Microscopic Off-axis Holographic Image Compression with JPEG 2000

Tim Bruylants\textsuperscript{a,c}, David Blinder\textsuperscript{a,c}, Heidi Ottevaere\textsuperscript{b}, Adrian Munteanu\textsuperscript{a,c} and Peter Schelkens\textsuperscript{a,c}

\textsuperscript{a}Vrije Universiteit Brussel, Dept. of Electronics and Informatics (ETRO), Pleinlaan 2, B-1050 Brussels, Belgium;
\textsuperscript{b}Vrije Universiteit Brussel, Brussels Photonics Team (B-PHOT), Pleinlaan 2, B-1050 Brussels, Belgium;
\textsuperscript{c}iMinds, Dept. of Multimedia Technologies (MT), Gaston Crommenlaan 8 (box 102), B-9050 Ghent, Belgium

ABSTRACT

With the advent of modern computing and imaging technologies, the use of digital holography became practical in many applications such as microscopy, interferometry, non-destructive testing, data encoding, and certification. In this respect the need for an efficient representation technology becomes imminent. However, microscopic holographic off-axis recordings have characteristics that differ significantly from that of regular natural imagery, because they represent a recorded interference pattern that mainly manifests itself in the high-frequency bands. Since regular image compression schemes are typically based on a Laplace frequency distribution, they are unable to optimally represent such holographic data. However, unlike most image codecs, the JPEG 2000 standard can be modified to efficiently cope with images containing such alternative frequency distributions by applying the arbitrary wavelet decomposition of Part 2. As such, employing packet decompositions already significantly improves the compression performance for off-axis holographic images over that of regular image compression schemes. Moreover, extending JPEG 2000 with directional wavelet transforms shows even higher compression efficiency improvements. Such an extension to the standard would only require signaling the applied directions, and would not impact any other existing functionality. In this paper, we show that wavelet packet decomposition combined with directional wavelet transforms provides efficient lossy-to-lossless compression of microscopic off-axis holographic imagery.

Keywords: wavelet, JPEG 2000, off-axis holography, holography, packet decomposition, directional transform, image compression

1. INTRODUCTION

Holography was first discovered in the late 1940s by Dennis Gabor as an unexpected side effect of his research with respect to improving electron microscopes. However, practical applications for recording 3D objects as optical holograms only became possible after the development of the laser in the beginning of the 1960s. And even then, the use of holograms remained limited due to being relatively expensive, complex and inefficient when compared to other simpler (and more limited) 3D recording techniques such as stereoscopic photography. Aside from the acquisition and reconstruction complexity, there is also the requirement for high-resolution recording media – typically 10 to 100 times higher than what is required for regular photography – allowing to accurately capture the interference pattern in order to have acceptable visual quality. Analog holographic recordings typically rely on a film with very high concentrations of light-reactive grains of silver halide. The advent of high-resolution digital light sensors provides a cheaper and more flexible alternative to analog holographic film, especially for dynamic recordings. This also creates the need for compression methodologies to cope with the huge amounts of recorded information. At present, the captured sensor data in holographic microscopy is losslessly compressed...
as classical image encoding algorithms, such as JPEG, are not adequate for coding holographic data due to quantization issues.

In this respect, we start by demonstrating how JPEG 2000\(^1\)\(^–\)\(^3\) can already be used for the efficient compression of digital microscopic off-axis holography images. With the releases of Part 1\(^1\) and Part 2\(^4\) of JPEG 2000, the JPEG committee (ISO/IEC SC29 WG1) aimed at offering a state-of-the-art successor for their very successful JPEG image compression standard. Aside from the obvious desire to significantly improve the compression efficiency over that of the original JPEG standard, WG1 also made the new standard versatile and future proof by providing many functional features including rate-distortion optimization, resolution and quality scalability, region-of-interest coding, progressive decoding, as well as by allowing for extending the standard at later stages through new parts and/or amendments to existing parts (e.g. JPIP, JPWL, Motion JPEG 2000, JP3D, ...).

The existing Arbitrary Wavelet Decomposition feature of Part 2 of JPEG 2000 can be used to efficiently cope with the uncommon frequency distributions of off-axis holographic image data. Moreover, it is known that such holographic images often contain distinct directional features that can be further exploited through the use of directional adaptive discrete wavelet transforms (DA-DWT).\(^5\)\(^–\)\(^8\) This paper proposes to extend the JPEG 2000 standard to enable DA-DWT functionality. Using our own JPEG 2000 compliant codec framework,\(^9\) extended with DA-DWT, we are able to provide the experimental results with improvements of up to 7dB in compression efficiency over the conventional JPEG 2000.

Section 2 provides a short introduction into microscopic holographic imaging. Section 3 describes how JPEG 2000 can be used – with and without extending the standard – to efficiently compress this type of image data. This is followed by the experimental results reported in section 5 and a concluding remarks given in section 6.

