Designing DCT Source Significance Information for Efficient and Robust Transmission

Pasteur Poda†, ††, *, Ahmed Tamtaoui †, El Houssine Bouyakhf††

† National Institute of Posts and Telecommunications Images Processing Laboratory Avenue Allal Al Fassi, Madinat Al Irfane 10100, Rabat, Morocco
†† Mohamed V University - Agdal Faculty of Sciences, LIMIARF 4 Avenue Ibn Battouta B.P. 1014 RP, Rabat, Morocco

Abstract

In this paper, we deal with an unequal error protection (UEP) method for DCT-based image or video coding and transmission. The source significance information (SSI) of the DCT data is designed not only to get profit of the intrinsic DCT sources contents distribution but also to meet a desired reconstruction quality of the image/video provided with a certain error protection. An adaptive algorithm is designed to achieve the elaboration of the SSI into levels of significance that rate compatible punctured convolutional codes protect against the channel perturbations. Simulations have concerned several images which are DCT coded and transmitted. The obtained results showed that the proposed method outperforms, in terms of efficiency, the classical equal error protection (EEP) and an existing UEP approach of DCT images transmission. Moreover the proposed UEP is as robust as the maximum error protection strategy done through the classical EEP.

Keywords: Discrete Cosine Transform, Human Visual System, Unequal Error Protection

* Correspondence: pasteurpoda@yahoo.fr
1 Introduction

In actual communication systems the transmission channel is noisy and has limited bandwidth that, moreover, has to be shared between several users. A classical solution of the digital communication theory has been to optimize both in source and channel coding of the information source. More recently, joint source-channel coding approaches have proved that better can be done in a simple way. Thus, one of these approaches is the hierarchical protection that includes the unequal error protection (UEP) method [2].

In the classical communication systems, no hierarchy of the image source information is assumed and the information bits are equally protected against channel errors. This way of proceeding is often referred to equal error protection (EEP) [8]. UEP has proved that useless protection can be avoided by protecting better the most significant information bits than the least ones. That is very interesting because of the economy of the bits resources which may result from without any loss on the image reconstruction quality.

Several approaches of UEP are encountered in the literature. Some UEP designs are integrated in an optimal tandem source-channel rate allocation method [8] or in a subband coding approach [16, 17]. In these designs, source and channel bits are distributed following a joint design of a rate-distortion function. In [5], UEP is integrated as a part of a rate control system where the elaboration of an error sensitivity metric provides the authors with the support for UEP. In [10], the video source is partitioned into GOPs (group of pictures) and within each GOP a partitioning into three levels of significance provides source significance information (SSI) material to the authors for an UEP of the pictures associated with each level of significance. In [18], the SSI material for UEP is supplied by an adaptive segmentation into blocks of variable sizes in the image spatial domain. In [22], the SSI support for the UEP running relies on the partitioning of the DCT frequencies into two levels of significance based on the natural unequal significance that exists between its DC and its AC coefficients. Recent researches on UEP deal with its improvement. Accordingly, in [3] a turbo coding UEP scheme is enhanced with a joint system of hierarchical quadrature amplitude modulation. The authors of [11] innovated in the achievement of UEP by proposing a system of unequal power allocation.

In this paper, we pursue the goal of improving, as in efficiency as in robustness, the UEP approach as proposed in [22]. We assume that, whenever the parameters of the channel change, the DCT SSI might vary to meet the desired reconstruction quality of the transmitted images. Thus we would like, by the means of an algorithm, to partition the DCT source of any image intended to be coded and transmitted into a structure of three hierarchical levels. Moreover, we assign the two most significant levels not to be always disjointed so that the SSI design can stand into two or three hierarchical levels to meet optimality. We also expect that the maximum transmission robustness yielded by the maximum error protection strategy with EEP be reachable with lower bits resources costs.
The proposed UEP design for image/video DCT data coding and transmission promises interesting applications in images or video communication over wired and wireless transmission environments. In wired transmission the channel is usually well known enabling a definitive design of an optimal SSI. In wireless environments the proposed algorithm might match very well, in a joint design, with the availability of channel state information (CSI) [7].

