PUBLIC

IST R&D. FP6-Priority 2.

SPECIFIC TARGETED RESEARCH PROJECT

Project Deliverable

SUIT Doc Number	SUIT_375
Project Number	IST-4-028042
Project Acronym+Title	SUIT- Scalable, Ultra-fast and Interoperable Interactive Television
Deliverable Nature	Report
Deliverable Number	D3.3
Contractual Delivery Date	July 31, 2007
Actual Delivery Date	July 31, 2007
Title of Deliverable	Scalable Joint Source and Channel Video Coding
Contributing Workpackage	WP3
Project Starting Date; Duration	01/02/2006; 27 months
Dissemination Level	PU
Author(s)	Maryse Stoufs, Adrian Munteanu, Pieter Heremans, Stewart Worrall, Zaheer Ahmad, Chee Hock Liew, Yves Dhondt

Abstract

In this deliverable, we propose a novel joint source-channel coding (JSCC)-methodology which minimizes the end-to-end distortion for the transmission over packet loss channels of scalable video encoded using SVC, the scalable extension of H.264/MPEG-4.

The proposed JSCC-approach performs channel protection using low-density parity-check codes and relies on Lagrangian-based optimization techniques to derive the appropriate protection levels for each layer produced by the scalable source codec. Our JSCC-approach for SVC can support spatial, temporal and quality scalability and can deliver an optimized channel protection in any scalable setting.

The deliverable also investigates use of unequal power allocation to improve the received video quality. The aim is to provide more protection to base layer data compared to the enhancement layers. Unequal power distribution among the video layers is carried out based upon the importance of layer under a total power constraint. An efficient two stage fast algorithm is proposed to reduce the complexity.

Keyword list: joint source-channel coding (JSCC), scalable video coding (SVC), unequal error protection, H.264/MPEG-4, Error Resilience.

Scalable Joint Source and Channel Video Coding (FEC-SVC)

SUIT_375

31-07-2007

Table of Contents

1 Introduction

1.1 Scope

This document is part of WP3.

In this document a joint source-channel coding (JSCC) methodology for scalable video transmission is proposed. The JSCC-methodology is proposed in conjunction with the scalable MDC-scheme described in deliverables D3.1 and D3.2 (Design of scalable MD-SVC) to realise efficient error resilient coding of the video stream. Specifically, the proposed JSCC achieves optimal network resource allocation by jointly optimizing the source rate allocated to encode the source video with a scalable video coder (SVC) and the channel rate allocated to forward error correction coding (FEC). FEC is a common tool employed to add redundancy bits to the compressed source bits and enable error detection and correction. The JSCC-approach represents an alternative (and complementary) solution to the scalable MD-SVC system developed in D3.1 and D3.2. The results of the experimental evaluation of the proposed JSCC-methodology are planned to be reported in deliverable D3.4.

This deliverable also proposes an alternate approach to control the transmission errors of SVC packets for WiMAX communications channels. Transmission power is an important resource in wireless links, and its dynamic use may enhance the successful transmission of video packets. In scalable video coding, since different packets and layers have different importance, unequal power allocation (UPA) can be used to provide better protection against varying error prone channel conditions. However, the optimization of the UPA scheme algorithms should be efficient and fast for real time communications. Such UPA scheme will be discussed in section 3.3. The results of experimental evaluation and a simulator are planned to be reported in deliverable D3.4.

Contributors to this deliverable are IBBT and the UoS.

1.2 Objective

The main objective of the JSCC-methodology described in this deliverable is to provide optimized resilience against transmission errors in scalable video streaming over variable-bandwidth error-prone channels. Thereby, the proposed JSCC-methodology minimizes the end-to-end distortion for the transmission of scalable video over packet loss channels and optimally adds redundant forward error correction (FEC) codes to protect the transmitted stream against transmission errors. The proposed approach relies on Lagrangian-based optimization techniques to derive the appropriate protection levels for each layer produced by the scalable source codec. It constructs convex rate-distortion hulls for each frame, irrespective of the target rate. This allows the pre-computation of the convex rate-distortion hulls for typical packet loss channels, such that the extraction of a near-optimal JSCC-allocation can be achieved on the fly for any target rate or packet-loss rate.

The second objective of this deliverable, the development of an unequal power allocation scheme, is to provide a way to minimize the end-to-end distortion of transmitted scalable video based on the characteristics of the wireless channel. The algorithm will try to protect the base layer over the enhancement layers in an efficient way. In this deliverable, the focus will be mainly on quality scalability for CIF resolution sequences.

2 Scalable joint source-channel coding

2.1 INTRODUCTION

Transmission of digital video over heterogeneous, error-prone networks to a large variety of devices requires on-the-fly adaptation of the original encoded bitstream to meet the requirements set by the end-user's connection and the terminal characteristics. This can be efficiently realized by using scalable video coding. In this context, a scalable extension to the state-of-the-art H.264/MPEG-4 AVC video coding standard, also known as SVC (Scalable Video Coding)[1], has been recently developed by the Joint Video Team (JVT) of the ISO/IEC Moving Pictures Experts Group (MPEG) and the ITU-T Video Coding Experts Group (VCEG). This scalable video compression technique provides quality, resolution and frame-rate scalability while its rate-distortion performance is on par with that of single layer H.264.

From a complementary perspective, when transmitting the encoded video over error-prone networks, packets can get lost or corrupted due to congestion, channel fading,… Encoded video bitstreams, such as SVC-streams, are typically very sensitive to information loss as this leads to desynchronization of the entropy decoder and eventually to decoder failure. Adapted error protection mechanisms are therefore of vital importance to protect the compressed bitstream against severe degradations caused by packet losses in the network. Error protection in packet-based systems is typically enabled by means of forward error correction (FEC) codes, which add redundancy bits to the compressed source bits and enable error detection and correction. In this area, many developments have been realized since the invention of capacity-approaching codes like the turbo codes in 1993 by Berrou and Glavieux [2] and the independent re-discovery of the low-density parity-check (LDPC) codes by MacKay in 1995 [3], originally introduced by Gallager in the early 1960's [4]. Driven by these results other similar codes have recently been proposed [5-7].

According to Shannon's separation principle [8], the source and channel coding can be performed separately with optimality. However, this principle only holds under the assumption of asymptotically long block lengths of transmitted data and unlimited complexity and delay. In practical applications, a joint source-channel coding (JSCC), allowing for the optimal allocation of the available bits between the source and channel codes, has shown to provide better results. Also, since the source is encoded by a scalable codec producing layers with different levels of importance, our JSCC-design should incorporate unequal error protection (UEP) [9] of the source packets.

JSCC has been an active area of research in the context of the protection of scalable multimedia material when targeting transmission over error-prone channels. One of the first techniques in this area is the approach of [10] proposing an optimal bit allocation mechanism for the transmission of 3D-subband coded video over binary symmetric channels. In [10], operational distortion functions are derived for each subband and the JSCC-solution relies on a Lagrangian-optimization with two Lagrange multipliers. For video coding based on the discrete cosine transform, JSCC-methods have been for instance proposed in [11, 12] to recover from packet losses. In [11], Reed-Solomon codes are used for channel protection and an optimized rate-allocation technique that relies on multiple description coding principles is proposed. The method of [11] starts from the knowledge of the target rate (or equivalently from the number of fixed-length packets to be transmitted) and develops a Lagrangian-based solution for the JSCC-problem taking into account the channel parameters. In [12], the JSCC-method first generates 'universal rate-distortion characteristics' of the individual layers of the scalable representation. These universal rate distortion characteristics are obtained by performing time-consuming simulations to derive the distortion resulting from independent bit errors on each layer for various bit error probabilities. For scalable wavelet-based coded images, JSCC-methodologies have been proposed in [13-17].

In the context of SVC, to the best of our knowledge, there is only one JSCC-algorithm proposed in the literature, recently presented in [18]. The authors of [18] propose a rate allocation approach for the transmission of SVC-encoded bitstreams in variable-length channel packets over MIMO systems. In [18], the total expected end-to-end distortion is estimated by taking into account the

probabilities of losing temporal and/or FGS layers of each frame. For each layer, a dependent rate-distortion curve is built using an exhaustive sweep over the application layer parameters, i.e. the group of picture (GOP) size, quantization parameters, and physical layer parameters, i.e. symbol constellation and channel protection strength. The end-to-end distortion is thus computed exhaustively for each parameter combination [18], and used in a Lagrangian optimization approach [12] to derive the optimal bandwidth allocation.

In contrast to the approach of [18], where a computationally prohibitive exhaustive search for the optimum solution is performed, in this deliverable, we develop a novel forward programming-based optimal JSCC-methodology based on Lagrangian-optimization and derive a Viterbi-based search algorithm, which (*i*) significantly reduces the computational complexity of the rate allocation and (*ii*) minimizes the end-to-end distortion when transmitting SVC-streams over packet-loss channels. In contrast to [18], in our approach we account for the possible truncation of FGS-layers and use channel packets of fixed-length, since this provides the advantage of an easier cross-layer design [19]. We consider a transmission scheme without any retransmission mechanism, and perform channel coding using state-of-the-art LDPC codes [4]. Compared to previously proposed JSCC-methodologies for scalable bitstreams [10, 11, 13-16], the proposed JSCC-approach does not rely on the knowledge of the target rate. The convex hulls for a certain packet loss rate can therefore be re-used in a bi-section method to meet any desired target bit-rate. Also, we show that our JSCC-methodology can be used in any scalable setting, allowing the determination of an optimized JSCC-solution for any resolution, frame-rate and quality.

This section is structured as follows. Section 2.2 gives an overview of the recently developed SVC-standard, while 2.3 details the channel codes used in our system. Thereafter, in 2.4, we define and propose a solution to the JSCC-allocation problem and derive a Viterbi-base search algorithm, which provides a practical alternative to the classical, computationally prohibitive exhaustive search method. Finally, in 2.5, we provide a detailed discussion of our algorithm, followed by its extension towards achieving full scalability, described in 2.6.

2.2 SOURCE CODING: SCALABLE VIDEO CODING

SVC [1] generates a fully-scalable bit-stream consisting of a base layer, representing the version of the input video with the lowest supported resolution and quality and a set of enhancement layers, representing versions of the input sequence with a higher resolution and/or quality. Figure 1 shows the SVC-encoder structure that provides full scalability.

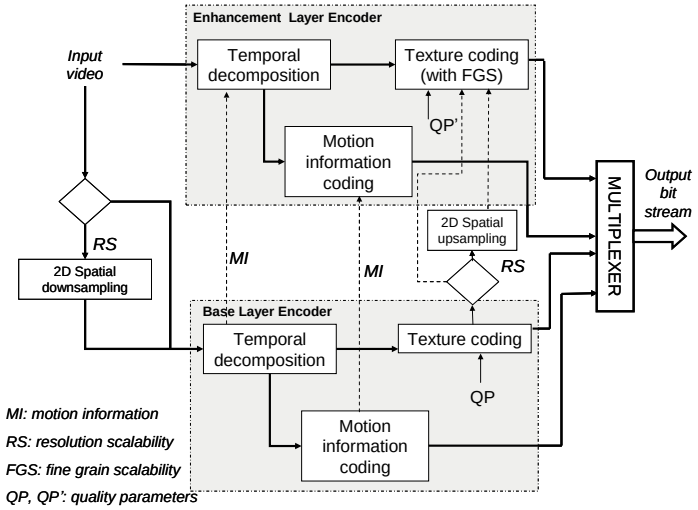

Figure 1: SVC-encoder structure providing spatial, temporal and quality scalable streams

Spatial scalability is achieved by spatial downsampling of the input video and using different encoder loops for each resolution. To encode each quality/resolution layer, the frames are first temporally decomposed by using hierarchical bi-directional motion-compensated prediction [1]. By omitting the lower levels of the hierarchy, temporal scalability can be supported.

