the most important one. The specific significance of a given Partition B and C is dependent on the importance of the associated Partition A, but they are not decodable in the absence of

Partition A. If Partition B of a slice is present along with A, the intra-frame MB update is added to the reconstructed frame, while if Partition C is present, the MCER of a slice is reconstructed and added to the motion compensated slice. As a result, the data of Partition A has to be protected with the highest priority, followed by Partition B and C. In the SVD-aided SM based MIMO system, Partition A should be transmitted via the eigen-beam having the highest-power, while Partition B and C should be transmitted with the aid of the eigen-beams having a lower power. However, due to the limit number of Partion A bits, some Partion B or C bits also transmitted through the highest-power eigen-beam in partice in order to efficiently utilized the limited bandwidth.

III. SVD-AIDED SPATIAL MULTIPLEXING

System Model: The system structure of the SVD-aided SM based MIMO system having n_T transmit antennas, n_R receive antennas and using the L highest-power eigen-beams is shown in Fig 1. At the transmitter side, the video bits are split into L' s ($1 \leq L \leq \min\{n_T, n_R\}$) streams according to their relative importance and mpped to the eigen-beams generated according to the $(n_R \times n_T)$ -element MIMO channel matrix \mathbf{H} . Then they are separately channel encoded, modulated and amplified according to the power allocation factor $\sqrt{p_l}$, while obeying $\sum_{l=1}^L p_l = 1$. By invoking the SVD for the channel matrix \mathbf{H} , the following equation is obtained [5]:

$$\mathbf{H} = \mathbf{S}\mathbf{V}\mathbf{D}^H, \quad (1)$$

where the $(n_R \times n_R)$ -element matrix \mathbf{S} and the $(n_T \times n_T)$ -element matrix \mathbf{D} are the Right-Hand-Side (RHS) and the Left-Hand-Side (LHS) unitary matrix generated from MIMO channel decomposition. The ordered singular values v_i ($v_i > v_j$; if $i < j$; $i, j \in \{1, \min\{n_T, n_R\}\}$), which correspond to the square root of the eigen-values λ_i of $\mathbf{H}\mathbf{H}^H$, constitute the $(n_R \times n_T)$ -element diagonal matrix \mathbf{V} . The MIMO channel is decomposed into L eigen-beams upon multiplying the signal to be transmitted by the first L columns of the RHS unitary matrix \mathbf{D} (denoted as \mathbf{D}_L) at the transmitter and upon multiplying the received signal by the first L rows of the LHS unitary matrix' Hermitian \mathbf{S}^H (denoted as \mathbf{S}_L^H) at the receiver. The resultant eigen-beams are orthogonal to each other and exhibit different channel gains, which are characterized by the corresponding singular-values. The video streams are demodulated and decoded separately. In order to reduce the BER and enhance the bandwidth efficiency, a Soft-Input-Soft-Output (SISO) demodulator and channel decoder are used exchanging extrinsic information with the aid of iterative decoding algorithms.

Analysis: When using the SVD, the received signal \hat{s}_l of the l -th eigen-beam is written as:

$$\hat{s}_l = v_l \sqrt{p_l} s_l + n_l, \quad (2)$$

where s_l denotes the transmitted signal, p_l is the transmitted power allocated to the l -th eigen-beam, v_l represents the l -th singular value, and n_l is the Gaussian noise. The specific singular value v_l , the allocated power value p_l and the variance

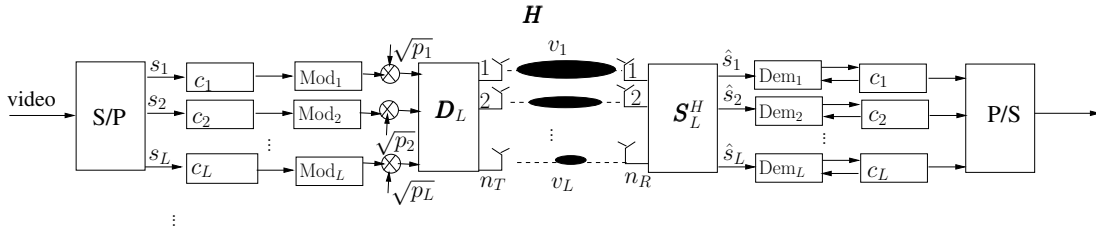


Fig. 1. System structure of the SVD-aided MIMO system having n_T transmit antennas n_R receive antennas, and L eigen-beams activated for transmission.

of the noise jointly determine the BER performance of the l -th eigen-beam provided.

