Wave Atoms for Lossy Compression of Digital Holograms

Tobias Birnbaum∗†, Ayyoub Ahar∗†, Colas Schretter∗†, David Blinder∗†, Tomasz Kozacki‡, and Peter Schelkens∗†

∗Vrije Universiteit Brussel – ETRO
Pleinlaan 2, 1050 Brussels, Belgium
tobias.birnbaum@vub.be
†imec
Kapeldreef 75, 3001 Leuven, Belgium
t.kozacki@mchtr.pw.edu.pl
‡Warsaw University of Technology
8 Sw. Andrzeja Boboli St., 02-525 Warsaw, Poland

Abstract
Compression of digital holograms is a major challenge that needs to be resolved to enable the efficient storage, transmission and rendering of macroscopic holographic signals. In this work, we propose to deploy the wave atom transform that has been utilized before for interferometric modalities such as acoustic and seismic signals. This non-adaptive multi-resolution transform has good space-frequency localization and its orthonormal basis is suitable for sparsifying holographic signals. By replacing the CDF 9/7 wavelet transform stage in a JPEG 2000 codec with the proposed wave atom transform, we did assess its suitability for coding complex amplitude wavefronts. Experimental results demonstrate improved rate-distortion performance with respect to JPEG 2000 and H.265/HEVC for a set of computer-generated, diffuse, macroscopic holograms.

1 Introduction
Holography provides the means to capture a planar cross section of the fully general plenoptic function [1], in theory indistinguishable from reality. All visual cues, such as depth perception and continuous parallax can be accounted for without any eye vergence-accommodation conflict. Holograms also record the phase information of light wavefronts, in addition to intensities. Under the wave model of light transport, light can be described as a complex-valued scalar field such that the knowledge of amplitude (square-root of the intensity) and phase is complete1.

Phase information cannot be measured directly by current digital sensors and therefore indirect measurement strategies must be deployed, such as multi-exposure in-line interferometry or off-axis geometries. When holographic signals are generated synthetically; however, it is not a problem to compute complex-valued incident wavefronts on a virtual detector plane, using numerical simulation of light transport. Conceptually, the most simplistic computer-generated hologram (CGH) is calculated using splatting or sometimes ray-tracing. Every fundamental unit of a given 3D point cloud model is the origin of a complex-valued spherical wave. Its value on any given plane (or manifold) can be described by a point-spread function. The sum, (i.e., the interference), of all those point-spread functions in the detector plane form the hologram of the whole object. If this detector is placed at a large distance away from the object that shall be recorded, the hologram will be visually different from a natural image, due to the interference effects, see Fig. 1.

1Hence the name ”Hologram” derived from the Greek translation of ”whole recording”. 
The well-known grating equation [2] links diffraction angles to spatial frequencies in the hologram. The larger the highest detectable frequencies are, the larger are the maximal diffraction angles, and therefore the field of view under which a given scene can be observed through the hologram. Despite the different hologram types that exist, the Nyquist-Shannon sampling theorem will limit the maximal frequencies detectable by a hologram of a certain size.

To give an example of the required amounts of data: a plane-wave recorded, digital hologram of a 10 cm object with diffuse reflectance properties and with field of view as large as $\pm 36^\circ$, recorded with visible light requires resolutions of $\geq 100$ Mpixels. This necessitates the development of visually lossless compression techniques [3].

We like to distinguish two different approaches for hologram compression: either encoding the complex-valued hologram as it is recorded on the detector plane or using first a numerical propagation scheme to refocus the hologram to a chosen in-focus object plane before encoding. The second variant assumes that the scene is for the most part located within the in-focus plane: an hypothesis that does not hold for macroscopic holograms obtained from deep scenes.

A great part of the work on hologram compression in the detector plane has been based on standard image and video compression, and their extensions, applied to the real and imaginary parts or amplitude and phase separately [4–6]. Leaving the design of the image coding techniques unaltered, i.e., adhering to the $\sim 1/f^2$ decay of spectral amplitude with spatial frequency $f$, will fail on the considered holograms, which possess an almost homogeneous spectral amplitude distribution. Improvements on the compression efficiency could be achieved either by generalization of the wavelet transforms [7, 8] or by devising completely different designs such as [9, 10]. The latter ones consider a viewing perspective as given and encode only the relevant information from the 4D space-frequency domain, using Gabor wavelets. As they encode only parts of the information they are useful for compression in a client-server modality but do not improve on the compression of the full hologram for the purpose of preservation.