Each spatial resolution layer is composed of a base layer and some quality enhancement layers that may be either coarse-grain SNR scalable (CGS), medium-grain SNR scalable (MGS) or fine-grain SNR scalable (FGS) [1]. Although the current draft of the SVC-standard does not include FGS anymore, ongoing core experiments on FGS are being conducted in order to simplify it such that it can be introduced as an amendment to the standard in the future. In this deliverable, we formulate our JSCC-approach for video streams encoded with FGS quality-enhancement layers, and show that our JSCC-methodology can also be applied on CGS and MGS encoded videos.

In case of FGS, up to three quality layers for each spatial resolution layer are supported. These layers are constructed by encoding successive refinements of the transform coefficients, starting with the minimum quality provided by AVC compatible intra/residual coding. This is done by repeatedly decreasing the quantization parameter QP by 6, thereby approximately dividing the quantization step size by two, and applying a modified entropy coding similar to bitplane coding. The FGS-layers only contain refinements for the texture data and provide progressive refinements layer which can be truncated at any byte. Each layer of each slice is stored in a separate network abstraction layer unit (NAL unit). The setup of an SVC-encoder producing a bit-stream with a single resolution and two FGS enhancement layers is shown in Figure 2.

Similar to FGS, the CGS and MGS enhancement layers are constructed by encoding successive refinements of the transform coefficients. However, in contrast with FGS, the CGS and MGS enhancement information are not encoded progressively, meaning that the generated enhancement layers can not be truncated. When encoding CGS enhancement layers, every produced enhancement layer is encoded as a new layer, such that the number of extraction points is limited to the number of layers that have been coded. As the total number of spatial and coarse-grain SNR scalable layers is limited to 8 [26], the number of extraction points is significantly limited. On the other hand, each spatial or CGS layer can have up to three MGS quality-enhancement layers. Hence, with MGS encoding, the number of allowed extraction points can be significantly increased [26] in comparison to CGS encoding.

Similar to the non-scalable H.264, SVC is organized into two conceptual layers: the video coding layer (VCL) and the network abstraction layer (NAL). The VCL defines the syntax elements necessary to efficiently represent the video content, while the NAL formats the VCL representation of the video and provides header information for conveyance by a variety of transport layers [20].

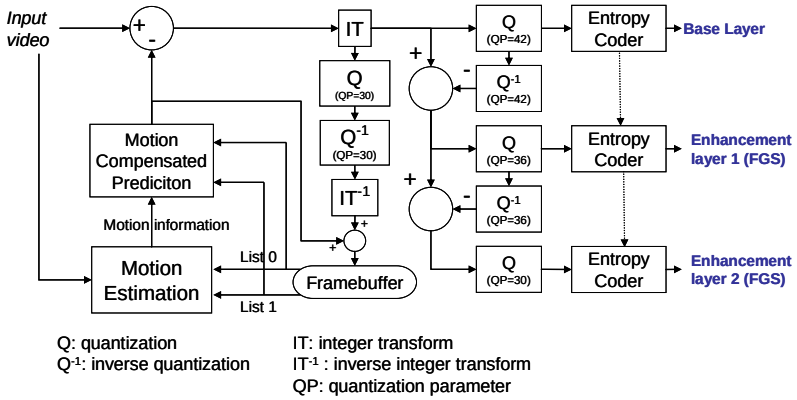

Figure 2: SVC-encoder producing one spatial resolution and two FGS enhancement layers

2.3 CHANNEL CODING: LOW DENSITY-PARITY CHECK CODES

For the protection of source information of size $K \mathrm{x} 1$, we use regular LDPC-codes over GF(2) [21]. An LDPC-code is a linear block code defined by a sparse parity-check matrix H of dimension $N' \mathrm{x} M$ with $M = N' - K$ that allows near-capacity data transmission. The parity-check matrix H of a regular LDPC-code is characterized by a low and fixed number of ones in the rows (also called left or variable degree d_v) and a low and fixed number of ones in the columns (also called right or check degree d_c). The sparsity of the parity-check matrix is a key property that allows for the algorithmic efficiency of the LDPC-codes. In order to construct regular (d_v, d_c) LDPC-codes which

exhibit good performance, we use the Progressive Edge Growth (PEG) algorithm of [22]. This algorithm presents good properties in terms of girth and minimum distance which are dominant criteria to achieve codes with a good error correction performance.

Encoding of the codewords is performed with the dual of the parity-check matrix, called the generator matrix G. We construct the matrix G in systematic form, meaning that the output after encoding includes the original source data. Iterative decoding is performed using a log-domain sum-product algorithm [21]. Since we focus on the transmission of codewords of fixed-length N, for each different protection level, the parity-check and generator matrices with different dimensions are first generated and the last redundant bits are then punctured in such a way that codewords of fixed-length N are achieved (with $N < N'$).

2.4 JOINT SOURCE-CHANNEL CODING

2.4.1 Problem Formulation

We consider a video sequence consisting of F frames that are encoded with a single resolution layer and a number of FGS quality-enhancement layers. Transmission over a packet loss channel with total capacity R_{tot} and packet loss parameter ε is considered. In order to recover from packet losses, we include in our design a row-column bit-interleaver before the actual transmission of the codewords. The interleaver translates the packet loss channel into an equivalent binary erasure channel (BEC) with bit erasure parameter ε, since a packet loss results in bit erasures distributed over the codewords. This is illustrated in the example of Figure 3 where packet number 6 gets lost.

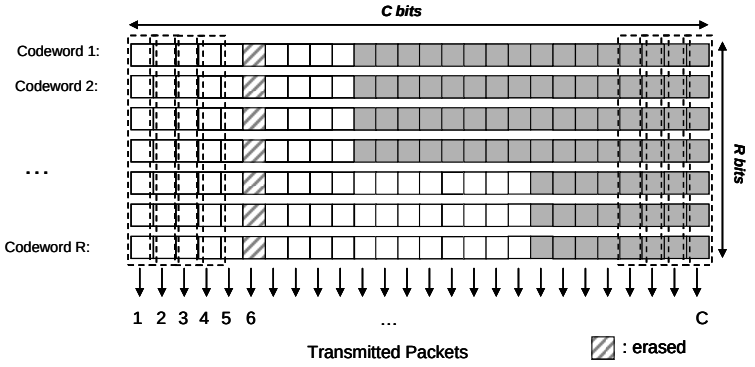

Figure 3: Row-column bit interleaver.

Protection against binary erasures is achieved with d LDPC-codes. We define the code rate r as: $r = k/N$, where k is the number of source bits and N is the total number of bits in a codeword.

The JSCC-method needs to find, for every encoded frame l, the optimum source and channel rates $R_{s,l}$ and $R_{c,l}$ as well as the corresponding total number of codewords M_l of fixed-length N to be transmitted such that the overall expected distortion is minimized. Additionally, for every frame of the video sequence, the optimal protection at the level of every codeword needs to be determined. We define the path Π_{M_l} as the set of code rates $\Pi_{M_l} = \left(r_{l,0}, r_{l,1}, ..., r_{l,M_l}\right)$ assigned to the M_l codewords of frame l, with the convention $r_{l,0} = 0$. Unequal error protection (UEP) can be imposed by forcing the code rates within each frame to be non-decreasing, i.e. $r_{l,0} \le r_{l,1} \le ... \le r_{l,M_l}$.

As in [23], we assume that the total expected distortion of the SVC bitstream can be expressed as a sum of the individual frame distortions. With $R_{s_i,l}$, $R_{c_i,l}$ the source and channel rates respectively used in codeword i of frame l, the JSCC-problem can be defined as the minimization of the distortion:

$$\overline{D_{tot}} = \sum_{l=1}^{F} \overline{D_l}(R_{s,l}, R_{c,l}), \qquad (1)$$

under the constraint that the target bit-rate is met:

$$R = \sum_{l=1}^{F}(R_{s,l} + R_{c,l}) = \sum_{l=1}^{F}\left(\sum_{i=1}^{M_l} R_{s_i,l} + \sum_{i=1}^{M_l} R_{c_i,l}\right) \le R_{tot}\,. \tag{2}$$

2.4.2 Solution

In order to solve the constrained optimization problem defined by (1) and (2), we first derive a recursive formula for the average expected frame distortion. Then, the evolution of the average expected frame distortion is shown to be convex with monotonically decreasing slopes. Finally, an optimal solution for the optimization problem is proposed and a near-optimal solution that significantly reduces the computational complexity is presented.

2.4.2.1 Recursive distortion formulation

The average expected distortion of frame l when transmitting M_l codewords of an SVC–encoded frame is given by:

$$\overline{D_l}(R_{s,l}, R_{c,l}) = \sum_{m=0}^{M_l}\left(\prod_{i=0}^{m}(1 - p_f(r_{l,i}, \varepsilon))\right)\cdot p_f(r_{l,m+1}, \varepsilon)\cdot D_l(\sum_{i=0}^{m} R_{s_i,l})\,, \tag{3}$$

where $p_f(r_{l,i}, \varepsilon)$ denotes the probability of losing codeword i of frame l when transmitted over a BEC with erasure parameter ε. $D_l(\sum_{i=0}^{m} R_{s_i,l})$ is the source distortion given that all codewords up to codeword m were received. Initial parameters are given by $r_{l,0} = 0$, $p_f(r_{l,0}, \varepsilon) = 0$, $p_f(r_{l,M_l+1}, \varepsilon) = 1$ and $R_{s_0,l} = R_{c_0,l} = 0$.

As we consider the transmission of codewords of fixed-length N, $D_l(\sum_{i=0}^{m} R_{s_i,l})$ can be equivalently formulated as a function of the code rates $r_{l,i}$, $0 \le i \le M_l$ of frame l: $D_l(\sum_{i=0}^{m} r_{l,i})$. Equivalently, $\overline{D_l}(R_{s,l}, R_{c,l})$ can be expressed as: $\overline{D_l}(r_{l,0}, r_{l,1}, ..., r_{l,M_l})$. Equation (3) then becomes:

$$\overline{D_l}(r_{l,0}, r_{l,1}, ..., r_{l,M_l}) = \sum_{m=0}^{M_l}\left[\prod_{i=0}^{m}(1 - p_f(r_{l,i}, \varepsilon))\right]\cdot p_f(r_{l,m+1}, \varepsilon)\cdot D_l(r_{l,0} + r_{l,1} + ... + r_{l,m}) \tag{4}$$

Denote $\alpha_{l,m} = \prod_{i=0}^{m}(1 - p_f(r_{l,i}, \varepsilon))$ with initial value $\alpha_{l,0} = 1$ and denote $\tilde{r}_{l,k} = \sum_{i=0}^{k} r_{l,i}$. We can then write: $\alpha_{l,m+1} = \alpha_{l,m}.(1 - p_f(r_{l,m+1}, \varepsilon)) \Leftrightarrow \alpha_{l,m}.p_f(r_{l,m+1}, \varepsilon) = (\alpha_{l,m} - \alpha_{l,m+1})$ for all $m, 0 \le m < M_l$.