The Cumulative Density Function (CDF) and Probability Density Function (PDF) of the ordered eigen-values were illustrated in [5] for various antenna configurations. In order to provide further intuitive insights, the histograms of the singular values recorded for a spatially independent Rayleigh fading MIMO channel matrix having (4×4) -elements are illustrated in Fig. 2, indicating that the distribution of the largest singular value v_1 is Gaussian-like and its mean value is 3.0959. By contrast, the distribution of the smallest singular value v_4 is Rayleigh-like, with a mean value of less than one. Fig. 3 shows the uncoded BPSK modulated BER performances of the SVD-aided SM system for the equal-power allocation scenario of $p_1 = p_2 = p_3 = p_4 = 0.25$, portraying both the overall average BER performance and the performance of each eigen-beam. For comparison, two standard benchmark curves recorded for BPSK modulated diversity orders of 1 and 16, respectively, are also shown along with the uncoded BPSK modulated Bell Lab Layered Space-time (BLAST)-type SM system [6]. The following observations may be made:

- 1) The slopes of the BER curves recorded for the individual eigen-beams at high SNRs are different, which indicates that the achievable diversity order or gain of each eigen-beam is different.
- 2) At high SNRs, the BER curve of the v_1 beam is parallel to the standard BER curve associated with the diversity order of 16, while the BER curve of the v_4 beam is parallel to the standard BER curve having a diversity order of 1. This indicates that the highest-power eigen beam achieves the full attainable diversity order and exhibits some extra coding gain, while the lowest-power eigen-beam has a diversity order of one and lacks coding gains.
- 3) The overall BER performance of the SVD-aided SM system is dominated by the performance of the lowest-power eigen-beam and hence on average it is worse than that of the BLAST-type SM system.

Based on these BER-related observations, the designs of the number of activated eigen-beams L , the modulation level of each beam M_l and the power allocated to each beam p_l is considered for the sake of improving the attainable BER performance. Since an improvement of the BER performance is expected to result in an enhanced video quality - although

this relationship is not necessarily true for all scenarios - it is worth using the attainable BER performance as our first design metric. It is clear that in the interest of improving the overall BER performance, the lowest-power eigen-beam should ideally be deactivated, since its performance is worse than that of a single-path Rayleigh channel. However, to achieve the same total data rate, as the system avoiding the deactivation of the low-quality eigen-beams, higher-order modulation schemes are required, which are more complex and may again result in a BER performance degradation. A plausible solution is to assign a high-order modulation scheme and commensurately more power to the eigen-beams having a higher eigen-value, which may be represented as $L < \min(n_T, n_R); p_1 \geq p_2 \geq \dots \geq p_L; M_1 \geq M_2 \geq \dots \geq M_L$. Based on this strategy, in Section IV, different configurations having a diverse number of activated eigen-beams, modulation schemes and power allocation arrangements will be compared in order to further investigate the associated system design trade-offs.

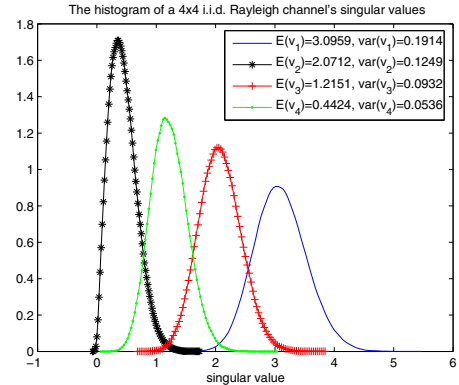


Fig. 2. Eigen-value PDF of a (4×4) -element MIMO channel. The channel matrix \mathbf{H} consists of 16 independent complex-Gaussian, i.e. Rayleigh variables with zero mean and unit variance.