With the current work we add the wave atom transform [11] to this list. It has not
been investigated yet for the purpose of holographic data compression. The existence of a critically sampled multi-resolution representation, that is near-optimal localized in space and frequency, which on top can be computed using Fast Fourier Transforms (FFTs), make it however very attractive.

In the remainder of this paper, we first introduce the wave atom transform in Section 2, focusing on its useful properties in an holography context. Next, we detail the proposed compression pipeline for holographic content in Section 3. In Section 4, the proposed method is benchmarked against the JPEG 2000 standard and the H.265/HEVC video compression standard configured in intra-coding mode. Finally, a summary and an outlook to future work is provided in Section 5.

2 Wave atom transform

Wave atoms\[^{[11, 12]}\] are in essence a non-adaptive construction of wave packets with compact support. Starting simple, real-valued wave atoms in the 1D case at scale \(j\), centered around \(\pm \omega_{j,m} = \pm \pi 2^j m\) in frequency domain and around \(x_{j,n} = 2^{-j} n\) in real space with indices \(m, n\) respectively, are given as:

\[
\psi_{j,m,n}(x) = 2^{j/2} \mathcal{F}^{-1}(\hat{\psi}_0^m)(2^j x - n) \tag{1}
\]

with the symmetric, 2 bump "mother" wave atom

\[
\hat{\psi}_0^m(\omega) = e^{-\omega/2} [e^{i \alpha_m} g(\epsilon_m (\omega - \pi (m + 0.5))) + e^{-i \alpha_m} g(\epsilon_{m+1}(\omega + \pi (m + 0.5)))] \tag{2}
\]

for \(j \geq 0, 0 \leq C_1 2^j \leq m \leq C_2 2^j\), arbitrary positive constants \(C_1, C_2, \alpha_m = \pi/2(m + 0.5), \epsilon_m = (-1)^m\). \(g\) is any real-valued \(C^\infty\) bump function with compact support on an interval of length \(2\pi\) that satisfies

\[
\sum_m |\hat{\psi}_m^0(\omega)|^2 = 1 , \tag{3}
\]

and preserves ortho-normality via an asymmetry support condition

\[
\forall \omega \in [-\frac{\pi}{3}, \frac{\pi}{3}] : g(-2\omega - \frac{\pi}{2}) = g(\omega + \frac{\pi}{2}) . \tag{4}
\]

Typical candidates for \(g\) are based on local cosine bases and the specific choice for \(g\) used in our and the reference implementation can be found in the file \(g.m\) of the Matlab version of the wave atom toolbox \[^{[13]}\]. The scales in space and frequency domain can not be chosen arbitrarily but must obey the parabolic scaling condition. This condition tiles the space and frequency domains such that at each scale the effective wavelength within the individual bumps equals the square root of the diameter of their essential support size. We refer the reader to \[^{[11]}\] for more details.

An example for a wave atom in space and frequency domain is given in Fig. 2a.

Gathering all the allowed scales and translations of the given variant yields an ortho-normal basis with multiple scales and basis elements that oscillate in different
directions. While directionality is less interesting in 1D it becomes relevant for any dimension \( \geq 1 \). The extension is readily done by taking the tensor product of the 1D transform, or in other words applying the transform subsequentially per dimension. While various variants of the transform can be obtained for by splitting up the respective bumps into individual atoms, we prefer the real-valued representation with \( 2^{\text{dim}} \) bumps per wave atom, as it provides us with a critically sampled representation.

The reference literature advocates isotropic tilings in both space or frequency domain but we may choose the aspect ratio according to our data, to avoid unnecessary, and potentially massive, zero-padding up to square dimensions.