A recursive formula for the average expected distortion $\overline{D_l}$ of a reconstructed frame when taking path Π_{M_l} can then be found as follows:

$$\begin{cases} \overline{D_l}(r_{l,0}, r_{l,1}, ..., r_{l,M_l}) = \left[\sum_{m=0}^{M_l-1}(\alpha_{l,m} - \alpha_{l,m+1}).D_l(\tilde{r}_{l,m})\right] + \alpha_{l,M_l}.D_l(\tilde{r}_{l,M_l}) \\ \overline{D_l}(r_{l,0}, r_{l,1}, ..., r_{l,M_l-1}) = \left[\sum_{m=0}^{M_l-2}(\alpha_{l,m} - \alpha_{l,m+1}).D_l(\tilde{r}_{l,m})\right] + \alpha_{l,M_l-1}.D_l(\tilde{r}_{l,M_l-1}) \end{cases}$$

$$\Leftrightarrow \overline{D_l}(r_{l,0}, r_{l,1}, ..., r_{l,M_l}) = \overline{D_l}(r_{l,0}, r_{l,1}, ..., r_{l,M_l-1}) + \alpha_{l,M_l}.\left(D_l(\tilde{r}_{l,M_l}) - D_l(\tilde{r}_{l,M_l-1})\right), \text{ such that finally:}$$

$$\overline{D_l}(\Pi_{M_l}) = \overline{D_l}(\Pi_{M_l-1}, r_{l,M_l}) = \overline{D_l}(\Pi_{M_l-1}) + \alpha_{l,M_l-1}.(1 - p_f(r_{l,M_l}, \varepsilon)).\left(D_l(\tilde{r}_{l,M_l}) - D_l(\tilde{r}_{l,M_l-1})\right). \tag{5}$$

In general, the average expected distortion when sending k codewords for frame l with $1 \le k \le M_l$ is:

$$\overline{D_l}(\Pi_k) = \overline{D_l}(\Pi_{k-1}, r_{l,k}) = \overline{D_l}(\Pi_{k-1}) + \alpha_{l,k-1}.(1 - p_f(r_{l,k}, \varepsilon)).\left(D_l(\tilde{r}_{l,k}) - D_l(\tilde{r}_{l,k-1})\right). \tag{6}$$

The source rate-distortion points $\left(r_{l,k}, D_l(r_{l,k})\right)$ of the SVC stream can be found for each progressive refinement layer as proposed in [23, 24]. The distortion reduction at each quality level q for any frame l with dependency identifier d and temporal level t is calculated as [23, 24]:

$$D(d,q,l) = (\log_{10} D_{ind}(q) - \log_{10} D_{ind}(q-1)) + (\log_{10} D_{dep}(t) - \log_{10} D_{dep}(t-1)).$$

The first term expresses the distortion reduction resulting from decoding quality-layer q, while the second term expresses the distortion reduction resulting from decoding temporal layer t. Conversion to a classical MSE distortion can be achieved by removing the log-function in the above formula. As indicated in [23], the source rate-distortion points $\left(r_{l,k}, D_l(r_{l,k})\right)$ lie on a convex hull with monotonically decreasing slopes.

2.4.2.2 Convexity of the average expected distortion $\overline{D_l}(\Pi_k)$

Using the recursive formula for the average expected frame distortion (6), it can be proven that the average expected distortion $\overline{D_l}(\Pi_k), 1 \le k \le M_l$ of a frame is always convex with monotonically decreasing slopes, no matter what path Π_k is taken. This is formulated in the following lemma.

Lemma 1:

Define $\lambda_{\overline{D_l}}(\Pi_k) = \dfrac{\overline{D_l}(\Pi_{k-1}) - \overline{D_l}(\Pi_k)}{r_{l,k}}, 1 \le k \le M_l$, with the convention $\lambda_{\overline{D_l}}(\Pi_0) = \lambda_{\overline{D_l}}(r_{l,0}) = \infty$. If $D_l(r_{l,k})$ is convex with monotonically decreasing slopes, then $\lambda_{\overline{D_l}}(\Pi_{k-1}) > \lambda_{\overline{D_l}}(\Pi_k)$ for any $k, 1 \le k \le M_l$.

The proof is given in Appendix A(a). A similar conclusion with bounds on the allowable code rates can be drawn when transmitting an increasing number of fixed-length codewords. This is formulated by the following lemma, which is proven in Appendix A(b).

Lemma 2:

Define $\gamma_{\overline{D_l}}(\Pi_k) = \dfrac{\overline{D_l}(\Pi_{k-1}) - \overline{D_l}(\Pi_k)}{k-(k-1)}$, which is the slope of the distortion when sending fixed-length codeword k after $k-1$. Assume that $D_l(r_{l,k})$ is convex and of the form $D_l(r_{l,k}) = \kappa\sigma^2 2^{-2r_{l,k}}$. A sufficient condition for $\gamma_{\overline{D_l}}(\Pi_k)$ to be monotonically decreasing with k is $r_{l,k} > \log_4 e / e \approx 0.2654$ for all $k, 1 \le k \le M_l$.

Based on these lemmas, we will analyze the convexity of the rate-distortion functions involved in the JSCC-problem and propose a Lagrangian-based optimization approach for determining an optimal solution.

2.4.2.3 Optimal JSCC

We observe that for each transmitted number of codewords k, there exists a path Π_k^* such that $\overline{D_l}(\Pi_k)$ is minimal (see Figure 4(a)). However, as illustrated in Figure 4(a), one cannot assume the existence of a *unique* path that provides the best protection for any number of transmitted codewords.

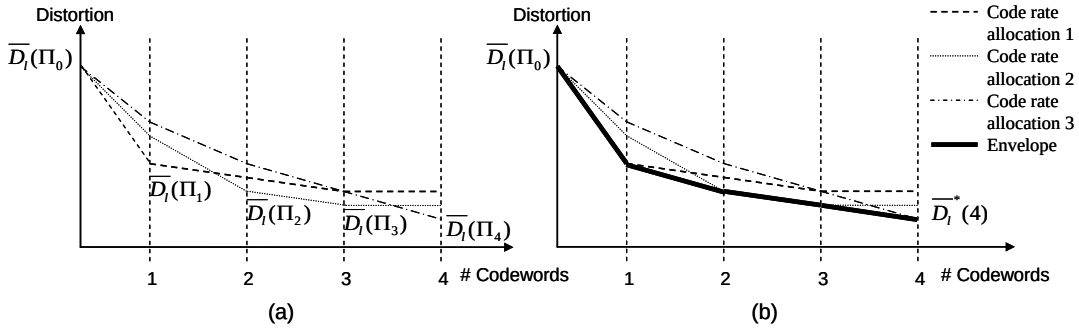

Figure 4: (a) Example of the paths which can be achieved when subsequent codewords are transmitted with different protection levels. (b) Example of the derivation of the virtual envelope from the real convex hulls.

Still, the global optimization problem can be solved by retaining from all constructed convex hulls, the paths which result in a minimal distortion at each subsequently transmitted codeword. Thus, let $\Pi_k^* = \arg\min_{\eta_{l,1},...,\eta_{l,k}} \overline{D_l}(r_{l,1},...,r_{l,k})$ be the path that minimizes the distortion if k codewords are transmitted, and construct the virtual envelope $\overline{D_l}^*(k)$ defined by the rate-distortion points $\left(k, \overline{D_l}(\Pi_k^*)\right), 0 \le k \le M_l$, with the convention $\overline{D_l}(\Pi_0^*) = D_l(0)$. Let Π_{k-1}^*, Π_k^* and Π_{k+1}^* be the paths resulting in a minimal distortion for the transmission of $k-1$, k and $k+1$ codewords respectively. The following holds.

Proposition 1:

Under the conditions of Lemma 2, if $\Pi_{k-1}^* \subset \Pi_k^*$ and $\Pi_{k-1}^* \subset \Pi_{k+1}^*$ for some $k \ge 1$ then the slopes $\gamma_{\overline{D_l}^*}(k), \gamma_{\overline{D_l}^*}(k+1)$ of the virtual envelope in points k and $k+1$ respectively satisfy $\gamma_{\overline{D_l}^*}(k) > \gamma_{\overline{D_l}^*}(k+1)$.

This is proven in Appendix B. We observe that if $\Pi_{k-1}^* \subset \Pi_k^*$ for any $k \ge 1$, then the virtual envelope is convex with monotonically decreasing slopes. However, if there exists k such that Π_{k-1}^* is not included in Π_k^* and Π_{k+1}^*, then the convexity of the virtual envelope is not guaranteed. This is illustrated in Figure 5.

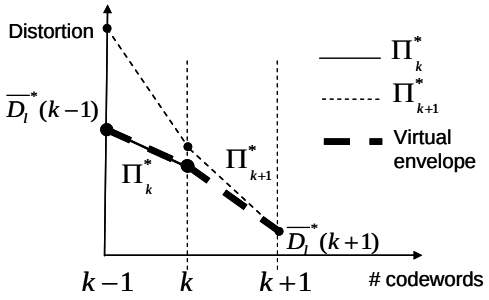

Figure 5: Illustration of the possible non-convexity of the virtual envelope

Table 1: Punctured LDPC-codes used for ε = 5% and ε = 10% packet loss channels. K: source bytes, M: parity bytes, P: punctured bytes, $p_f(\varepsilon)$: probability of loosing a codeword when transmitted over a packet loss channel with parameter ε.

$\varepsilon =$ 5 % losses				$\varepsilon =$ 10 % losses			
K (bytes)	**M (bytes)**	**P (bytes)**	**$p_f(\varepsilon)$ (%)**	**K (bytes)**	**M (bytes)**	**P (bytes)**	**$p_f(\varepsilon)$ (%)**
205	51	154	0.00E+00	194	62	132	0.00E+00
209	47	162	1.00E-06	198	58	140	4.84E-05
211	45	166	2.83E-04	200	56	144	1.21E-03
213	43	170	3.24E-02	202	54	148	3.63E-02
215	41	174	1.83E-01	204	52	152	1.49E-01

.

We analyzed the frequency of occurrence of this non-convex situation. Starting from a source model given by $D(R)=\sigma^2 2^{-2R}$ and setting $p_f(r_{l,i},\varepsilon)$ as given in Table 1 for BECs introducing 5% and 10% of erasures, we derive the virtual envelopes using our recursive formula (6). The resulting slopes of the virtual envelopes are shown in Figure 6. It can be seen that in both situations at some points the slopes of the virtual envelope are not monotonically decreasing. However, globally, the slopes are decreasing with the number of codewords sent. An optimal rate-allocation can therefore be found by (1) removing from the virtual envelope the points that do not lie on a convex hull with monotonically decreasing slopes, and by (2) subsequently applying a classical Lagrangian-optimization method on the produced convex envelopes computed for each frame. The points that are retained define the set of allowable truncation points, similar to JPEG2000's PCRD algorithm [25]. We observe that in each of these truncation points the distortion $\overline{D_l}^*(k)$ is minimal, hence a Lagrangian-based optimization approach operating on the virtual envelopes will lead to a globally-optimal solution.

