Source coding of holographic data: challenges, algorithms and standardization efforts

Schelkens, Peter; Blinder, David; Birnbaum, Tobias; Symeonidou, Athanasia; Ahar, Ayyoub; Schretter, Colas

Published in:
Proceedings of Speckle 2018: VII International Conference on Speckle Metrology

DOI:
10.1117/12.2322721

Publication date:
2018

Citation for published version (APA):
Schelkens, P., Blinder, D., Birnbaum, T., Symeonidou, A., Ahar, A., & Schretter, C. (2018). Source coding of holographic data: challenges, algorithms and standardization efforts. In M. Józwik, L. R. Jaroszewicz, & M. Kujawińska (Eds.), *Proceedings of Speckle 2018: VII International Conference on Speckle Metrology: VII International Conference on Speckle Metrology* (Vol. 10834). [108340Y] (Proceedings of SPIE - The International Society for Optical Engineering; Vol. 10834). SPIE. https://doi.org/10.1117/12.2322721

Source coding of holographic data: challenges, algorithms and standardization efforts

Peter Schelkens, David Blinder, Tobias Birnbaum, Athanasia Symeonidou, Ayyoub Ahar, et al.

Peter Schelkens, David Blinder, Tobias Birnbaum, Athanasia Symeonidou, Ayyoub Ahar, Colas Schretter, "Source coding of holographic data: challenges, algorithms and standardization efforts," Proc. SPIE 10834, Speckle 2018: VII International Conference on Speckle Metrology, 108340Y (7 September 2018); doi: 10.1117/12.2322721

Event: SPECKLE 2018: VII International Conference on Speckle Metrology, 2018, Janów Podlaski, Poland

Source coding of holographic data: challenges, algorithms and standardization efforts

Peter Schelkens[a,b], David Blinder[a,b], Tobias Birnbaum[a,b], Athanasia Symeonidou[a,b],
Ayoub Ahar[a,b], and Colas Schretter[a,b]

[a]Vrije Universiteit Brussel, Dept. of Electronics and Informatics, Pleinlaan 2, B-1050 Brussels, Belgium
[b]imec, Kapeldreef 75, B-3001 Leuven, Belgium

ABSTRACT

Holographic imaging modalities are gaining increasing interest in various application domains ranging from microscopy to high-end autostereoscopic displays. While much effort has been spent on the development of the optics, photonics and micro/nano-electronics that enable the design of holographic capturing and visualization devices, relatively few research effort has been targeted towards the underlying signal processing. One significant challenge relates to the fact that the data volumes needed in support of this kind of holographic applications is rapidly increasing: for visualization devices, and in particular holographic displays, unprecedented resolutions are desired resulting in huge bandwidth requirements on both the communication channels and internal computing & data channels. An additional challenge relates to the fact that we are handling an interference-based modality being complex amplitude based in nature. Both challenges lead to the fact that for example classic data representations and coding solutions fail to handle holographic data in an effective way. This paper attempts to provide some insights that enable to alleviate or a least reduce these bottlenecks and sketch an avenue for the development of efficient source coding solutions. Moreover, it will also outline the efforts the JPEG committee is undertaking in the context of the JPEG Pleno standardization programme to roll out a path for data interoperability of holographic solutions.

Keywords: Holography, compression, coding, standardization, JPEG Pleno

1. INTRODUCTION

Holographic imaging modalities are gaining increasing interest in various application domains ranging from microscopy to high-end autostereoscopic displays. In contradiction to classical stereoscopic and autostereoscopic light field displays, holographic displays bring the inherent advantage that they are not subject to the vergence-accommodation conflict.[1] Moreover, a general property of holographic systems is as well that they deliver a very high depth resolution due to the fact that the phase of the signal is captured, resulting in nanometer precision, which is particularly important for microscopy and non-destructive testing applications. While much effort has been spent on the development of the optics, photonics and micro/nano-electronics that enable the design of holographic capturing and visualization devices, unproportionally few research effort has been targeted towards the underlying signal processing.