IV. SIMULATION RESULTS AND DISCUSSIONS

The H.264/AVC JM 13.2 reference codec operated at $R_f = 15$ frames/s using the (176×144) -pixel Quarter Common Intermediate Format (QCIF) "Akiyo" video sequence, encoded at a near-constant bitrate of $R_b = 64$ kbit/s was transmitted through a MIMO system equipped with $n_T = 4$ transmit and $n_R = 4$ receive antennas. The half-rate Recursive Systematic Code

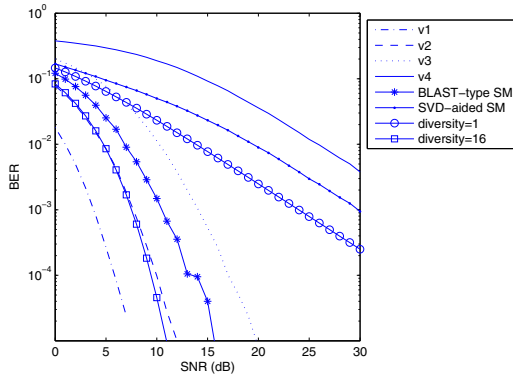


Fig. 3. The uncoded BER performance of the BLAST-type SM and of the SVD-aided SM system, respectively.

(RSC) RSC(215) was employed in all system configurations in order to focus our investigations on the UEP performance provided by the SVD-aided MIMO channel. Iterative decoding was employed at the receiver side and the number of iterations was set to $I = 2$ for striking a good compromise between enhancing the achievable BER performance and limiting the decoder's complexity.

Source Codec	H.264/AVC
Number of transmit antennas (n_T)	4
Number of receive antennas (n_R)	4
Target coded bits transmission rate	2 BPS
Number of Iteration	2
Interleaver length	1000 ~ 3000
SVD-aided SM configuration 1	
λ_1 , QPSK, RSC(215), $p_1=0.6$, Partition A and B	
λ_2 , BPSK, RSC(215), $p_2=0.2$, Partition C	
λ_3 , BPSK, RSC(215), $p_3=0.2$, Partition C	
SVD-aided SM configuration 2	
λ_1 , QPSK, RSC(215), $p_1=0.25$, Partition A and B	
λ_2 , BPSK, RSC(215), $p_2=0.25$, Partition C	
λ_3 , BPSK, RSC(215), $p_3=0.5$, Partition C	
$\lambda_1 > \lambda_2 > \lambda_3$	
SVD-aided SM configuration 3	
λ_1 , QPSK, RSC(215), $p_1=0.5$, Partition A and B	
λ_2 , QPSK, RSC(215), $p_2=0.5$, Partition C	
BLAST-type SM configuration	
BPSK	
RSC(215)	
MMSE detector	

TABLE I
SIMULATION PARAMETERS OF SIMULATION I

Let us now continue by addressing our design dilemmas posed in Section I, namely the design of the SVD-related parameters for the sake of improving the attainable video performance. The Peak-Signal-to-Noise-Ratio of the luminance (PSNR-Y) [2] is used as our video quality evaluation metric. The video data are first split into two streams according to their error sensitivity. One streams consists of the Partion A and B bits, while the other stream consists of the Partion C bits. The reason for combining the bits of Partion A and B is that on

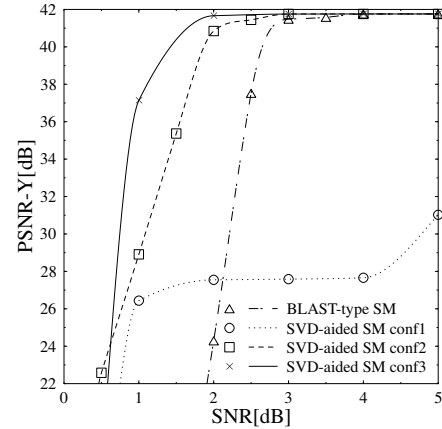


Fig. 4. PSNR-Y performance versus SNR of the SVD-aided SM based video systems using Configurations 1, 2 and 3, as well as of the BLAST-type SM based video system, respectively.