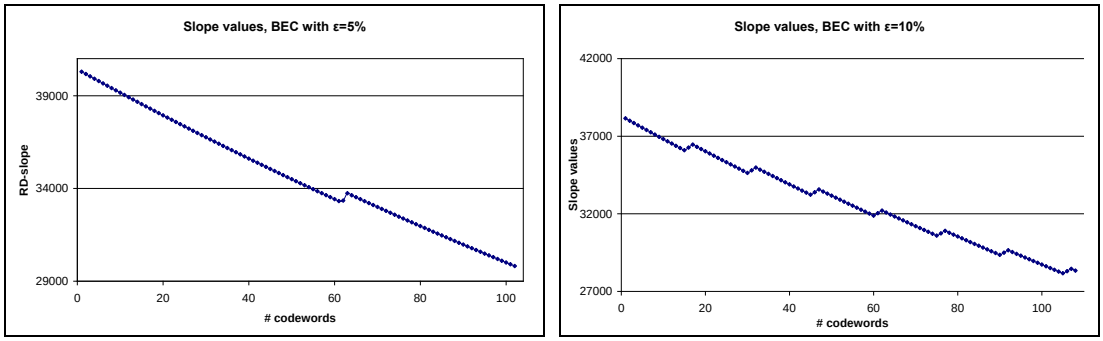

Figure 6: Evolution of the slope of the virtual envelope for a Gaussian source rate-distortion model.

2.4.2.4 Near-optimal JSCC

In order to construct optimal virtual envelopes for each frame l it is necessary to generate all possible code rate combinations and identify the optimum paths Π_k^* for which $\overline{D_l}^*(k)$ are global minima for any k.

From Figure 7 it can be noticed though that even when imposing the UEP constraint ($r_{l,k} \le r_{l,k+1}, \forall k$) this results in an explosive growth of the search space, even for a limited number of protection levels.

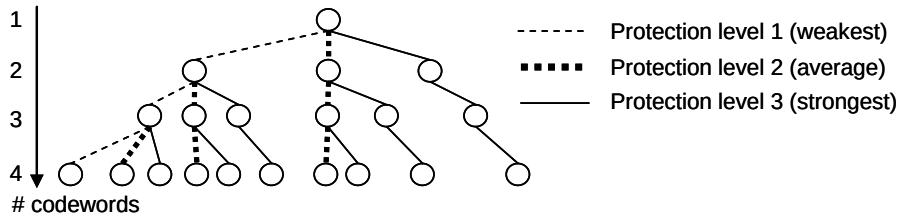

Figure 7: Example of the explosive growth of the search space when using exhaustive search.

In order to reduce the computational complexity associated with an exhaustive search for the optimal solution, we propose the use of a simplified Viterbi-search method, as explained next.

Let R be the set of all possible protection levels $r_{l,k}$ and $d = card(\mathrm{R})$. According to the recursive formula (6), the algorithm starts by evaluating $\overline{D_l}(\Pi_1^{[p]}) = \overline{D_l}(0) - (1 - p_f(r_{l,1}, \varepsilon))\left(D_l(0) - D_l(r_{l,1})\right)$ for all possible values of $r_{l,1}, r_{l,1} \in \mathrm{R}$. The path $\Pi_1^{[p]} = \left(0, r_{l,1}\right), 1 \le p \le d$ corresponds to the *p*-th smallest code-rate $r_{l,1}, r_{l,1} \in \mathrm{R}$. From all $\Pi_1^{[p]}$, the minimal path $\Pi_1^{*} = \arg\min_p\left(\bar{D}_l\left(\Pi_1^{[p]}\right)\right)$ is determined and stored, while $\left(1, \bar{D}_l^{*}(1)\right) = \left(1, \bar{D}_l\left(\Pi_1^{*}\right)\right)$ defines the first rate-distortion point of the virtual envelope. This is illustrated in Figure 8(a) for three protection levels.

For the determination of the second point of the virtual envelope, the algorithm computes $\overline{D_l}(\Pi_2^{[q]})$ based on (6) for all the possible paths $\Pi_2^{[q]} = \left(\Pi_1^{[p]}, r_{l,2}\right), 1 \le q \le d$, where q indicates the *q*-th smallest code-rate $r_{l,2}, r_{l,2} \in \mathrm{R}$, and $\Pi_1^{[p]}, 1 \le p \le d$ are the paths retained at the previous step of the algorithm. One notes that we follow the UEP constraint, that is $r_{l,2} \ge r_{l,1}$, equivalent to $q \le p$. The paths $\Pi_2^{[q]}, 1 \le q \le d$ that lead to the smallest distortions $\overline{D_l}(\Pi_2^{[q]})$ at each protection level q of codeword 2 are stored and used as starting points for the next recursion of the algorithm, as illustrated in Figure 8(b). Also, the minimal path $\Pi_2^{*} = \arg\min_q\left(\bar{D}_l\left(\Pi_2^{[q]}\right)\right)$ is determined and stored, while $\left(2, \bar{D}_l^{*}(2)\right) = \left(2, \bar{D}_l\left(\Pi_2^{*}\right)\right)$ defines the second point on the virtual envelope for frame *l*. The algorithm continues recursively (see Figure 8(c)) to construct the virtual envelope for each frame l until the maximally-available source rate for each frame is exhausted.

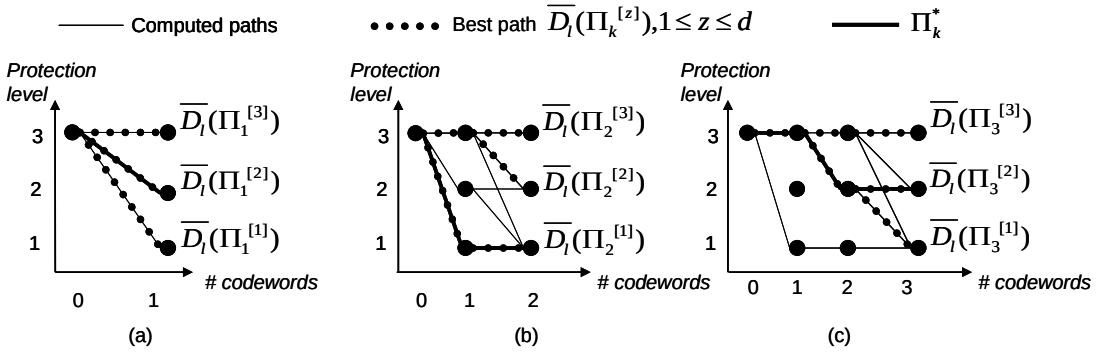

Figure 8: Proposed Viterbi-based search method with three protection levels (1-3: lowest-to-highest protection levels). The thin lines indicate the computed paths at each step of the algorithm; the dotted lines correspond to the best paths leading to each protection level; the bold lines indicate the overall minimum path at each codeword.

An important advantage of our method is that in contrast to previously proposed scalable JSCC-methods for fixed-length codewords [11, 13-16], the computation of the virtual envelopes is independent of the target rate and should therefore be performed only once for a channel with

packet loss parameter ε. One can produce the convex envelopes once for typical channel losses, store them as look-up tables and re-use them for the extraction of any target bit-rate.

2.5 ALGORITHM DISCUSSION

In Figure 9 we show a block diagram of our proposed JSCC-methodology. For each frame $l, 1 \le l \le F$ a convex virtual envelope is constructed from the source rate-distortion points by using equation (6) together with our Viterbi-based algorithm, as described in section 2.4.2.4. Finally, the target rate R_{tot} is method is met by applying a bi-section method on the constructed convex virtual hulls. Since the SVC reference decoder fails if the base layer information is missing and no error control is performed, we guarantee the arrival of the base layer of each frame by protecting it with the highest protection level. This also guarantees a minimum quality at the receiver. If the size of the base layer is not an integer multiple of the K source bytes that are needed for generating the codeword, then the number of base-layer channel codewords that is generated is rounded up and the last codeword is partially filled with bytes of the first FGS layer.

The source distortion at the level of each additional codeword is achieved by linear interpolation of the source rate-distortion points. Using the constructed convex virtual envelopes, a global bi-section search is then performed to meet the desired target rate R_{tot}. This bi-section method can also be applied on each GOP, instead of applying it on the whole sequence, to adapt to the possibly rapid variations in the available channel bandwidth.

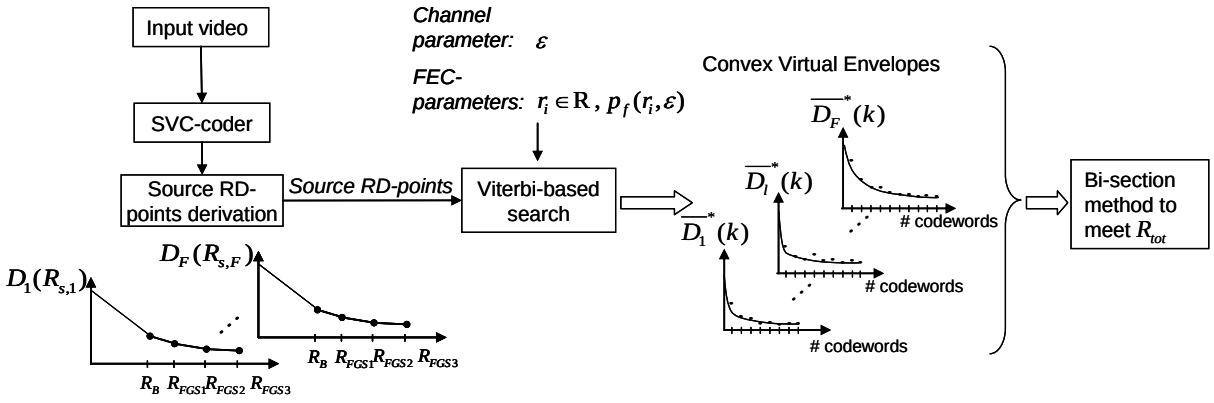

Figure 9: Block diagram of the proposed JSCC-algorithm where the JSCC-allocation uses a Viterbi-based search method requiring the source rate-distortion (RD)-points and the channel and codeword parameters to derive convex virtual envelopes.

2.6 EXTENSION OF THE PROPOSED JSCC-METHOD

In the following we show that the applicability of our proposed JSCC-methodology is not only limited to quality scalable streams with one spatial resolution and multiple FGS layers, but that it can also be used to determine optimized channel protection for SVC streams with temporal, spatial and quality scalability.

2.6.1 Temporal scalability

As explained previously, the proposed JSCC-method constructs convex virtual envelopes for each frame. Therefore, determining an optimized channel protection at a certain target frame-rate and rate R_{tot} is easily supported by considering only the frames that contribute to the targeted frame-rate, and thus by only using the convex virtual envelopes for these frames.

2.6.2 Spatial scalability

Consider a video that is encoded with S spatial layers, where each spatial layer contains one base layer and Q_S FGS enhancement layers. The SVC-standard allows interlayer prediction from the

low-resolution spatial layers to obtain the higher-resolution layers, as illustrated in Figure 10(a) for two spatial layers.

In this resolution-scalable setting, our JSCC-method can be applied to obtain an optimized channel protection as follows. Decoding a spatial resolution $S', S' \leq S$ requires decoding the lower spatial resolutions that have been used in interlayer prediction. In order to avoid the error-propagation across spatial resolutions, the low resolution information used in interlayer prediction needs to be protected using the highest protection level. The quality layers at the lower resolutions that are not used in interlayer prediction are dropped as they do not contribute to the decoding of the video sequence at the targeted resolution.

In the example of $S' = 2$ illustrated in Figure 10(b), the total rate needed to protect the low spatial-resolution is denoted by $R_{res_0}^{protected}$. Only the base layer and FGS layers FGS_1, FGS_2 are protected using the highest protection level as they are used in the prediction of the high-resolution base layer, while FGS_3 is dropped. The remaining rate $R'_{tot} = R_{tot} - R_{res_0}^{protected}$ needs then to be allocated using our JSCC-approach, enabling an optimized channel protection of the higher spatial resolution.