The remainder of this paper is built around four sections. The section 2 deals with the discrete cosine transform (DCT) and its current applications. Then, follow the proposed approach and algorithm. In the fourth section, an overview of the DCT data communication system is done while the simulations results are discussed in section 5.

2 DCT and current applications

Discrete Cosine Transform (DCT) is developed by Ahmed, Natarajan and Rao [12] and has seen its first applications in images compression by Chen and Pratt [4]. DCT is present in several current standards of images and video coding [21] such as JPEG (Joint Photographic Experts Group), MPEG (Moving Pictures Experts Group), H.261 and its evolutions. DCT is mathematically formulated for bi-dimensional (2D) signals, \( Y(x,y) \) of size \( N \times M \), as drawn up in (1) and its inverse transform is expressed as in (2).

\[
S(u,v) = \frac{2}{MN} C(u)C(v) \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} Y(x,y) \cos \left( \frac{\pi}{N} u \left( x + \frac{1}{2} \right) \right) \cos \left( \frac{\pi}{M} v \left( y + \frac{1}{2} \right) \right) \tag{1}
\]

with

\[
C(0) = 2^{-1/2}, \quad C(w) = 1, \quad w = 1, 2, ..., N-1 \quad \text{and} \quad u = 0, ..., N-1 \quad \text{et} \quad v = 0, ..., M-1
\]

\[
Y(x,y) = \frac{2}{\sqrt{MN}} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} C(u)C(v)S(u,v) \cos \left( \frac{\pi}{N} u \left( x + \frac{1}{2} \right) \right) \cos \left( \frac{\pi}{M} v \left( y + \frac{1}{2} \right) \right) \tag{2}
\]

Recently, images and video coding applications evolved to meet their transmission issues. In [22], Xiang et al dealt with JPEG images transmission based on an UEP approach of the DCT components. They designed a SSI that relies on the partitioning of DCT frequencies into two levels of significance separating the DC components to the AC’s. They argue such a hierarchy of significance of the DCT source with two considerations.

The first is relative to the fact that, in the 2D DCT done in JPEG coding, the DC coefficient is a measure of the average value of the 64 image samples within a block of size 8×8 and contains a significant fraction of the total image energy.
Their second consideration relies on the higher error sensitivity of the DC coefficients due to the strong correlation introduced by the differential pulse code modulation (DPCM) coding. Accordingly when errors occur on the DC bits in one block, then they also induce decoding errors in subsequent blocks. On the other hand, decoding errors in the AC bits only affect local blocks.

Leaning on their DCT source partitioning approach, they allocate a lower turbo coding rate to the highly sensitive DC components and a higher coding rate to the least sensitive AC components. Their simulations results showed that their UEP schemes outperform the classical equal error protection.

Beyond their considerations, we think that additional assumptions may hold right when taking into account a coding based on human visual system.

3 The proposed approach of DCT SSI design

3.1 HVS-based Foundations

The human visual system (HVS) approach of image coding has brought significant rate-distortion enhancement in the JPEG standard [20]. It is nowadays acknowledged that the quality measures based on HVS approach are in better correlation with human observers’ score than classical measures such as PSNR (peak signal to noise ratio) or MSE (mean square error) [1,19].