Nonetheless, new signal processing insights are required at many stages of the holographic processing pipeline: advanced holographic tomographic capturing techniques, sparse reconstruction methods, computer-generated holography algorithms to produce holograms from other input sources such as point clouds and light fields, advanced rendering techniques to provide suitable holographic data to display, quality assessment metrics and methodologies, speckle denoising algorithms... One significant challenge relates to the fact that the data volumes needed in support of this kind of holographic applications is rapidly increasing. For visualization devices, and in particular holographic displays, unprecedented resolutions are desired resulting in huge bandwidth requirements on both the communication channels and internal computing & data channels. An additional challenge relates to the fact that we are handling an interference-based modality being complex amplitude based in nature.

Further author information: (Send correspondence to Peter Schelkens)
E-mail: peter.schelkens@vub.be, Telephone: +32 2 629 16 81

Speckle 2018: VII International Conference on Speckle Metrology, edited by Malgorzata Kujawińska,
Leszek R. Jaroszewicz, Proc. of SPIE Vol. 10834, 108340Y · © 2018 SPIE
CCC code: 0277-786X/18/$18 · doi: 10.1117/12.2322721

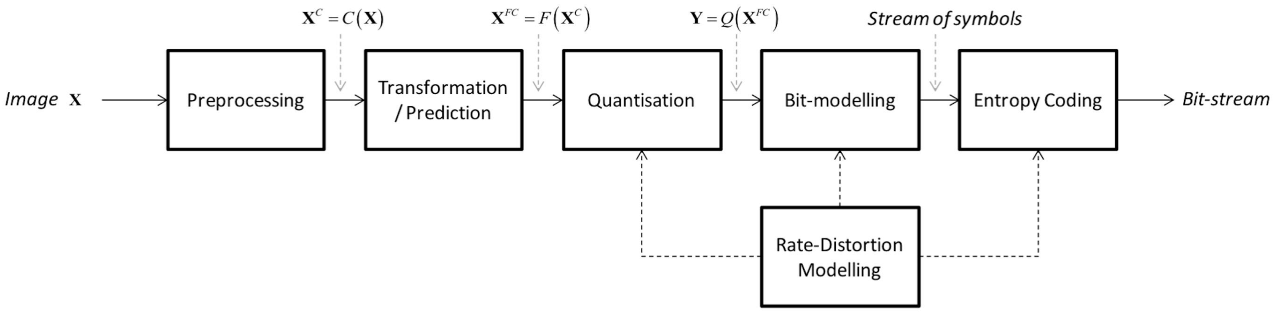

Figure 1. Generic architecture of an image coding system

Both challenges lead to the fact that for example classic data representation and coding solutions fail to handle holographic data in an effective way.

This work attempts to provide some insights that enable to alleviate or a least reduce these bottlenecks and to sketch an avenue for the development of efficient source coding solutions. One of the larger international initiatives that is focused on this particular problem, and that attempts to provide interoperability mechanisms to exchange holographic data between the different types of devices that emerges is the JPEG Pleno standardization programme. The JPEG committee[2] is undertaking in this context efforts to roll out a plan for data interoperability of holographic solutions and other plenoptic modalities.

The paper is structured as follows. In Section 2, the architecture for encoding image and video data is presented at a high abstraction level. This will be used as basis for detailing in the subsequent Section 3 the challenges faced when handling holographic data in a coding context. In addition, potential avenues are being outlined. In Section 4, an overview is provided of the JPEG Pleno standardization effort that among others also attempts to provide a framework for the encoding and exchange of holographic image data. Finally, in Section 5, conclusions are drawn for the domain of holographic image and video coding.

2. COMPRESSION BASICS

To understand well the challenges faced when compressing holograms, it is useful that we first summarize the core principles on which typical image and video coding systems are based. Actually, almost any coding system that is processing audio or visual information is following this scheme. The basic functionality of a compression system is to attempt to remove all redundancy – read correlations – from the signal that is subject to compression. This goal is achieved in multiple steps, which are discussed in the following paragraphs. Fig. 1 conceptually illustrates a complete coding system.