In general, error-propagation across spatial resolutions can be avoided by protecting the lowest $S'-1$ spatial layers using the highest protection level, and apply the proposed JSCC-approach for an optimized channel protection of the targeted spatial-resolution layer S'.

We conclude that our proposed JSCC-method can be employed to derive an optimized protection for any temporal, resolution and quality scalable bit-stream.

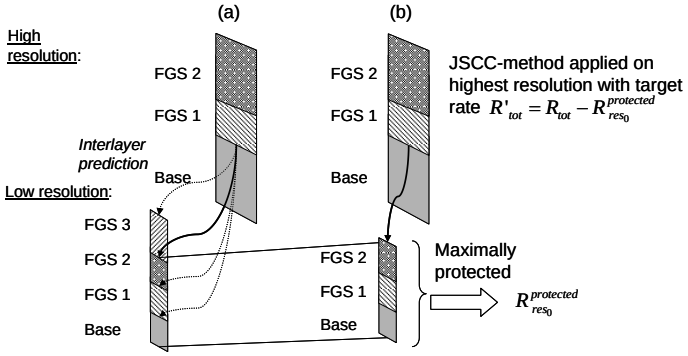

Figure 10: Example of spatial scalability with one low and one high resolution and with each spatial resolution encoded with one base and multiple FGS enhancement layers. The dotted arrows show the interlayer prediction possibilities for the prediction of high spatial resolution from the lower resolution version.

2.6.3 Coarse Grain Scalability and Medium Grain Scalability

We highlight the fact that Coarse-Grain Scalable (CGS) and Medium-Grain Scalable (MGS) streams can also be supported by the proposed JSCC-approach. Despite of the limited granularity and of the fact that the CGS/MGS layers cannot be truncated, our JSCC-method can still be applied, as briefly explained next.

For MGS/CGS encoded video the virtual envelope of each frame is computed as proposed in 2.4.2.4 for FGS-encoded video. We must observe that in this case the assumed source rate-distortion function has a staircase shape, that is, in between two quality layers the source distortion is assumed to be constant. The fact that this function is not convex anymore does not limit the applicability of the proposed JSCC-algorithm. A constant source distortion will lead to a constant expected distortion $\overline{D_l}(\Pi_k)$ (see equation (6)), which is equivalent to obtaining sets of points that are not lying on a convex hull. These points are automatically eliminated by the algorithm. Only when $D_l(\underline{r}_{l,k}) - D_l(\underline{r}_{l,k-1}) \neq 0$ (corresponding to the end of a CGS/MGS layer) we can have $\overline{D_l}(\Pi_k) < \overline{D_l}(\Pi_{k-1})$, resulting into a possible eligible point on the virtual envelope $\overline{D_l}^*(k)$. One

concludes that the algorithm automatically selects truncation points that correspond to the end of the quality layers, and that CGS/MGS streams are supported similar to FGS.

2.7 EXPERIMENTAL RESULTS

The experimental results planned for the proposed JSCC system will be carried out in the immediate future. The results and the optimization of the entire system will be reported in deliverable D.3.4. The testing conditions will include both lossless as well as lossy transmission scenarios. Also, a complexity analysis of the proposed JSCC rate allocation against state-of-the-art is foreseen.

3 Unequal Power Allocation for Video Transmission

3.1 Introduction

In OFDM, the available bandwidth is divided into a number of subcarriers, which are orthogonal to one another. The high data rate channel is split into many low rate data channels. Users share the bandwidth according to their quality of service demand. For a given frequency bandwidth, the time domain is divided into frames which are usually of equal duration. A typical downlink and uplink frame structure used in Wimax is shown in Figure 11. A subchannel in the frame can be used by one user or a group of user. Similarly, data can be transferred to a user over various subchannels. Therefore, in such broadcast wireless communication systems, it is always preferable to precisely control the power of each subchannel in the frame. The transmitted power for each user should be high enough so that reception of data is within acceptable level of quality, yet the power should not be too high to cause unintended interference with other users. The power allocation can also be used as a mean to improve the error resilience and robustness, in particular the use of joint power and system resources has been recommended in the literature [27-29]. The rest of the section discusses the use of power allocation to improve the quality of SVC data.

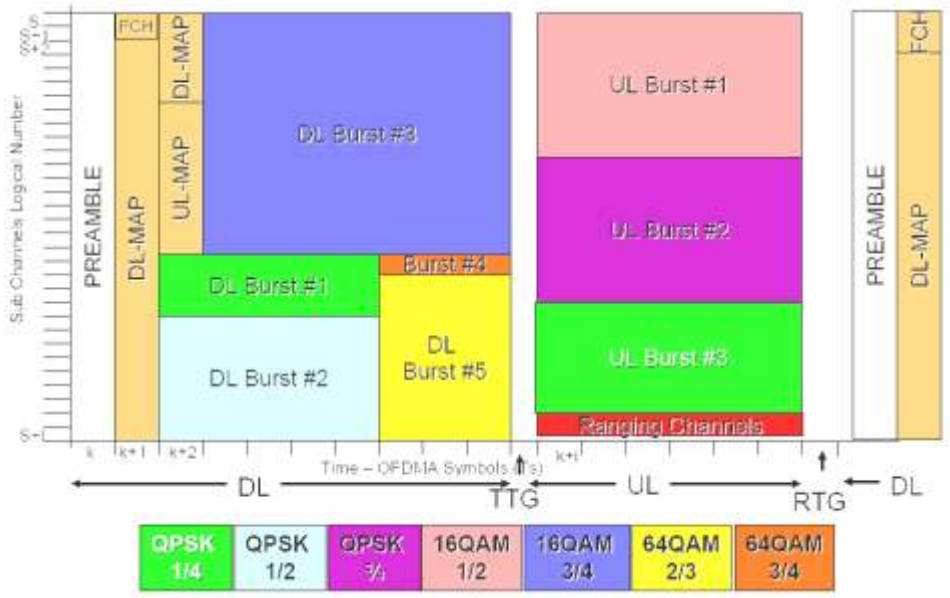

Figure 11 Typical frame structure of Wimax.

3.2 Optimal Power Allocation Schemes

Research in the field of power allocation for wireless links has produced a wide range of schemes, such as optimal allocation of power [30], rate adaptive dynamic power allocation [31], efficient power distribution in OFDM multiuser systems [32]. In these schemes, dynamic resource allocation techniques have been adapted to exploit multiuser diversity of the system. In fixed resource allocation systems, it is not possible to optimize system resources with changing channel conditions. However, the dynamic adaptation of system resources such as power allocation need precise channel estimation, timely feedback information and fast optimization processing. Optimization attempts with more than one variable result in complex formulation statements. Therefore, for real time data transmission, practical allocation of power remains an area of further research.

Power allocation has also been used as error control tool to reduce the interference between the transmitted symbols or frames [33] and to reduce the distortion of the video at the receiver [34]. Power control techniques have often been combined jointly with other resources for error resilience and robustness. Use of power allocation in conjunction with unequal error protection has been

reported in [29]. In [35], joint subcarriers and power allocation algorithms have been reported for OFDM video transmission. The transmit power levels are assigned to various channels subject to acceptable received video quality, providing different levels of services based upon the priority needs of users. However, simultaneous control of transmission rate and power levels requires a framework with fast and accurate algorithm.

3.3 Proposed Unequal Power Allocation Scheme

The power allocation schemes found in the literature provide optimal solutions and effective use of power to protect the important data for transmission of multimedia data over the wireless link. However, the algorithms proposed are computationally complex. In a multiuser environment, assigning unequal power to different subcarriers based upon the importance of layers of the scalable video coded bitstream, with the constraints of the total transmission rate and total transmit power, requires significant processing power. Let R_T be the total bandwidth rate available for a total of M users in the cell. The total transmit power is P_T. Different users may have data rates, and different quality of service requirement such as QCIF, CIF or 4CIF display. However, here we do not consider temporal scalability to avoid further complexity. In the broadcast scenario, same data is transmitted to all users in the cell, but different levels of SNR scalability may be required. The number of subchannels allocated to the m*th* user depends upon the rate and power requirement of the user with the constraints of total power and total bandwidth.

$$\min_{r,\,p} N(r,p) \quad subject\ to \quad \sum_{k=1}^{M} p_k \le P_T \ and \sum_{k=1}^{M} r_k \le R_T \tag{7}$$

Where N is the number of subchannels available to a user. The solution of Equation (7) will result in optimal power distribution to M users in the cell. One way to solve this problem is to minimize the Lagrangian:

$$J(r,p,\lambda_1,\lambda_2) = N(r,\,p) + \lambda_1 \sum_{k=1}^{M} r_k + \lambda_2 \sum_{k=1}^{M} p_k \tag{8}$$

There are many solutions available in the literature to find the solution with one constraint of rate only or with both rate and power constraints. [29;36]. The Lagrangian solution to this optimization problem needs several iterations with suitable assumptions for the convergence of the result. Furthermore, for each user, different layers of video have to be transmitted in different subchannels with transmission rate r_k, and power constraint p_k. For each user receiving more than one video layer, unequal power allocation has to be done such that distortion of received video is minimum. The distortion of received video packet depends upon quantization step, channel condition, frame type, and importance of the layer.

$$D\left(QP, PLR_L, L, f, m\right) = \sum_{j=0}^{m-1} \left[\begin{array}{l} \left(\prod_{i=1}^{L} \left(1 - PLR_i\right) \right) \left(D_{QP}\left(QP, f, m, o_m\right) + D_{EP}\left(f, m, o_m\right) \right) \\ + \left(1 - \prod_{i=1}^{L} \left(1 - PLR_i\right) \right) \times D_{EC}\left(f, m\right) \end{array} \right] \tag{9}$$

Where QP is the quantization parameter, PLR_L is the packet loss rate for the L*th* layer, f is the frame number, m is the macroblock number, D_{QP} is the distortion due to compression, o_m is the mode of block m, D_{EP} is the distortion from error propagation, and D_{EC} is the distortion from error concealment.

It is assumed that the scheme is to operate on pre-encoded schemes, and therefore PLR mainly depends upon the transmission rate, transmission power, interference from other users, and losses in the channel. However, power is the dominant factor in out power allocation problem, and we denote this distortion by D(r,p), where r denotes the rate information. The problem statement to minimize this distortion due to channel characteristics can be defined as:

$$\min_{r,\,p} D(r,p) \quad subject\ to \quad \sum_{k=1}^{M} p_k \le P_{ST}\ and \sum_{k=1}^{M} r_k \le R_{ST} \qquad (10)$$

where P_{ST} and R_{ST} are the total power and rate budget for a user. In the next section, a suboptimal solution to this problem is proposed for fast implementation.

3.3.1 Distribution of Subcarriers Among Users

In an OFDM system, the available bandwidth is divided into various subchannels, which are orthogonal to one another. Users may receive data using different subcarriers. Based upon the requirement of the quality of service, subchannels are assigned to each user. An OFDM system may allocate the sub carrier resources adaptively among the users, based upon the channel conditions. This results in optimal use of the resources amongst its users in time varying wireless channels.