For coding purposes, we value the existence of a non-adaptive, non-redundant transform that focuses information into the most suitable scales. While the non-redundant transform is most convenient whenever the overall objective is compression, the non-adaptivity makes the transform fast. Especially when considering that the wave atom transform can be implemented in a computationally efficient manner with (I)FFTs. Being able to decompose data at multiple scales allows in principle the transmission of holographic content with a restricted angular field of view as an early prior to the final view. But what makes wave atoms most suitable for the encoding of holograms is their near-optimal simultaneous localization in space and (spatial) frequency, cf. Fig. 2. It allows to sparsify holograms irrespective of their space-frequency footprints which varies for different hologram types, supported angular fields of view, and distances between object and hologram plane. The stringent but, in terms of natural image processing, untypical space-frequency behaviour of holograms is well documented[14].

The most common alternative with equally good space-frequency localization properties supporting multiple scales, are Gabor wavelets and Gabor frames, which, however, require redundancies over-completeness per scale for reasons of numerical stability.

3 Coding scheme

After introducing the wave atom transform and its relevant properties for hologram compression, in this section, we elaborate on the proposed coding scheme. For conve-
For each part do:

1. 2x 1D orthogonal Wave atom transform
2. Scalar quantization global
3. Entropy coding (EBCOT)

Figure 3: Overview of the Wave Atom Coding (WAC) scheme. Parts in orange are taken from the JPEG 2000 standard.

nient referencing, we term our method Wave Atom Coding (WAC). The core coding architecture is based on the wave atom transform and the bit modelling and entropy coding (EBCOT) of a JPEG 2000 coding engine [15], as illustrated in Fig. 3.

The holographic data is ingested into the coding engine using a real-imaginary representation, i.e. one plane is representing the real component and one plane is representing the imaginary component of the holographic signal. This representation is not only compliant with the wave atom transform, but is also more effective than the amplitude-phase representation since phase information is very hard to handle with 'classic' transforms due to phase wrapping phenomena [3].

First, a separable wave atom transform [13] is performed utilizing a 1D kernel in horizontal and vertical dimension. Please note that as the transform coefficients are real-valued, the independently transformed real and imaginary components produce again real-valued floating point data. Due to the specific time-frequency partitioning the wave atom transform is not defined for arbitrary data lengths, henceforth the data may be zero-padded before transformation up to the next admissible value. The padding is small, for example $< 12\%$, $< 6.25\%$, $< 3.124\%$ for signals of length $N \geq 258$, $\geq 1025$, $\geq 4096$, respectively. Since wave atoms are so well localized in space-frequency any padding will solely result in additional zero entries present in the transform domain.

After transformation, the data is fed into the back-end of a JPEG 2000 encoder. The major advantage of the JPEG 2000 part 2 pipeline is that it provides the necessary tools to handle the float-valued wavelet coefficients. The uniform, scalar dead-zone quantizer will transform floating-point to integer data, whereas EBCOT subsequently performs the rate-distortion optimization as well as entropy coding.

The EBCOT module treats the padded wave atom transformed image as one sub-band and hence equally weights the different wave atom subbands. The main difference related to the default JPEG 2000 encoder is that the 2D CDF 9/7 wavelets transform is replaced by a wave atom transform. The WAC backend and the JPEG 2000
Figure 4: Center reconstructions of the test holograms with focus at the front.

encoding are based on a standard compliant implementation of JPEG 2000 by Bruylants et al. [16].

4 Numerical experiments

4.1 Test data

In the experiments, we deploy four computer-generated Fourier holograms with a spatial resolution of 2048 × 16384 pixels. The recorded scenes, cf. Fig. 4, measure 4 – 5 cm in diameter and range in depth from 4.5 cm up to 31.5 cm. The holograms were transformed into a compact space-bandwidth representation [14]) immediately after generation. This means, their 4D space-frequency footprint remains close to the respective axes, resulting in a near-uniform distribution of spatial frequency content over their whole spatial extents. Holograms of this kind, are readily obtained in color from optical recordings [17] after spatial filtering, too.

To enable reuse of components from natural image compression codecs, as well as to allow for a fair comparison of the compared codecs, the DC term was removed from these floating-point holograms and their range was clipped to \( \approx 3.35 \sqrt{\sigma^2} \). This particular clipping point simultaneously minimizes the quantization and clipping errors assuming Gaussian phase noise on the object surface recorded by the hologram. Finally the real and imaginary parts were jointly 8-bit quantized.