When captured or generated, image data $\mathbf{X}$ is often represented in Red-Green-Blue (RGB)-format, requiring B bit per colour component for representation. Due to the nature of the capturing – partially overlapping spectral bandpass filters – this data needs to be further decorrelated and hence for coding, it is converted to a luminance-chrominance (YCbCr) representation $\mathbf{X}^C = C(\mathbf{X})$. This representation has also the advantage that one can exploit the higher sensitivity of the Human Visual System (HVS) for luminance than for chrominance information. Roughly, one can state that the contrast sensitivity of the HVS for luminance information spans a two times as large spectral envelope as for chrominance data. A common practice is consequently to subsample the chrominance space with a factor two and as such achieve an additional data reduction. It is important to understand that the latter procedure is inherently lossy in nature. These steps are typically addressed as **Preprocessing** steps in a coding system.

At this stage the data is ready for launching the core and most demanding processes that are part of a coding system. In the **Transformation/Prediction** component, the codec will attempt to remove the remaining spatial and temporal redundancy from the signal. Ideally, the transform or prediction coefficients that are produced by this process $\mathbf{X}^{FC} = F(\mathbf{X}^C)$ should depict a sharply-peaked Laplacian point-density function (pdf). If the transform or prediction model is perfect – i.e. a perfect energy compaction by the transform into one coefficient or a perfect prediction resulting in a zero prediction error – maximum compression or an infinite compression factor can be achieved. Hence, this illustrates the importance of having a suitable signal model on which the transform and prediction tools are based to facilitate successful compression. In a practical case, we are of course limited to more generic signal models to be able to support a sufficiently wide range of use cases for off-the-shelf codecs.

For *spatial redundancy* reduction of classical imagery this translates into the deployment of image transforms that are based on for example cosine, wavelet, directional wavelet bases, each having their particular advantages and disadvantages which are beyond the scope of this paper. In addition, spatial predictors are being used that attempt to exploit local data redundancy by utilizing spatial, directional predictors that predict the current pixel values based on the values of previously encoded, neighbouring pixels. Previously encoded pixels, because the *causality principle* needs to be respected. In more advanced codecs, like H.265/HEVC, directional spatial prediction and spatial transforms are combined.

To exploit the *temporal redundancy*, most popular codecs are using motion estimation and compensation mechanisms. Based on a reference frame and in a block-based approach, image blocks in the frame to be encoded – called macroblocks – are predicted from the best matching block in the reference frame. The motion estimation process identifies the best matching blocks by deploying in the meanwhile very advanced motion models. The prediction error – i.e. the difference between the motion compensated macroblock and the original macroblock in the current frame - is subsequently further processed by a spatial transform to further compact the energy in the prediction error frame. An important remark to make here is that for light field data, codecs will exploit in a similar fashion correlations between the subaperture views. More advanced techniques deploying for example spatiotemporal transforms have been proposed as well, but none of them are currently beating the hybrid codec architectures outlined above. Please note that the motion information has to be signalled as well and that part of encoded bit budget will have to be spent on this information.

Up to this stage all operations are basically lossless, besides some potential finite precision computational issues – which can be solved as well – and potential chrominance subsampling. Hence, the obtained transform/prediction error coefficients and motion vector information allow for a near-lossless reconstruction of the data. To obtain higher compression, the accuracy with which this information is to be encoded has to be reduced, the process enabling this is called **Quantization**; $\mathbf{Y} = Q(\mathbf{X}^{FC})$.

Quantization will result in a reduced number of bits needed to represent the information and potentially also a particular ordering of these bits to support for example quality scalability. The latter functionality will take care that the coefficients are gradually refined when more bits are being received by the decoder. This process - called **Bit-modelling** – is particularly enabled by the quantization process.

