

Proceedings

Analog Video Encoding and Quality Evaluation [†]

Jose Balsa ^{*}, Óscar Fresneda, José A. García-Naya, Tomás Domínguez-Bolaño and Luis Castedo

CITIC Research Center, Universidade da Coruña (University of A Coruña), 15071 A Coruña, Spain

^{*} Correspondence: j.balsa@udc.es

[†] Presented at the 2nd XoveTIC Conference, A Coruña, Spain, 5–6 September 2019.

Published: 23 July 2019

Abstract: The most widespread analog video encoding systems in the literature are based on the use of the 2D and 3D DCT. These systems use both transformations indistinctly without assessing their suitability. In this paper, we present procedures to compress video using 2D and 3D-DCT and we evaluate the video quality for different compression levels.

Keywords: analog video encoding; video quality evaluation

1. Introduction

Nowadays, the multimedia use of computers and mobile devices is increasing because of its applications. Within this, the transmission of information quickly and without delay is essential for applications and users. In this environment, video transmission plays a predominant role with a great challenge due to the relationship between transmitting a large amount of data and compressing it by introducing a high computational load.

In recent years, if we talk about video transmission and compression, we are talking about digital processes. These systems are capable of delivering high performance in the vast majority of possible scenarios. Although it is true that these digital systems present some well-known problems. On one hand, video compression techniques look for spatial and temporal correlations and require a high computational load. On the other hand, if data cannot be recovered without errors, retransmissions are needed, which degrades the delay of the communication link.

An alternative approach to digital systems is to use analog transmissions, which provide low delay and low complexity. Most of the existing works regarding analog video encoding and transmission propose hybrid analog-digital schemes, where the analog part consists on the use of the discrete cosine transform (DCT). The idea of using the DCT is that the components at higher frequencies correspond to the most important visual information; hence, some of the coefficients can be discarded without affecting the image quality. Some of the proposed systems use the 2D-DCT [1], whereas others consider the 3D-DCT [2,3]. However, in most of them, the digital part is a key component. In order to compare the image quality after the transmission there exist different metrics: peak signal-to-noise ratio (PSNR) is considered in [1–3], whereas structural similarity (SSIM) and signal-to-distortion ratio (SDR) is employed in [3].

In this context, we compare both transformations, 2D and 3D DCT, in videos with movement and static scenarios. We have to establish some metrics to evaluate and compare the system evaluation parameters such as the compression ratio, related to the transmitted frequencies, and image quality, measured in terms of PSNR and SSIM.

2. System Description

We propose two analog schemes for video encoding: one using the 2D-DCT and another one employing the 3D-DCT. The 2D-DCT system encodes each individual video frame using the analog scheme proposed in [4] for still images, although in this case the correlation between frames is not

considered. In this system, each frame is divided into 8×8 blocks and the DCT transformation is applied to each block. The resulting DCT coefficients are stacked onto a vector following a zigzag pattern. Thus, the resulting vector will be sorted from low to high frequencies. The symbols corresponding to the higher frequencies are discarded, thus compressing the image and reducing its visual quality.

Regarding the 3D-DCT system, the entire video is firstly divided into sequences of 8 frames each. Next, each sequence is divided into blocks of 8×8 pixels, thus the whole video is split in cubes with dimension $8 \times 8 \times 8$. We now define the concept of *symbol* as the pixel intensity (luminance in our case) expressed as an integer number ranging between 0 and 255. A weight is then assigned to each symbol according to its low or high frequency using the 3D pattern defined in [5] to rearrange the symbols into a sequence from the lower to the higher frequencies. As in the 2D-DCT case, the symbols corresponding to the higher frequencies are discarded, hence compressing the image and reducing its visual quality.

Next, a comparison with the original video sequence is carried out to determine the video quality related to the compression factor. The metric consists in comparing each original frame with the compressed one and averaging out the result. More specifically, in this paper both the SSIM and PSNR are considered since they are the most used metrics to perform this type of comparisons [1–3].

3. Results

The results presented here are the product of averaging out three different video sequences. We have also taken into account different scenarios to get a fair comparison. We tested resolutions ranging from SD to 1080p, as well as with different static and motion sequences.

Figure 1 shows the results obtained from the system simulations. The results show that the quality of the video is higher for the 3D-DCT than for the 2D-DCT and for both metrics: PSNR and SSIM. Note that in Figure 1a, when the compression factor becomes 1, the PSNR approaches infinite.