The problem statement is to distribute N number of subcarriers among L layers of video for M users as shown in Figure 12. The transmitter has the information of the channel conditions, which is updated periodically via feedback channels. First N number of subchannels are distributes among M users, according to rate requirement. We keep the equal power for all users to find out fast suboptimal solution, such that:

$$p_{M,N} = \frac{p_{Total}}{N} \qquad (11)$$

Where $P_{M,N}$ is the power of a m*th* user's n*th* subchannel. The rate assigned to a user for an AWGN channel is given by:

$$R_m = \sum_{n=1}^{N} \log_2\left(1 + P_{m,n} \frac{\left|G_{m,n}\right|^2}{N_o \frac{B}{N}}\right) \qquad (12)$$

Where $G^2_{m,n}$ is the n*th* channel gain of m*th* user with noise density of N_o. The overall available bandwidth is given by B, and R_m is the transmission rate of the user m. For M users, the optimization of the power can be expressed as:

$$\max P_{m,n} \sum_{m=1}^{M} \sum_{n=1}^{N} \log_2\left(1 + P_{m,n} \frac{\left|G_{m,n}\right|^2}{N_o \frac{B}{N}}\right) \qquad (13)$$

The solution to this equation may be found using methods given in [31]. However, a suboptimal solution may be found for fast allocation of subchannels to users. The first step of algorithm is the initialization of power levels, and then determination of the number of subchannels for each user in the cell.

a. Initialize the set of users, A = {1,2,3,…M}, and $R_m = 0$
b. For m=1 to M {
    i. Find n satisfying $G_{m,n} \ge G_{m,j}$ for all $j \in A$
    ii. Calculate R_m using (12)
c. While $A \neq \emptyset$
    i. Find m so that $R_m \le R_i$ for all $i \le M$
    ii. For the above m, find n satisfying $G_{m,n} \ge G_{m,j}$ for all $j \in A$
    iii. Update R_m and A using (12), and A=A-{n}

The above algorithm will assign number of subcarriers to each user in the cell, in proportion with the rate requirement. The next section describes the allocation of unequal power to different layers of SVC video based upon the importance of data.

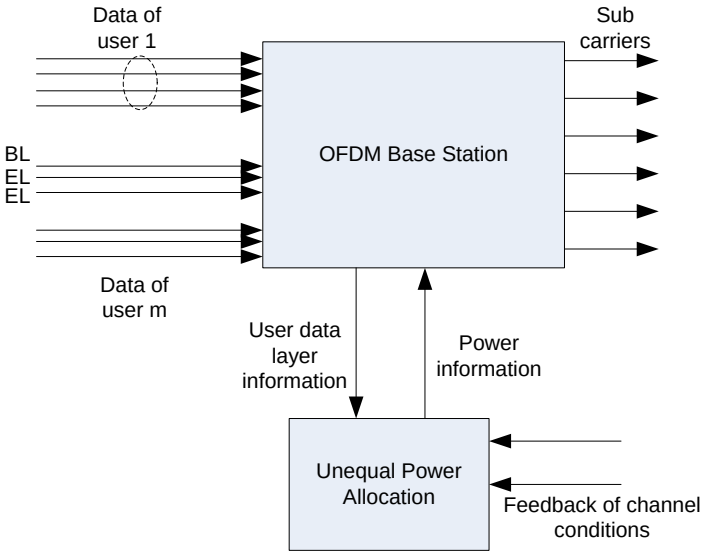

Figure 12 Flow of information in the proposed unequal power allocation scheme

3.3.2 Distribution of Subcarriers of a User Among Video Layers

After all the users have been assigned the subcarriers, each user has its power budget and rate budget. Let's assume $N_{m,s}$ subchannels are available to the m*th* user for transmission of data, from the algorithm of section 3.3.1. An extractor in the base station analyses the importance of the data packet. Base layer data is given priority over enhancement layer data. Based upon the feedback about the packet loss ratio, the UPA algorithm calculates power for the subchannels with the help of a lookup table as shown in Table 2. Such lookup tables will be obtained after simulations of various channel conditions using the SUIT simulator. The table translates the packet loss rate information into the distortion levels, which in turn is used to distribute the power among L video layers for the particular user. The objective of the UPA algorithm at this stage is to minimize the received distortion, and is expressed as:

$$\min_{p} \; Distortion\,(p_{m,s}) \quad subject\ to \quad \sum_{i=1}^{L} P_i \le P_{m,S} \tag{14}$$

where $P_{m,S}$ is the total power budget for user m. We assume all the subcarriers assigned to a user have equal power. This means allocation of two subcarriers to a particular layer doubles the power, and so forth. Hence, power increment is available in fixed step increments of ΔP_s. The UPA algorithm is provided with the information about the number of layers being transmitted to the user, the condition of the subcarrier channels $N_{m,s}$, and the number of subcarriers available.

- Initialize power for all subchannels $N_{m,s}$ to $P_{m,s}/N_{m,s}$
- Distribute video layers equally among the subchannels $N_{m,s}$
- At each feedback interval, distribute the power among layers based upon feedback information of estimated distortion:
    - If $D_{estimated} < D_{required}$
        - Increment $N_{m,s,b}$ according to Table 2
        - Decrement $N_{m,s,e}$ according to Table 2

Where $N_{m,s,b}$ is the number of subchannels assigned to base layer, and $N_{m,s,e}$ is the number of subchannels assigned to enhancement layer.

Table 2 Example of lookup table for distortion of video received at different modulation scheme and channel conditions

SNR Level	PLR %	MCS 5	
		$D_{estimated}$	BL:EL
13.25	0.001	33.4	1:1
11.6	0.01	27.5	1:1
9.95	11.5	14.9	2:1
8.3	40.8	9.8	2:1
6.65	49.5	8.5	3:1
5	50	8.5	4:1

BL:EL = base layer : enhancement layer = ratio of distribution of subchannels between base layer and enhancement layer.

3.3.3 Evaluation Methodology

Figure 13 shows the simulation platform of the system for the evaluation of the unequal power allocation scheme. The video was compressed offline using the SUIT H.264 SVC codec. During transmission, the bit stream on the video server can be accessed by using the bit stream extractor. Rate control is performed at this point or in other words the extracted SVC NAL units are truncated to meet the channel bandwidth requirement. RTP packetizer implements two schemes specified in RFC3984:

1. Single NAL Unit
2. Fragmentation Units Scheme A (FU-A).

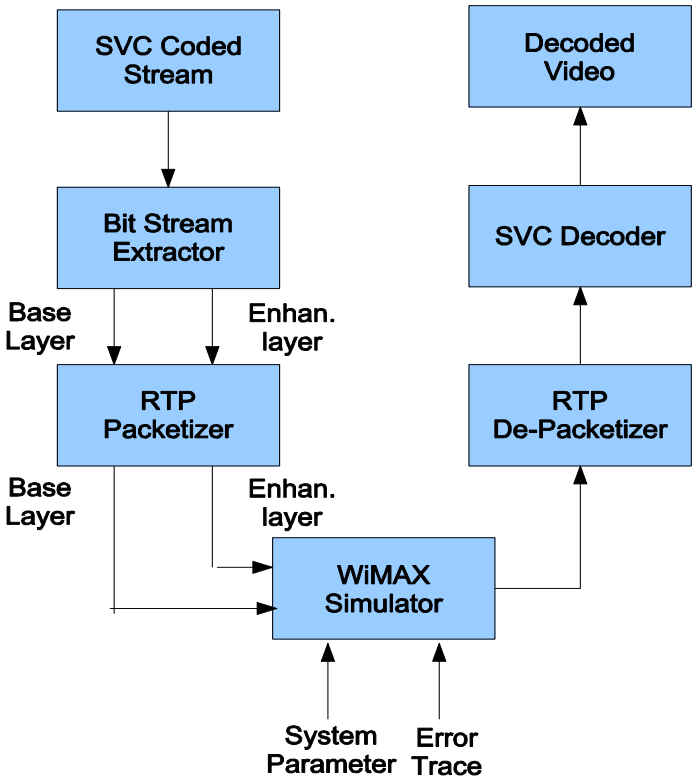

Figure 13 Simulation platform for testing of proposed UPA scheme.

In the simulator, we have adopted a two stage process. The first stage is the generation of an error trace using a WiMAX baseband simulator for a range of SNR values. The details of the simulation procedures are described in [37]. The parameter used for generation of the error trace is in Table 3 and the corresponding Block Error Rate (BLER) curves are shown in the figure.

Table 3 Wimax simulation parameters for error trace

Parameter	Values
Length of Trace (s)	15
Bandwidth	10 MHz
Permutation	PUSC
Channel Coding	CTC
Terminal Speed (km/h)	60
Test Environment	ITU Vehicular A
MCS Mode	16QAM-1/2, 64QAM-1/2

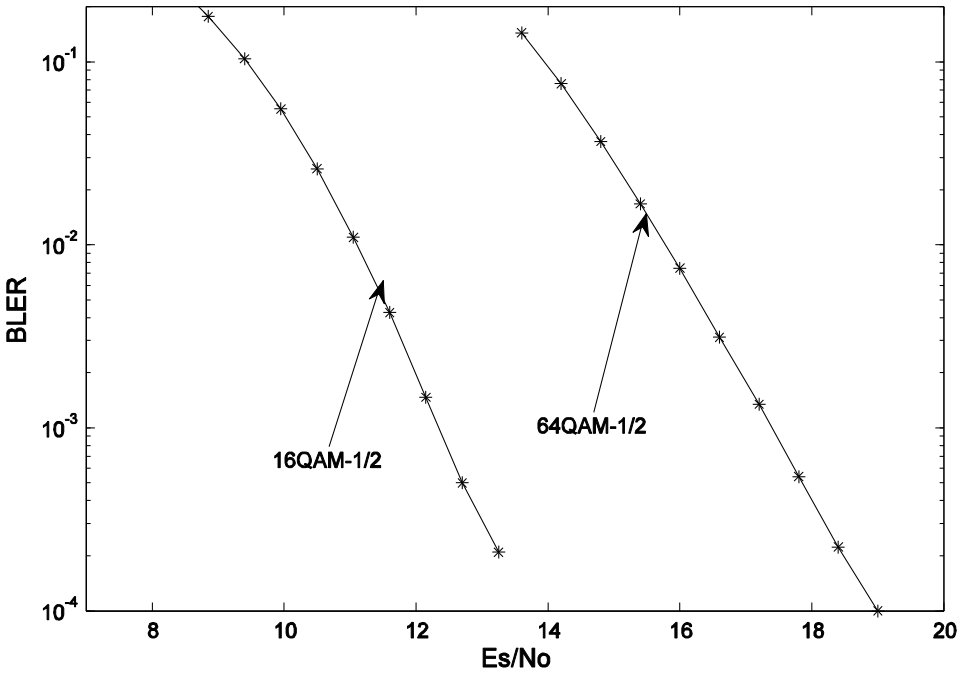

Figure 14 Bock error rate curves for the error trace

In the second stage of the process, we will perform system level simulations, considering the impact of path loss, shadowing loss and multipath loss. The parameters of the system level simulation are shown in Table 4.

With the model above, it is possible to simulate a time varying channel. Note the multipath effect is already simulated during trace generation process using the ITU Vehicular A model. Thus given the total transmit power and the corresponding path loss and shadowing loss, it is possible to calculated the total received power. The total received power can be mapped to an SNR value using thermal noise density in the above table. For the given SNR value, the pre-simulated error trace that has the closest SNR will be used to corrupt the video stream transmitted through the simulator.

For the concept of Unequal Power Allocation (UPA), we can easily calculate the power allocated to two channels, which are then mapped to corresponding SNR values. Similarly, these SNR values

indicate which two error traces to used for corruption of base layer and enhancement layer streams.