The stream of symbols that is generated after bit-modelling needs to be converted subsequently into variable length codes, whereas short codewords are assigned with symbols that have a high probability of occurrence and long codewords are assigned to symbols with low probability. Optimal codewords will assure that the resulting bitstream will approach the entropy of the signal, i.e. the theoretical minimum number of bits required to losslessly encode the source (i.e. set of input symbols). Various techniques are utilized, but the most commonly used techniques in compression standards for **Entropy Coding** are Huffman and arithmetic encoding.

Finally, one module was not discussed yet; **Rate-Distortion Optimization**. This component will take care that given the available bit budget, i.e. bitrate, which can be expressed in bit-per-second or bit-per-pixel, a maximal reconstruction quality or minimum distortion can be achieved. In this optimization process, additional characteristics of the human visual system, content properties, display characteristics etc. can be account for. The whole quest towards better codecs is basically obtaining better rate-distortion performance. This component steers the Transformation, Prediction, Quantization, Bit-modelling and Entropy coding units of the system.

The above – though high-level – understanding will help us to provide more insight what is now particularly challenging to encoding holographic data.

3. CODING OF HOLOGRAPHIC DATA

3.1 Challenges

Holographic data comes with many challenges that basically destabilize the foundations of the efficient and strong architectures we have been constructing for classical image and video coding.

A fundamental characteristic trait is that we are not handling classical integer intensities anymore, but **complex-valued hologram representations**. Hence, real-imaginary or amplitude-phase representations need to be deployed, raising immediately the question of how to handle this data. Encoding the real and imaginary components of the holographic signal as independent components during encoding? And how to handle the phase information in the latter case, knowing that inevitably phase wrapping issues are going to pop-up when transforming, predicting and quantizing pixel values and transform/prediction coefficients?

Secondly, high quality holograms in terms of visual quality and large angular field-of-view (AFOV) support, requires up to 5000-10000 lines per mm, or holograms with a pixel pitch that is at least half of the wavelength of lowest spectral frequency – blue light – that is supported. Designing large displays results in images that are **1 Terapixel and beyond for tabletop displays**. Hence, codecs have to be extremely fast with minimal computational complexity and memory footprint per pixel that needs to be processed.

Next, holography is based on the assumption that light propagates according to a **wave-based light propagation model** and not ray-based as is the case for classical and light field imagery. This has many consequences. One important effect is that interference patterns are being recorded and most popular transforms and prediction mechanisms are poorly handling this behaviour.

Even more disastrous, **motion estimation and compensation fail** because of the wave propagation model. If the diffraction angle is sufficiently large and/or the hologram recording plane is sufficiently distant from the scene, every pixel in the hologram plane will contain information about every point of the 3D scene; yes, also motion information. Classical block matching algorithms will not do anymore, as traditional motion compensation will not do either.

3.2 State-of-the-art coding solutions

Over the last decade, several solutions have been proposed that attempt to provide a solution for encoding holographic data.[3] They can be classified in three categories.

The first and largest category of solutions operate on the hologram plane, basically **encode the raw holographic data**, and hence, attempt to handle the unavoidable fringe patterns. Consequently, they attempt to match better the particular signal models of holograms, which are depending as well on the type of holograms – e.g. in-line, off-line or computer generated holograms with varying properties. Typically, this results in tuning the transform component of existing codecs, like JPEG 2000 or H.265/HEVC, by replacing it with transforms that can better handle the directionality or more uniform spectral envelop of holographic signals. Typical examples are: Fresnelets,[4] bandelets,[5] directional-adaptive wavelets,[6] arbitrary packet decompositions,[6] vector quantization lifting schemes,[7] wave atoms,[8] Gabor atoms,[9] modulo wavelet transform to handle phase data,[10] and mode-dependent directional transform-based HEVC.[11] These solutions all deliver often a significant improvement with respect to their non-adapted counterparts in terms of rate-distortion performance; however, they are not able to interpret the 3D scene that was captured by the hologram and hence their performance is ceiled to that extend.