Leaning, on the one hand, on Mannos and Sakrison HVS model [9] and, on the other hand, on Daly’s modulation transfer function [6] drawn up in (3), the authors of [20] proposed the weighting matrix $H$, given in (4), from which they derived the quantization table $Q_{HVS}$ of the expression (5) for the DCT 8×8 images coding. So if $x(u,v)$ is an input of the quantizer, then its corresponding output $x_q(u,v)$ follows the equation (6).

$$H(p) = \begin{cases} 2.2(0.192 + 0.114f(p))\exp\left(\left(0.114f(p)\right)^2\right) & \text{if } f(p) < f_{\text{max}}, \\ 1.0 & \text{otherwise,} \end{cases}$$ (3)
Designing DCT source significance information

\[ H = \begin{bmatrix}
1.0000 & 1.0000 & 1.0000 & 1.0000 & 0.9599 & 0.8746 & 0.7684 & 0.6571 \\
1.0000 & 1.0000 & 1.0000 & 1.0000 & 0.9283 & 0.8404 & 0.7371 & 0.6306 \\
1.0000 & 1.0000 & 0.9571 & 0.8898 & 0.8192 & 0.7371 & 0.6471 & 0.5558 \\
1.0000 & 1.0000 & 0.8898 & 0.7617 & 0.6669 & 0.5912 & 0.5196 & 0.4495 \\
0.9599 & 0.9283 & 0.8192 & 0.6669 & 0.5419 & 0.4564 & 0.3930 & 0.3393 \\
0.8746 & 0.8404 & 0.7371 & 0.5912 & 0.4564 & 0.3598 & 0.2948 & 0.2480 \\
0.7684 & 0.7371 & 0.6471 & 0.5196 & 0.3930 & 0.2948 & 0.2278 & 0.1828 \\
0.6571 & 0.6306 & 0.5558 & 0.4495 & 0.3393 & 0.2480 & 0.1828 & 0.1391 
\end{bmatrix} \]  

\[ Q_{HVS}(u,v) = \frac{q}{H(u,v)}, \text{ for } 0 \leq u, v \leq 8 \]  

\[ x_q(u,v) = \text{round}\left[ \frac{x(u,v)H(u,v)}{q} \right] \]  

In (6), one can easily notice that every DCT component is quantized according to its visual significance since all the DCT coefficients are not of equal visual significance. The proposed approach of DCT data SSI design, operating with the DCT sources partitioning into levels of significance, intends to get profit of this reading as a further consideration to those assumed in the Xiang’s approach.

3.2 The proposed DCT data SSI design

In [15], we reported a design of the DCT source significance (information) partitioning into two, three and four levels for the UEP. Further simulations, which are not involved hereby, have shown us that the partitionings into three and four levels are quite of equal performance in terms of robustness and efficiency. In this paper, we get profit of these previous works to formulate the source significance hierarchy structure according to the expression in (7).

The proposed approach and the one reported by Xiang have in common the running of UEP for images coding and transmission by partitioning the DCT frequencies into levels of significance to constitute the SSI. What could be different relies on the frontiers between the different levels of significance. Since in the Xiang’s approach, the number of the significance levels and their frontiers are fixed for ever, his resulting SSI can be referred to as a fixed SSI (FSSI). At the contrary the proposed SSI is adapted to several parameters such as the desired reconstruction quality (distortion), the error protection codes (rate compatible punctured convolutional (RCPC) codes [7]) and the channel perturbations intensity (SNR: signal to noise ratio). Therefore, the resulting frontiers of the different levels of significance have to vary; hence the proposed SSI could be called adaptive SSI (ASSI).

In section 2, we mentioned two major considerations which argue the Xiang’s approach of the DCT source partitioning. Approving the relevance of these
considerations, we go forward in thinking that additional considerations may lead to think the things differently. As matter of fact, while reading the components of the weighting matrix \( H \), we observe that some AC components possess the same perceptual significance weight as the DC component. That means that their contribution is not always so negligible as assumed in the Xiang’s approach. These AC components are also significant in rendering the reconstructed image visual quality. Consequently they deserve to be as well protected as the DC. Moreover the zigzag scanning of the AC components within a block sweeps them following a decreasing order of their energy. So, even if we reduce the DCT coefficients to the AC’s alone, the first AC in zigzag order are not rigorously of equal significance as the last ones. That’s an additional consideration that leads us to design the SSI into levels of significance, for a block of size \( 8 \times 8 \), as described by the expression (7). The DCT source significance hierarchy as suggested by Xiang could be described by the formulation of (8).