4.2 Testing conditions

The proposed WAC is tested against two standardized coding solutions: JPEG 2000 and H.265/HEVC. WAC is configured with a code block size of 64 × 64. The Fraunhofer implementation of H.265/HEVC (HM 16.19) [18] was used and configured as follows: main-RExt profile with level 8.5, intra period of one and 4:0:0 input chroma format, whilst all other options have been left at their default. For JPEG 2000 the standard compliant implementation of Bruylants et al. [16] was used. The JPEG 2000 code block size was fixed as well at 64 × 64 and a four-level CDF 9/7 decomposition was performed.
4.3 Experimental analysis

We compare first the different codecs in terms of rate-distortion behavior. The Peak Signal-to-Noise Ratio (PSNR) is utilized for distortion measurement, which is albeit its many drawbacks still a generally accepted measure for objective quality evaluation of holograms. The PSNR is measured in the hologram plane before encoding and after decoding. The rate-distortion curves in Fig. 5 show that across all holograms, WAC is competing well with H.265/HEVC for bitrates $\leq 0.75$ bpp and outperforms H.265/HEVC for larger bitrates. WAC depicts significant gains compared to JPEG 2000 for bitrates larger than 0.5 bpp.

![Rate-distortion curves (PSNR)](image)

Figure 5: Rate-distortion curves (PSNR) for WAC, JPEG 2000 and H.265/HEVC.

To assess the visual quality of the holograms, they were numerically reconstructed using Fresnel backpropagation in an object plane position at the front of the 3D object/scene contained by the hologram. An aperture of $2048 \times 2048$ was used. It is noteworthy that for the chosen type of Fourier holograms the degradation of the corner views across all methods was quite small due to the use of the compact space-bandwidth representation. As a general comment, we would like to state that...
with H.265/HEVC and WAC all views reconstructed from the holograms "Airplane", "Ball", and "Earth" were visually almost indistinguishably for bitrates larger than or equal to 2 bpp with PSNRs ≥ 40. For the reconstructed views from the "Chess" hologram only WAC did yield PSNRs larger 40 dB at 2 bpp.

Fig. 6 displays the reconstructions obtained from the "Airplane" hologram using WAC, JPEG 2000, and H.265/HEVC at 0.5 bpp. Visibly, WAC is almost on par with H.265/HEVC, albeit the latter codec uses more advanced coding tools. While H.265/HEVC distributes background noise evenly over the whole hologram and shows first signs of vignetting, WAC preserves the full object and introduces more structured noise. Both show still decent quality. In direct comparison with JPEG 2000, that is CDF 9/7 wavelets instead of wave atoms, we can observe how heavy aliasing and signal dispersion are caused due to the poor high-frequency localization properties of a classical wavelet transform. These assessments remain true for all distances and viewing angles and matches the rate-distortion ranking.
5 Conclusion

The wave atom transform, which was previously a popular choice for reconstruction problems in for example acoustic and seismic imaging, was applied to the compression of holographic signals. We detailed its interesting mathematical properties, among which – most notably – are the near-optimal space-frequency localization, its multi-resolution support, and the existence of an orthonormal basis. By contrasting the wave atom transform with CDF 9/7 wavelets, the former top candidate for macroscopic hologram compression, we showed that wave atoms are a much better suited non-adaptive sparsifying transform for representing complex-valued fringe patterns in holograms. Both in terms of subjective and objective quality, the proposed wave atom codec outperforms JPEG 2000, deploying a CDF 9/7 wavelet transform, and H.265/HEVC, with respect to the latter particularly for bitrates greater than 0.75bpp.

Funding

The research leading to these results has received funding from the European Research Council under the European Union’s Seventh Framework Programme (FP7/2007-2013)/ERC Grant Agreement nr. 617779 (INTERFERE), from the Research Foundation - Flanders (FWO), nr. G024715N, and from Narodowe Centrum Nauki (NCN), Poland (2014/15/N/ST8/03117, 2015/17/B/ST8/02220), Politechnika Warszawska.

References