The second category does not suffer from this shortcoming since it opts to **encode the data in object plane**, either by using encoding the source data issued later on to generate the hologram by using computer-generated holography (CGH) techniques, or by performing a numerical back-propagation of the hologram to the object plane. The first case is interesting since for CGH the source data can be encoded more efficiently and e.g. holograms can be generated on the fly from for example compressed textures and depth maps.[12] The latter case has the advantage that for a roughly flat scene, traditional codecs will still behave well due to limited presence of fringes. In case of deep scenes, they fail though. This approach is also still subject of debate in the coding community. Notwithstanding the fact that the encoding works well for "flat" objects, like holographic microscopy, back-propagating for encoding and forward propagation for decoding and display represents a significant computational cost. Hence, it is really use case dependent whether it is an acceptable strategy to encode the object plane or the CGH source data.

The last category, attempts to find an **hybrid solution** by combining the best of two worlds. At the transform level, it was recently proposed to deploy non-linear canonical transforms.[13] Hereby, the back-propagation of the hologram is locally steered based on a piece-linear approximation of the depth map of the 2D scene such that the obtained representation is fully in focus. Thereafter a classical coding strategy can be deployed.

In terms of video coding or coding of larger holograms, so far, no significant efforts have been reported to develop more efficient handling of holographic data.

3.3 Research target

First of all, it should be clear that sticking with nowadays widespread coding technologies will not do the job. We need to rethink thoroughly the current codec architectures to enable them to tackle the various outlined challenges. Although important steps have been taken in terms of understanding holographic signals and the associated requirements on the codec technology, nowadays solutions still insufficiently address the challenges we are facing. To handle deep scenes from a compression efficiency point-of-view, pursuing an object-based or a depth-steerable transform appears to be a wise challenge to pursue. This will also facilitate more efficient motion estimation and compensation. Motion estimation can in most case be perfectly done in the CGH source domain and hence does not require any new technology. We need to migrate too towards an heterogeneous solution whereas compression efficiency and computation complexity are traded between coder and decoder based on the constraints imposed by the use case.

In conjunction and accounting for the immense amount of data that has to be processed, data delivery to high-end displays needs to be considerably revisited. Similar as it is the case for light field displays, displaying and consequently also decoding information of which only a small fraction is instantly consumed does not make much sense. Therefor, decoders should be able to extract only that piece of data that is required to support the angular field of view required at a certain time interval for each individual viewer. Applying such strategies will already alleviate much pressure on the computational capacity, introducing though strict requirements on acceptable latency on the distribution channels.

4. JPEG PLENO

We need to realize that we can expect a plethora of realizations of holographic displays, each supporting for example potentially different angular fields-of-view, spatial resolutions, pixel pitches. Hence, a much larger diversity in terms of potential display parameterizations as for regular 2D displays emerges, likewise it is the case for light field displays.[14] In terms of addressing interoperability, this poses a significant challenge for standards attempting to address the compression of holograms.

Recently, the JPEG committee (ISO/IEC JTC1/SC29/WG1) initiated the JPEG Pleno standardization framework[15] that will facilitate the capturing, representation and exchange of not solely of holograms, but also point cloud and light field imaging modalities.[16] In its philosophy, these imaging modalities are understood to be light representations all inspired by the plenoptic function, regardless of which modality was used to capture or create parts of the entire content. This mindset recognizes that conversions of - or between - different modalities are possible and often useful. The standard aims at (1) defining tools for improved compression while providing advanced functionalities at system level and (2) supporting data and metadata manipulation, editing, random access and interaction, protection of privacy and ownership rights as well as other security mechanisms.

In the context of the standardization process, JPEG is also setting up a public JPEG Pleno database[15] with plenoptic test data to support these efforts. This database has to be sufficiently heterogeneous in terms of types of holographic content and covering different application domains.

The JPEG Pleno standard will exist out of different parts, where Part 1 covers the overall architectural description and Part 2 is currently focused on light field coding. Additional parts of the standard, currently in an exploratory phase, will address point cloud and holographic modalities and system layer aspects.

5. CONCLUSIONS

Coding of holographic signals remains a huge challenge. In recent years, important steps have been set to better understand these signals from a compression perspective. Several promising strategies to handle this data have been proposed in recent years, out of which scene segmentation with support of associated flexible transform are important pathways. Moreover, we have to think more carefully about the architectural aspects of a codec in terms of supporting functional aspects such as random access and distributed processing capacities.