Table 4 Parameter list for the system level simulation to be used for testing proposed UPA scheme

Parameter	Values
Path loss model	128.1 + 37.6 x log10(d) dB
Shadowing loss model	Log-normal with variance of 64 dB
Multipath loss model	ITU Vehicular A
Carrier Frequency	2GHz
Bandwidth	10MHz
BS Antenna Gain	14 dBi
MS Antenna Gain	0 dBi
MS Noise Figure	9 dB
Thermal Noise Density	-174 dBm / Hz
BS Transmit Power	43 dBm

3.4 Results

The proposed UPA scheme has been tested with two CIF sequences ‘foreman’ and ‘city’, with a limited set of distortion tables. The sequences are encoded using the SUIT encoder, with 2 SNR scalable layers; including one base layer and one enhancement layer. The WiMAX error patterns, developed in SUIT, have been used to simulate wireless link. The ‘quantexp’ parameter is set to 7, so that enhancement layer has a significant effect on the decoded output. To simulate the UPA scheme, a packet error trace program has been used, that accepts two error trace files, and drops the NAL units in case of any error in the packet. In order to compare the psnr results, frame error concealment has been implemented at the decoder. In case of an error in the base layer NALU, picture data is copied from the previous frame to the current frame for the affected area.

Figure 15 shows the test results for the ‘foreman’ and ‘city’ sequences, with and without the proposed UPA scheme for a WiMAX simulated channel with a modulation scheme of MCS5. For test purposes, we have kept the ΔP_s step size equal to 1.65. However, in future simulations, detailed sets of lookup tables similar to Table 2, will be used. As can be seen in Figure 15, the performance of the UPA scheme is up to 10 dBs better than EEP transmission at the same transmission rate. The main reason for this because the proposed UPA scheme protects the base layer effectively, which results in more graceful degradation of decoded video performance in error prone channel conditions.

The proposed unequal power allocation scheme will be tested more thoroughly so that more subjective and objective results can be included in D3.4. Further work may be carried out to include the temporal scalability scenarios. The main emphasis will be to provide acceptable optimal quality of service at least for the base-layer for all users in the cell

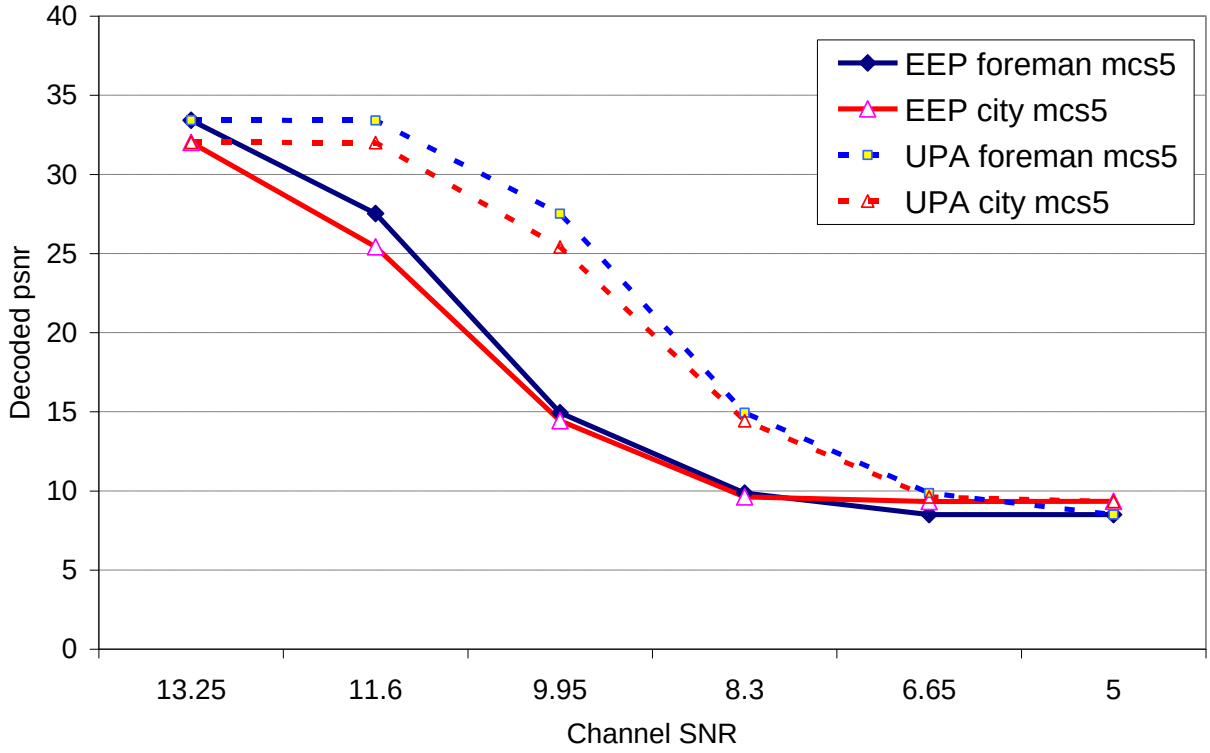

Figure 15 Performance comparison of the UPA écheme for 300 frames of CIF sequences 'foreman' and 'city', The ΔP_s = 1.65, and msc5 Wimax error trace is used with simple frame concealment.

.

4 Conclusions

In Section 2, we have presented a Lagrangian-based optimization JSCC-methodology for transmission in fixed-length packets of SVC encoded video over packet loss channels. A novel approach for solving the SCC-problem is proposed, wherein the convex rate-distortion hulls are constructed for each frame, irrespective of the target rate. We have shown that the presented JSCC-solution can be used in any scalable setting, allowing the determination of an optimized channel protection for resolution, frame-rate and quality scalable streams. Compared to previously proposed JSCC-techniques, our algorithm can be used to pre-compute the convex rate-distortion hulls for typical packet loss rates. Adaptation to fluctuating target rates or packet-loss rates can be achieved on the fly, thereby significantly lowering the computational complexity of the rate-allocation in the JSCC-system. We conclude that the proposed JSCC-approach retains all the scalability functionalities of SVC and enables optimized error-resilience in error-prone transmission conditions.

Ongoing work aims at (1) evaluating experimentally the proposed JSCC system both in lossless and lossy transmission conditions, (2) assessing its complexity against state-of-the-art, and (3) optimizing the entire system. The results of these investigations are planned to be reported in deliverable D.3.4.

An unequal error protection scheme for SVC video packets has been proposed in Section 3, which allocates power unequally to the different NALUs. In OFDM, a number of subchannels are available to transmit the video packets. The proposed UPA scheme first allocates subchannels to users based upon the quality of service requirements. Then subcarriers are distributed among the layers of a user based upon the importance of the layers and the total power constraints. The UPA algorithm updates the power allocation periodically using feedback information about the channel conditions. In contrast to available power optimization solutions in the literature, the proposed scheme uses a lookup table for fast allocation of power. The lookup tables translates the estimated video distortion into power levels. Such tables will be generated by using the SUIT simulator for various channel conditions and modulation schemes. Some initial test results show significant improvements in decoded video quality. Further results will be reported in D3.4.

5 Acronyms

AVC: Advanced Video Coding

BEC: Binary Erasure Channel

BL: Base Layer

BLER: Block Error Rate

BS : Base Station

CGS: Coarse-Grain Scalable

EBCOT : Embedded Block Coding with Optimized Truncation

EEP: Equal Error Protection

EL: Enhancement Layer

FGS: Fine-Grain Scalable

GOP: Group Of Pictures

JPEG : Joint Photographic Experts Group

JSCC : joint source-channel coding

JSVM: Joint Scalable Video Model

JVT : Joint Video Team

LDPC: Low-Density Parity-Check

MGS: Medium-Grain Scalable (MGS)

MPEG : Moving Picture Experts Group

MSE : Mean Square Error

NAL: Network Abstraction Layer

OFDM: Orthogonal Frequency-Division Multiplexing

PEG : Progressive Edge Growth

PSNR : Peak Signal to Noise Ratio

QAM: Quadrature Amplitude Modulation

QLA : Quality Level Assigner

QP: Quantization parameter

RD: Rate Distortion

SEI : Supplemental Enhancement Information

SVC : scalable video coding

UEP : unequal error protection

UPA : Unequal Power Allocation

VCEG : Video Coding Experts Group

VCL : Video Coding Layer

6 Appendices

6.1.1 Appendix A

6.1.1.1 Proof Lemma 1

Using our recursive formula we can state:

$$\lambda_{\overline{D}_l}(\Pi_k) = \frac{\overline{D_l}(\Pi_{k-1}) - \overline{D_l}(\Pi_k)}{r_{l,k}} = \alpha_{l,k-1} \cdot (1 - p_f(r_{l,k}, \varepsilon)) \cdot \frac{\left[D_l(\tilde{r}_{l,k-1}) - D_l(\tilde{r}_{l,k}) \right]}{\tilde{r}_{l,k} - \tilde{r}_{l,k-1}}$$

$$\Leftrightarrow \lambda_{\overline{D}_l}(\Pi_k) = \alpha_{l,k-1} \cdot (1 - p_f(r_{l,k}, \varepsilon)) \cdot \lambda_{D_l}(\tilde{r}_{l,k})$$

$$\Leftrightarrow \lambda_{\overline{D}_l}(\Pi_k) = \alpha_{l,k} \cdot \lambda_{D_l}(\tilde{r}_{l,k})$$

Since $1 - p_f(r_{l,k}, \varepsilon) \le 1, \forall r_{l,k}$, it implies that $\alpha_{l,k} = \alpha_{l,k-1}.(1 - p_f(r_{l,k}, \varepsilon)) \le \alpha_{l,k-1}$ and because $D_l(\tilde{r}_{l,k})$ is convex with monotonically decreasing slopes: $\lambda_{D_l}(\tilde{r}_{l,k}) = \lambda_{D_l}(\tilde{r}_{l,k-1} + r_{l,k}) < \lambda_{D_l}(\tilde{r}_{l,k-1})$. We obtain that $\lambda_{\overline{D}_l}(\Pi_{k-1}) > \lambda_{\overline{D}_l}(\Pi_k)$, meaning that the average expected distortion $\overline{D_l}(\Pi_{M_l})$ of a frame is always convex with monotonically decreasing slopes. Q.E.D.