\[
\begin{align*}
S_n : & \quad \eta^0 = \{ DC \} \\
& \quad \eta^i = \{ n \text{ first AC in zigzag order, } 0 \leq n < 63 \} \\
& \quad \eta^2 = \text{the remainder}
\end{align*}
\]

\[
\begin{align*}
S_X : & \quad \eta^0_X = \{ DC \} \\
& \quad \eta^i_X = \{ \text{all the AC} \}
\end{align*}
\]

In the design of (7) we allow the two first levels of significance \( \eta^0 \) and \( \eta^i \) to be merged in a single one or to remain separated, provided that the resulting partitioning (two or three levels) is the optimal one. \( \eta^0 \) and \( \eta^i \) are merged when a same RCPC code is used to encode them. Flexibility is taken into account in such a design to bring the optimal partitioning according to the image with its specific contents. To achieve the SSI as formulated in (7) we propose an algorithm that the following subsection is about.

### 3.3 Algorithm for the DCT source significance hierarchy elaboration

The proposed algorithm is similar – although the context is different - to the simulated annealing algorithm (SAA) since it is a Monte Carlo algorithm [2]. Used for the index assignment optimization in vector quantization, SAA aims at optimizing the design of the vector quantization codebook. The proposed algorithm that we denote by SPA (source partitioning algorithm), aims at optimizing the DCT source partitioning into levels of significance for the UEP. SPA has the task of seeking the smallest value \( n^* \) of \( n \) that satisfies a desired global distortion \( D^* \) constraint under a predefined fixed channel state expressed by a signal to noise ratio (SNR).
Designing DCT source significance information

The total end-to-end distortion $D$ is evaluated using the expressions (9) and (10) of the weighted peak signal to noise ratio ($wPSNR$) [19].

\[
\begin{align*}
  wPSNR &= 10 \log_{10} \left( \frac{255^2}{wMSE} \right) \\
  wMSE &= \frac{1}{N^2} \sum_{u=1}^{N} \sum_{v=1}^{N} H^2(u,v)(x_o(u,v) - \hat{x}_o(u,v))^2
\end{align*}
\]  

(9) (10)

where $wMSE$ is the weighted mean square error, $N \times N$ is the size of the image or video frame to be DCT coded, and $x_o$, $\hat{x}_o$ respectively the original image and its reconstructed version. Therefore, for any state $S_n$ (of the space $P$) of the DCT source partitioning that is coded and transmitted using the UEP with RCPC codes combination $\{R(\eta^0), R(\eta^1), R(\eta^2)\}$, a corresponding distortion $D(S_n)$ can be calculated using (9). The algorithm terminates when $D(S_n) \geq D^*$ or the partitioning space $P$ of $S_n$ is empty. Figure 1 describes the different steps to be processed in the running of this algorithm. $T_f$ characterizes the working SNR value of the additive white gaussian noise transmission channel involved in the simulations.
4 Overview of the communication system

The proposed DCT images coding and transmission system is like the block diagram depicted in figure 2. The SPA module provides the RCPC coder with the source significance information. The RCPC coder achieves the unequal error protection of the information bitstream following that they whether belong to a significance level $\eta^j$. The information bitstream of the DC components that are always of level $\eta^0$ in the ASSI schemes or $\eta_x^0$ in FSSI schemes are associated with the RCPC code of rate $R(\eta^0)$ of the triplets $(R(\eta^0), R(\eta^1), R(\eta^2))$ or the
RCPC code $R(\eta^0_X)$ of the couple $(R(\eta^0_X),R(\eta^1_X))$. The RCPC codes triplets $(R(\eta^0_0),R(\eta^1_0),R(\eta^2_0))$ obey to the relation $R(\eta^0_0) \leq R(\eta^1_0) \times R(\eta^2_0)$ in conformity with the possibility for the two first levels of significance $\eta^0$ and $\eta^1$ to be merged in a single one or to remain separated. In this way, we enlarge the scope of the expected optimal SSI design of (7). In the other hand the couple $(R(\eta^0_X),R(\eta^1_X))$ is obviously constrained to the inequality $R(\eta^0_X) \times R(\eta^1_X)$.