ACKNOWLEDGMENTS

The research leading to these results received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007- 2013)/ERC Grant Agreement n. 617779 (INTERFERE).

REFERENCES

[1] Kramida, G., "Resolving the vergence-accommodation conflict in head-mounted displays," *IEEE Transactions on Visualization and Computer Graphics* **22**, 1912–1931 (July 2016).

[2] "ISO/IEC JTC1/SC29/WG (JPEG) Website." https://www.jpeg.org/. (Accessed: 25 June 2018).

[3] Dufaux, F., Xing, Y., Pesquet-Popescu, B., and Schelkens, P., "Compression of digital holographic data: an overview," *SPIE Photonics Europe – Optics, Photonics and Digital Technologies for Imaging Applications* **9599** (2015).

[4] M. Liebling, T. Blu, M. U., "Fresnelets: New multiresolution wavelet bases for digital holography," *IEEE Transactions on Image Processing* **12**(1), 29–43 (2003).

[5] Bang, L. T., Ali, Z., Quang, P. D., Park, J.-H., and Kim, N., "Compression of digital hologram for three-dimensional object using wavelet-bandelets transform," *Opt. Express* **19**, 8019–8031 (Apr 2011).

[6] Blinder, D., Bruylants, T., Ottevaere, H., Munteanu, A., and Schelkens, P., "JPEG 2000-based compression of fringe patterns for digital holographic microscopy," *Optical Engineering* **53**(12), 123102 (2014).

[7] Xing, Y., Kaaniche, M., Pesquet-Popescu, B., and Dufaux, F., "Adaptive nonseparable vector lifting scheme for digital holographic data compression," *Appl. Opt.* **54**, A98–A109 (Jan 2015).

[8] T. Birnbaum, D. Blinder, C. S. and Schelkens, P., "Compressing macroscopic near-field digital holograms with wave atoms," in [*OSA Imaging and Applied Optics – Digital Holography and 3D Imaging (3DHD)*], (2018).

[9] A. El Rhammad, P. Gioia, A. G. M. C. and Pesquet-Popescu, B., "View-dependent compression of digital holograms based on matching pursuit," *SPIE Photonics Europe – Optics, Photonics and Digital Technologies for Imaging Applications* **10679** (2018).

[10] D. Blinder, H. Ottevaere, A. M. P. S., "Efficient multiscale phase unwrapping methodology with modulo wavelet transform," *Optics express* **24**(20), 23094–23108 (2016).

[11] Peixeiro, J. P., Brites, C., Ascenso, J., and Pereira, F., "Holographic data coding: Benchmarking and extending hevc with adapted transforms," *IEEE Transactions on Multimedia* **20**, 282–297 (Feb 2018).

[12] Senoh, T., Wakunami, K., Ichihashi, Y., Sasaki, H., Oi, R., and Yamamoto, K., "Multiview image and depth map coding for holographic tv system," *Optical Engineering* **53**(11), 112302 (2014).

[13] Blinder, D., Schretter, C., Ottevaere, H., Munteanu, A., and Schelkens, P., "Unitary transforms using time-frequency warping for digital holograms of deep scenes," *IEEE Transactions on Computational Imaging* **4**, 206–218 (June 2018).

[14] Z. Y. Alpaslan, H. S. E.-G. and Schelkens, P., "Information processing challenges of full parallax light field displays," *SPIE Photonics Europe – Optics, Photonics and Digital Technologies for Imaging Applications* **10679** (2018).

[15] "ISO/IEC JTC1/SC29/WG (JPEG) Website." https://jpeg.org/jpegpleno/. (Accessed: 25 June 2018).

[16] Ebrahimi, T., Foessel, S., Pereira, F., and Schelkens, P., "Jpeg pleno: Toward an efficient representation of visual reality," *IEEE MultiMedia* **23**, 14–20 (Oct 2016).