6.1.1.2 Proof Lemma 2:

We observe that $\gamma_{\overline{D}_l}(\Pi_k) = \lambda_{\overline{D}_l}(\Pi_k) \cdot r_{l,k}$. Hence, $\frac{\gamma_{\overline{D}_l}(\Pi_{k-1})}{\gamma_{\overline{D}_l}(\Pi_k)} = \frac{\lambda_{\overline{D}_l}(\Pi_{k-1})}{\lambda_{\overline{D}_l}(\Pi_k)} \cdot \frac{r_{l,k-1}}{r_{l,k}}$. From the proof of lemma 1 we know that $\lambda_{\overline{D}_l}(\Pi_k) = \alpha_{l,k-1}.(1 - p_f(r_{l,k}, \varepsilon)) \cdot \lambda_{D_l}(\tilde{r}_{l,k})$, which implies that:

$$\frac{\gamma_{\overline{D}_l}(\Pi_{k-1})}{\gamma_{\overline{D}_l}(\Pi_k)} = \frac{\lambda_{D_l}(\tilde{r}_{l,k-1})}{\lambda_{D_l}(\tilde{r}_{l,k})} \cdot \frac{r_{l,k-1}}{r_{l,k}} \cdot \frac{1}{1 - p_f(r_{l,k}, \varepsilon)}.$$

Also, the slope $\lambda_{D_l}(\tilde{r}_{l,k})$ can be approximated by the derivative of $D_l(\tilde{r}_{l,k})$. If $D_l(\tilde{r}_{l,k}) = \kappa\sigma^2 2^{-2\tilde{r}_{l,k}}$, then its derivative is given by: $D_l'(\tilde{r}_{l,k}) = -2\kappa\sigma^2 2^{-2\tilde{r}_{l,k}} \ln(2)$, implying that:

$$\frac{\gamma_{\overline{D}_l}(\Pi_{k-1})}{\gamma_{\overline{D}_l}(\Pi_k)} = 2^{2(\tilde{r}_{l,k} - \tilde{r}_{l,k-1})} \cdot \frac{r_{l,k-1}}{r_{l,k}} \cdot \frac{1}{(1 - p_f(r_{l,k}, \varepsilon))} = 2^{2r_{l,k}} \frac{r_{l,k-1}}{r_{l,k}} \cdot \frac{1}{1 - p_f(r_{l,k}, \varepsilon)}$$

From this, a sufficient condition such that $\gamma_{\overline{D}_l}(\Pi_k)$ is monotonically decreasing with k is $2^{2r_{l,k}} \frac{r_{l,k-1}}{r_{l,k}} \ge 1$. Let $r_{l,k} = \beta r_{l,k-1}$. This implies: $4^{r_{l,k}} > \frac{r_{l,k}}{r_{l,k-1}} = \beta \Leftrightarrow 4^{\beta \cdot r_{l,k-1}} > \beta \Leftrightarrow f(\beta) = \frac{\log_4 \beta}{\beta} < r_{l,k-1}$

A simple derivation shows that the maximum of $f(\beta)$ is achieved when $\beta = e$. This means that $r_{l,k} > \log_4 e / e \approx 0.2654$ which ends the proof. Q.E.D.

6.1.2 Appendix B

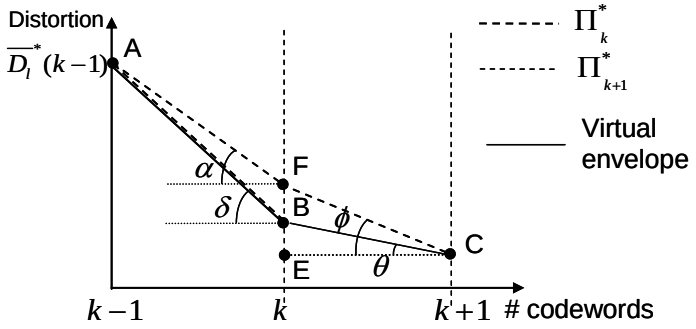

Figure 16: Convexity of the virtual envelope.

Proof:

From $\Pi_{k-1}^* \subset \Pi_k^*$ and $\Pi_{k-1}^* \subset \Pi_{k+1}^*$ we deduce that the virtual envelope defined by the points $(p, \overline{D_l}^*(p)), 0 \le p \le k+1$ coincides with the minimal path Π_{k-1}^* for all $p, 0 \le p \le k-1$, and is constructed using the minimal paths Π_k^*, Π_{k+1}^* for $k \le p \le k+1$. This is illustrated in Figure 16 for $k-1 \le p \le k+1$, where the envelope corresponds to the lines AB, BC. Let $\delta = \gamma_{\overline{D_l}^*}(k)$ and $\theta = \gamma_{\overline{D_l}^*}(k+1)$ be the slopes of the virtual envelope in points B and C respectively, and α, ϕ be the slopes of path Π_{k+1}^* at k and $k+1$ transmitted codewords respectively. We need to prove that $\theta < \delta$. One notes that $\theta = \phi$ and $\alpha = \delta$ correspond to the particular case $\Pi_k^* \subset \Pi_{k+1}^*$. In general, $\theta \le F\hat{C}E = \phi$ and $\phi < \alpha$ (cf. lemma 2) and $\alpha \le \delta$ (since $\overline{D_l}^*(\Pi_{k+1}^*) \ge \overline{D_l}^*(\Pi_k^*)$ and $\Pi_{k-1}^* \subset \Pi_k^*$, $\Pi_{k-1}^* \subset \Pi_{k+1}^*$), implying that $\theta < \delta$, which ends the proof. Q.E.D.

7 References

[1] T. Wiegand, G. J. Sullivan, J. Reichel, H. Schwarz, and M. Wien, Joint Draft 7 of SVC Amendment (revision 2), Klagenfurt, Austria, JVT-T201, 15–21 July, 2006.

[2] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near shannon limit error-correcting coding and decoding: Turbo codes," in Proc. IEEE International Conference on Communications, Geneva, Switzerland, vol. 2, pp. 1064-1070, 23-26 May 1993.

[3] D. MacKay and R. Neal, "Good Codes based on Very Sparse Matrices," Lecture Notes in Computer Science, vol. 1025, pp. 100-111, 1995.

[4] R. Gallager, Low-Density Parity-Check Codes, Massachusetts Institute of Technology, PhD thesis, 1963

[5] H. Jin, A. Khandekar, and R. McEliece, "Irregular Repeat Accumulate Codes," in Proc. 2nd International Symposium on Turbo codes and Related Topics, Brest, France, vol., pp. 1-8, September 2000.

[6] A. Shokrollahi, "Raptor codes," in Proc. IEEE International Symposium on Information Theory, vol., pp. July 2003 2003.

[7] M. Luby, "LT- codes," in Proc. 43rd Annual IEEE Symposium on the Foundations of Computer Science (STOC), vol., pp. 271-280, 2002.

[8] C. E. Shannon, "A mathematical theory of communication," Bell System Technical Journal, vol. 27, pp. 379-423, July 1948.

[9] A. Albanese, J. Blömer, J. Edmonds, M. Luby, and M. Sudhan, "Priority Encoding Transmission," IEEE Transactions on Information Theory, vol. 42, pp. 1737-1744, November 1996.

[10] G. Cheung and A. Zakhor, "Joint source/channel coding of scalable video over noisy channels," Proc. IEEE International Conference on Image Processing, vol. 3, pp. 767-770, 16-19 September 1996.

[11] R. Puri and K. Ramchandran, "Multiple Description Source Coding Through Forward Error Correction Codes," in Proc. 33rd Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, vol. 1, pp. 342 - 346, October 1999.

[12] L. Kondi, F. Ishtiaq, and A. Katsaggelos, "Joint Source-Channel Coding for Motion-Compensated DCT-based SNR Scalable Video," IEEE Transactions on Image Processing, vol. 11, September 2002.

[13] R. Hamzaoui, V. Stankovic, and Z. Xiong, "Fast algorithm for distortion-based error protection of embedded image codes," IEEE Transactions on Image Processing, vol. 14, pp. 1417-1421, October 2005.

[14] Z. Wu, R. Jandhyala, B. Ali, and M. W. Marcellin, "Joint Source/Channel Coding for Multiple Video Sequences With JPEG2000," IEEE Transactions on Image Processing, vol. 14, pp. 1020-1032, August 2005.

[15] B. A. Banister, B. Belzer, and T. R. Fischer, "Robust image transmission using JPEG2000 and turbo-codes," IEEE Signal Processing Letters, vol. 9, pp. 117-119, April 2002.

[16] S. Dumitrescu, X. Wu, and Z. Wang, "Globally optimal uneven error-protected packetization of scalable code streams," IEEE Transactions on Multimedia, vol. 6, pp. 230 - 239, April 2004.

[17] A. Munteanu, M. R. Stoufs, J. Cornelis, and P. Schelkens, "Scalable and Channel-Adaptive Unequal Error Protection of Images with LDPC Codes," Lecture Notes in Computer Science, vol. 4179/2006, pp. 722-733, 2006.

[18] M. K. Jubran, M. Bansal, R. Grover, and L. P. Kondi, "Optimal Bandwidth Allocation for Scalable H.264 Video Transmission Over MIMO Systems," in Proc. MILCOM2006, Washington, USA, vol., pp. 1-7, 23-25 October 2006.

[19] R. Hamzaoui, V. Stankovic, and Z. Xiong, "Optimized Error Protection of Scalable Image Bit Streams," IEEE Signal Processing Magazine, vol. 22, pp. 91-107, November 2005.

[20] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC Video Coding Standard," IEEE Transactions on Circuits and Systems for Video Technology, vol. 13, pp. 560 - 576, July 2003.

[21] S. Lin and D. J. Costello, Error Control Coding-Fundamentals and Applications, 2nd ed., Pearson, Prentice Hall, 2004.

[22] X.-Y. Hu, E. Eleftheriou, and D.-M. Arnold, "Regular and irregular progressive edge-growth tanner graphs," IEEE Transactions on Information Theory, vol. 51, pp. 386-398, January 2005.

[23] I. Amonou, N. Cammas, S. Kervadec, and S. Pateux, "Optimized Rate-Distortion Extraction with Quality Layers," in Proc. International Conference on Image Processing, Atlanta, GA, USA, vol. ICIP2006, pp. 173-176, October 2006.

[24] J. Reichel, H. Schwarz, and M. Wien, Joint Scalable Video Model JSVM-6, Draft Output Document from JVT, JVT, Geneva, Switzerland JVT-S202, April 2006.

[25] D. S. Taubman and M. W. Marcellin, JPEG2000: Image Compression Fundamental, Standards and Practice Dordrecht, Kluwer Academic Publishers, 2001.

[26] I. Amonou, N. Cammas, S. Kervadec, and S. Pateux, Enhanced SNR scalability for layered CGS coding using quality layers, Geneva, Switzerland, Input Document to JVT JVT-S044, 1-7 April 2006.

[27] T. Bruggen and P. Vary, "Unequal error protection by modulation with unequal power allocation," *Communications Letters, IEEE*, vol. 9, no. 6, pp. 484-486, 2005.

[28] E. Maani and A. K. Katsaggelos, "Joint source-channel coding and power allocation for video transmission over wireless fading channels,", 2 ed 2006, pp. 949-953.

[29] Z. Shengjie, X. Zixiang, and W. Xiaodong, "Joint error control and power allocation for video transmission over CDMA networks with multiuser detection," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 12, no. 6, pp. 425-437, 2002.

[30] C. Pirak, Z. J. Wang, K. J. R. Liu, and S. Jitapunkul, "Optimum Power Allocation for Maximum-Likelihood Channel Estimation in Space-Time Coded MIMO Systems,", 4 ed 2006, p. IV.

[31] S. Zukang, J. G. Andrews, and B. L. Evans, "Adaptive resource allocation in multiuser OFDM systems with proportional rate constraints," *Wireless Communications, IEEE Transactions on*, vol. 4, no. 6, pp. 2726-2737, 2005.

[32] S. Zukang, J. G. Andrews, and B. L. Evans, "Optimal power allocation in multiuser OFDM systems,", 1 ed 2003, pp. 337-341.

[33] G. Zihua and K. B. Letaief, "Optimal power allocation for multirate CDMA systems with interference cancellation,", 6 ed 2001, pp. 1885-1889.

[34] A. Karaer and C. N. Georghiades, "Optimum Bit-by-Bit Power Allocation for Minimum Distortion Transmission,", 4 ed 2006, pp. 1616-1621.

[35] C. Mohanram and S. Bhashyam, "A sub-optimal joint subcarrier and power allocation algorithm for multiuser OFDM," *Communications Letters, IEEE*, vol. 9, no. 8, pp. 685-687, 2005.

[36] V. Chande and N. Farvardin, "Progressive transmission of images over memoryless noisy channels," *Selected Areas in Communications, IEEE Journal on*, vol. 18, no. 6, pp. 850-860, 2000.

[37] "WIMAX Modelling, SUIT Deliverable Document 208,"Jan.2007.