In the simulation of the Xiang’s approach, all the zigzag scanned AC components are gathered in a single frame and the information bitstream resulting from source coding is coded using the RCPC code of rate $R(\eta^1_X)$.

In the proposed method, whenever the SPA provides us with the frontier between $\eta^1$ and $\eta^2$, two AC frames are constituted and separately processed in the RLE block. However, for optimality purpose, a single Huffman codebook is constructed for all the AC. The corresponding information bitstream are unequally protected against channel errors in accordance with their respective significance levels ($\eta^1$ or $\eta^2$) by using RCPC codes with respective rates $R(\eta^1_0)$ or $R(\eta^1_X)$.

In addition to the UEP methods, the classical EEP method using the convolutional code of rate 1/2 is simulated. In the three cases, the channel coder receives bits packets of limited size in order to reduce the errors propagation within a frame. The simulated transmission channel includes a QPSK (quadrature phase shift keying) modem and an additive white gaussian noise (AWGN).

5 Simulation results

Simulations are run for five well known images of size 512×512 pixels that are source coded at very close rates. The end-to-end distortion $D^*$ parameters used in the simulations of the SPA correspond to $wPSNR$ values in the range $[38, 45]$. For the visual appearance examination, figure 3 and figure 4 provide us with samples of images. The simulated RCPC codes derived from a mother code 1/2, a puncturing period $P = 4$, a memory $M = 6$ and a generator $G = [133 \ 171]$, so that the codes rates that are given by $P/P + l$, $l = 1,...,P$, correspond to the respective rates 4/5, 4/6, 4/7 and 4/8.
In recent UEP designs a total rate-distortion function, whenever available, is modelled to elaborate a cost function for the optimal redundancy allocation. However, when such a model misses one can opt for some empirical strategies as done in [8, 13]. As far as we are concerned hereby, we think that in the proposed UEP a total rate distortion function is complex to be designed and may necessitate a thorough study. That why an exhaustive search, among the RCPC codes combinations of the family with mother code 1/2, allowed us to find the optimal RCPC codes assignation to the different significance levels. Therefore, we obtain the optimal redundancy allocation grid that is reported in table 1. We understand by optimal redundancy allocation the use of a RCPC codes combination - such as $\left( R(\eta^0), R(\eta^1), R(\eta^2) \right)$ for ASSI or $\left( R(\eta^0_1), R(\eta^1_2) \right)$ for FSSI - that enables the coding of a given image with fewer bits resources and keeping the desired $D^*$ reconstruction quality of the image. It results from table 1 that the optimal DCT source partitioning is a 2-level significance hierarchy obeying to the formulation $S_n^*$ as drawn up in (11). This result reinforces our preoccupation expressed in section 3.2 about the fact that there are some AC which deserve to be as well protected as the DC. In table 1, $n_l$ and $n_u$ mean respectively the average (across the five images that are simulated) $n^*$ value of $n$ in $S_n^*$ for respectively the lower and upper bounds of the corresponding SNR interval.

The performances associated with the optimal redundancy allocation are provided through figure 5 to figure 8. The total bit rate (TBR), expressed in bits per pixel (bpp), is representative of the total transmission cost in terms of bits resources. These figures highlight comparative total transmission cost variation as a function of the channel status given in terms of signal to noise ratio. They also
point out the clear outperformances of the proposed ASSI. As matter of fact, the proposed ASSI is revealed more efficient than the other methods since it produces the lowest TBR. Moreover since a robust transmission is characterized by lower working SNR while keeping $D^*$, we conclude that ASSI is more robust than the FSSI and as robust as the maximum protection strategy solution carried by the classical EEP.