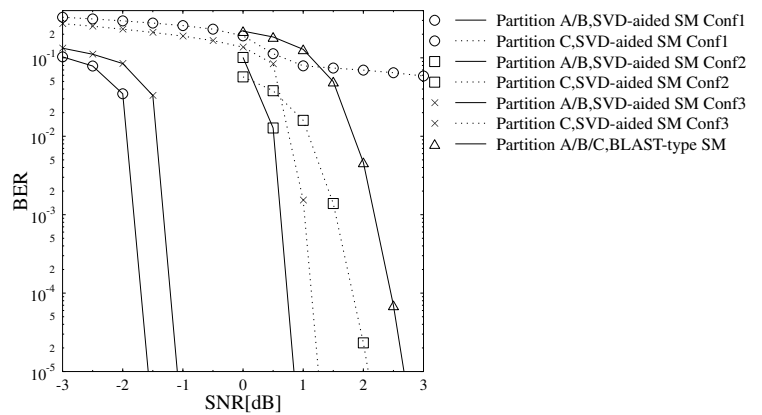


Fig. 5. BER performance versus SNR of the SVD-aided SM system using Configurations 1, 2 and 3 of Table I, as well as of the BLAST-type SM based video system. Two detection iterations were used for exchanging extrinsic information between the detector and the channel decoder in the schematic of Fig. 1.

average the total number of Partion A and B bits is similar to the number of Partion C bits. A further split might be needed according to the number of activated eigen-beams. Based on Section III, three SVD-aided SM configurations were set up, as illustrated in Table I. Naturally, the more error sensitive bits are transmitted through the eigen-beams having higher power in all configurations. The first two configurations of Table I activated the three highest-power eigen-beams. However, the first configuration allocated more power to the stronger eigen-beam and less to the weaker eigen-beam. By contrast, the second configuration of Table I allocated the power the

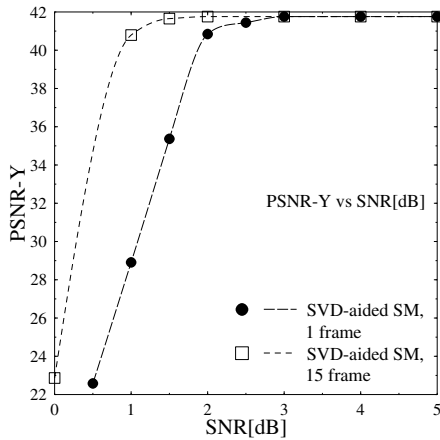


Fig. 6. PSNR-Y versus channel SNR performance of two system using either a single-frame or 15-frame transmission burst length. The SVD-aided SM based video Configuration 2 of Table I was used.

opposite way round. The third configuration activated only the two strongest eigen-beams and allocated the transmission power to the activated beams equally. Different modulation schemes were employed in the different configurations in a way to ensure that the transmission throughput remained 2 bits/symbol. For comparison, we employed an identical-throughput BLAST-type SM scheme using BPSK modulation, half-rate RSC coding and an MMSE detector as a benchmark. The source-coded video stream was processed on a frame-by-frame basis. As a result, the transmission burst duration was equal to $1/R_f = 66.7\text{ms}$ and the average number of bits per transmission burst was $R_b/R_f = 4266\text{bits}$. Our PSNR performance comparisons are shown in Fig. 4, while the BER performance comparisons are illustrated in Fig. 5.