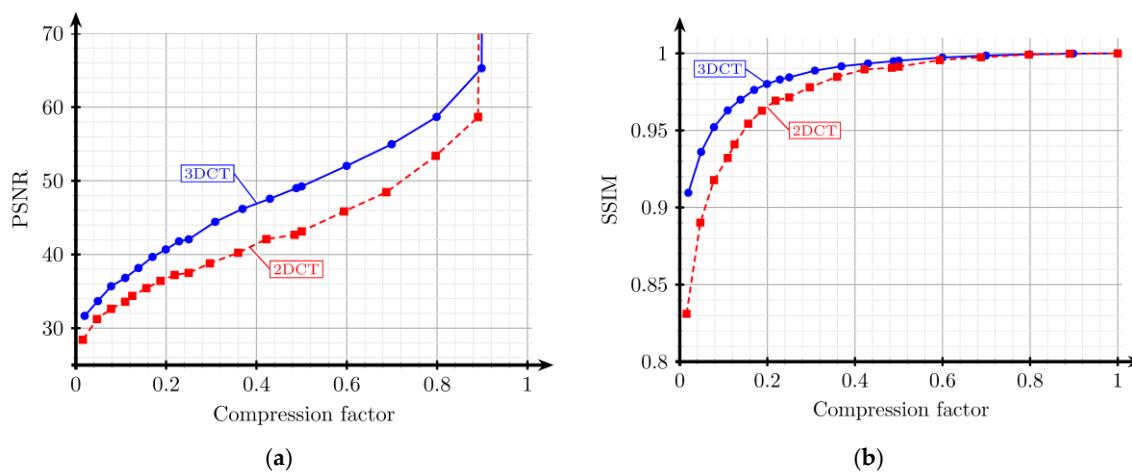


Figure 1. Quality comparison between 2D- and 3D-DCT versus the compression factor for the two considered metrics: (a) PSNR in decibels, and (b) SSIM.

4. Conclusions

The results show that with the same video quality it is possible to achieve a higher compression level when the 3D-DCT is employed instead of the 2D-DCT. This is because the 3D-DCT considers both spatial and temporal correlation at the same time, hence reducing the amount of redundant information in the three dimensions.

In view of the results, it can be safely concluded that the 3D-DCT can be considered as an improvement with respect to the 2D-DCT for video applications.

Author Contributions: Conceptualization, J.B., T.B., O.F. and J.G.; software, J.B.; investigation, J.B.; writing—original draft preparation, J.B.; writing—review and editing, T.B., O.F. and J.G.; funding acquisition, L.C.

Acknowledgments: This work has been funded by the Xunta de Galicia (ED431C 2016-045, ED431G/01), the Agencia Estatal de Investigación of Spain (TEC2016-75067-C4-1-R) and ERDF funds of the EU (AEI/FEDER, UE).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Fan, X.; Wu, F.; Zhao, D.; Au, O. C.; Gao, W. Distributed Soft Video Broadcast (DCAST) with Explicit Motion. In Proceedings of the 2012 Data Compression Conference, Snowbird, UT, USA, 10–12 April 2012.
2. Jakubczak, S.; Katabi, D. SoftCast: Clean-slate scalable wireless video. In Proceedings of the 2010 48th Annual Allerton Conference on Communication, Control, and Computing (Allerton), Allerton, IL, USA, 29 September–1 October 2010.
3. Liang, F.; Luo, C.; Xiong, R.; Zeng, W.; Wu, F. Hybrid Digital–Analog Video Delivery With Shannon–Kotel’nikov Mapping. *IEEE Trans. Multimed.* **2018**, *20*, 2138–2152.
4. Balsa, J.; Domínguez-Bolaño, T.; Fresnedo, Ó; García-Naya, J.A.; Castedo, L. Transmission of Still Images Using Low-Complexity Analog Joint Source-Channel Coding. *Sensors* **2019**, *19*, 2932.
5. Chan, R. K. W.; Lee, M. C. 3D-DCT quantization as a compression technique for video sequences, In Proceedings of the International Conference on Virtual Systems and MultiMedia VSMM ’97 (Cat. No. 97TB100182), Geneva, Switzerland, 10–12 September 1997.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).