Table 1
Grid of optimal redundancy allocation $\left\{ R(\eta^{\text{in}}), R(\eta^{\text{in}}) \right\}$ and associated channel SNR interval.

<table>
<thead>
<tr>
<th>SNR interval</th>
<th>FSSI</th>
<th>ASSI</th>
<th>$[n_1, n_u]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 8.2</td>
<td>N/A</td>
<td>N/A</td>
<td>(4/8, 4/7)</td>
</tr>
<tr>
<td>[8.2, 8.5]</td>
<td>N/A</td>
<td>(4/7, 4/6)</td>
<td>[27, 14]</td>
</tr>
<tr>
<td>[8.5, 8.8]</td>
<td>N/A</td>
<td>(4/7, 4/6)</td>
<td>[27, 14]</td>
</tr>
<tr>
<td>[8.8, 9.4]</td>
<td>(4/7, 4/6)</td>
<td>(4/6, 4/5)</td>
<td>[28, 16]</td>
</tr>
<tr>
<td>[9.4, 10.7]</td>
<td>(4/6, 4/5)</td>
<td>(4/6, 4/5)</td>
<td>[10, 10]</td>
</tr>
<tr>
<td>[10.7, 13]</td>
<td>uncoded</td>
<td>uncoded</td>
<td>uncoded</td>
</tr>
<tr>
<td>≥ 13</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$$S_n^* = \left\{ \eta^{\text{in}} = \{DC\} \cup \{the \ n \ first \ AC \ in \ zigzag \ order, 0 \leq n < 63\} \right\}$$

$$\left\{ \eta^{\text{in}} = \{the \ remainder\} \right\}$$

(11)

6 Conclusion

In summary, this paper dealt with the improvement of an existing UEP [22] design for DCT-based images or video coding and transmission. Since the source significance information of the existing design relies on the partitioning of the DCT frequencies into levels of significance we innovated with an algorithm for an adaptive partitioning of the DCT sources so that the transmission channel parameters and the reconstructed image quality are taken into account to yield the best partitioning. Simulations are run for several images and obtained results pointed out the outperformances of the proposed UEP compared to the existing UEP and the classical EEP.

Moreover, we think that the obtained robustness performances can be improved by replacing the mother code 1/2 of the RCPC codes with a more powerful (e.g. mother code 1/3). However the trade-off is in the increasing of the transmission cost. In mobile transmission environments, we expect the same qualitative performances with the introduction of hybrid ARQ/FEC [14] mechanism to face the fading effects. In systems like HSDPA (high speed downlink packet access) the availability of a channel state indicator makes the
integration of the proposed SPA attractive for DCT based images and video communication.

Acknowledgements

The authors thank the “Agence Universitaire de la Francophonie, Bureau Afrique de l’Ouest” for its financial support.

Figure 3. Lena image corresponding to a distortion $wPSNR(D^*) = 38.3$ dB
Figure 4. Masuda image corresponding to a distortion $w_{PSNR}(D^*) = 44.7 \text{ dB}$
Figure 5. Comparative efficiency (TBR) and robustness (SNR) performances of the three methods to reconstruct Lena (image) at a desired distortion given by $wPSNR(D^*) = 38.3 \text{ dB}$.

Figure 6. Comparative efficiency (TBR) and robustness (SNR) performances of the three methods to reconstruct Masuda (image) at the distortion given by $wPSNR(D^*) = 44.7 \text{ dB}$.
Figure 7. Comparative efficiency (TBR) and robustness (SNR) performances of the three methods to reconstruct House (image) at a desired distortion given by $w\text{PSNR}(D^*) = 43.2 \text{ dB}$.

Figure 8. Comparative efficiency (TBR) and robustness (SNR) performances of the three methods to reconstruct Cameraman (image) at the desired distortion given by $D^* = 40.9 \text{ dB}$.
References


Received: August 25, 2008