The simulation results demonstrated that:

- 1) As expected, an increased throughput is achieved by employing the proposed SVD-aided SM scheme as a result of using multiple antennas.
- 2) The SVD-aided SM scheme provides UEP for the partitioned video bit-stream, while the BLAST-type SM benchmark scheme provides Equal Error Protection (EEP). As a result, observe in Fig. 4 that the SVD-aided SM system achieved a near-unimpaired video quality associated with $\text{PSNR-Y}=40\text{dB}$ at 1dB lower SNR than the BLAST-type SM benchmark scheme.
- 3) Focusing our attention on the PSNR-Y and BER performance of the SVD-aided SM system using Configuration 2 of Table I represented by the line marked by squares in Fig. 4 and Fig. 5 respectively, it is clear that besides having a low BER ($< 10^{-2}$) for the Partion A and B bits, an acceptable BER performance (< 0.08) is required for Partion C for the sake of reaching a decodable PSNR-Y level. Moreover, when the BER of all the bits is

lower than 10^{-4} , the video quality appears subjectively unimpaired and $\text{PSNR-Y}\approx 42\text{dB}$ was obtained.

- 4) As a result, the SVD-related parameters should be designed for providing as strong a protection for the sentisitive video bits as possible, while maintaining an adequate transmission quality also for the less error-sensitive bits.

Let us now briefly consider the effects of different transmission burst lengths on the achievable video performance. The SVD-aided SM system using Configuration 2 of Table I was used. The PSNR-Y performance of having a transmission burst length of 15 video frames was then evaluated and compared to the results associated with using a single-video-frame transmission burst length. The corresponding results are illustrated in Fig. 6. It is shown that the system associated with a longer transmission burst length reaches a near-unimpaired average video performance of $\text{PSNR-Y}\approx 42\text{dB}$ at a channel SNR of about 1dB lower than the system having a single-video-frame burst length. This gain accrues from the enhanced error correcting capability of the channel code achieved as a result of its increased codeword length. The price paid for achieving this gain is the 15-fold increased delay of $15 \times 66.7\text{ms} = 1\text{s}$, which would only be acceptable in a YouTube-type non-interactive vedio-streaming.

V. CONCLUSIONS AND FUTURE WORK

In this study, we have investigated a SVD-aided SM based MIMO scheme designed for UEP-aided video transmission. The three main advantages of this scheme are its ability of providing UEP, that of simplifying the receiver and that of offering an increased throughput. Moreover, we have demonstrate that the best strategy for the SVD-aided SM system design is to provide as strong a protection for the sentisitive video bits as possible, while maintaining an adequate transmission quality for the less error-sensitive bits. In our future work, we will investigate the achievable performance in conjunction with imperfect channel knowledge instead of assuming perfect channel knowledge.

REFERENCES

- [1] L. Hanzo, F. C. A. Somerville, J. P. Woodard, and H. T. How, *Voice and Audio Compression for Wireless Communications*. Wiley-IEEE Press, 2007.
- [2] L. Hanzo, P. Cherriman, and J. Streit, *Video Compression and Communications: From Basics to H.261, H.263, MPEG2, MPEG4 for DVB and HSDPA-Style Adaptive Turbo-Transceivers*. Wiley-IEEE Press, 2007.
- [3] R. G. Maunder, J. Wang, S. X. Ng, L. L. Yang, and L. Hanzo, "On the performance and complexity of irregular variable length codes for near-capacity joint source and channel coding," *IEEE Transactions on Wireless Communications*, vol. 7, no. 4, pp. 1338–1347, Apr. 2008.
- [4] S. Ng, J. Y. Chung, and L. Hanzo, "Turbo-detected MPEG-4 video using multi-level coding, TCM and STTC," *IEE Proceedings Communications*, vol. 125, no. 6, pp. 1116–1124, 2005.
- [5] R. Kwan, C. Leung, and P. Ho, "Distribution of ordered eigenvalues of Wishart matrices," *Electronics Letters*, vol. 43, pp. 31–32, Mar. 2007.
- [6] G. J. Foschini, "Layered space-time architecture for wireless communication in fading environments when using multiple antennas," *Bell Labs Technical Journal*, vol. 2, pp. 41–59, 1996.