### 2. MICROSCOPIC HOLOGRAPHIC IMAGES

Holograms allow to fully describe the wavefront of a scene, and thereby to completely resolve captured 3D structures. This is an extremely useful property for many measurement and visualization applications. In particular, the use of digital holography offers many advantages for microscopic applications. Regular microscopes can only provide a two-dimensional snapshot of the intensity, with a single focal plane. Holographic microscopes, on the other hand, are able to capture the full wavefront emanating from the sample, offering several benefits and substantially expanding the available tools for data analysis:

- Besides the ordinary **bright field** image, phase information is available as well. The phase shift image data gives quantifiable information about optical distance. As such, topographical information about a sample is also present.
- The hologram can be digitally refocused at any depth, allowing for quick passive digital auto focus of a sample, as well as to view objects located at different focal planes.
- Holography is referred to as **lensless photography**. As such, it will not suffer from optical aberrations related to imperfections in the image forming lenses that are present in traditional microscopic setups.

Digital holograms are digitized recordings of a light field by means of interferometry. In the general case, a monochromatic and coherent laser beam is widened and further collimated, using a beam expander, to allow complete illumination of the sample object. This beam will be split by a beam splitter into two:

1. a reference beam to serve as a carrier wave, and
2. an object beam to illuminate the sample.

These two beams recombine onto an image sensor (typically a CCD), that captures the interference pattern resulting from the two superposed laser beams.

The recording device will only be able to measure the irradiance \(I\) of the interferogram, given by:

\[
I = (R + O)^2 = |R|^2 + |O|^2 + R \cdot O^* + R^* \cdot O
\]  \hspace{1cm} (1)
\( R \) and \( O \) represent the amplitudes of the reference and object beams respectively, and \( * \) is the complex conjugate operator.

Reproduction of the hologram requires utilizing the same reference beam \( R \) to illuminate the recorded hologram, effectively modulating the irradiance,\(^{10}\) which then gives:

\[
\Psi = R \cdot I = R \cdot (|R|^2 + |O|^2) + R^2 \cdot O^* + |R|^2 \cdot O
\]  

(2)

Given equation 2, it can be observed that the reproduced hologram will contain additional undesirable terms that will disturb the image to be viewed. Of the three terms of the equation, the first term produces zero-order diffraction and the second term produces a twin image. Only the third term, which is directly proportional to the sought object beam, represents the desired part of the reconstruction.

Separating these terms from each other can be done in a multitude of ways and is the reason why several variants of holographic microscopy exist – i.e. Fourier holography, image plane holography, in-line holography, Gabor holography, phase-shifting digital holography and off-axis (Fresnel) holography.\(^{11}\) The analyzed data in this paper was gathered using the off-axis holography configuration setup.

In off-axis holography, also known as Leith-Upatnieks holography, the reference beam will fall on the CCD at an offset angle \( \theta \neq 0^\circ \), rather than being collinear with the object-beam axis. The detected reference wave will therefore be a tilted plane wave, denoted by \( R e^{i\omega x} \). The spatial frequency \( \omega \) depends on the incidence angle \( \theta: \omega = \frac{2\pi}{\lambda} \sin \frac{\theta}{2} \). The resulting irradiance will now be the following:

\[
I = (R e^{i\omega x} + O)^2 = |R|^2 + |O|^2 + R e^{i\omega x} \cdot O^* + R^* e^{-i\omega x} \cdot O
\]  

(3)

Equation 3 shows that the terms are now effectively being separated in the frequency domain. In principle, a sufficiently high carrier frequency \( \omega \) allows the object beam to be recovered unambiguously.

The presence of these high frequencies in the recording will result in a power spectrum distribution which will significantly differ from the classical \( 1/f^{2} \) distribution found in regular imagery: the distribution will be neither isotropic nor will it have most of the energy found in the low frequencies. These properties motivated us to investigate the following two approaches,\(^{13}\) namely, directional wavelets and wavelet packet decompositions.

### 3. JPEG 2000

#### 3.1 Introduction

Our coding framework is in fact based on JPEG 2000 Part 1, which is a scalable wavelet based image codec, built around the two-tiered Embedded Block Coding by Optimized Truncation (EBCOT)\(^{14,15}\) entropy coder to code the wavelet coefficients. Tier-1 contains the actual entropy coder and Tier-2 is the rate-distortion optimizer that generates packets of encoded data.

The first step in encoding an image with JPEG 2000 consists of a multi-level 2-D discrete wavelet transform (DWT) following the Mallat dyadic decomposition style,\(^{16}\) where only the low-pass sub-bands are further decomposed in the subsequent levels. JPEG 2000 provides two wavelet kernels, both specified as lifting scheme implementations: a) the integer 5x3 kernel for lossless coding and b) the floating-point 9x7 kernel for lossy to near-lossless coding. Both kernels are specified in detail by the JPEG 2000 specification. Because the lifting coefficients of the 5x3 kernel are powers of two, and in order to be able to guarantee lossless reconstruction at the decoder side, the specification restricts its implementation to integer calculus only using well-defined rounding rules. The 9x7 kernel, on the other hand, offers much better energy compaction over the 5x3 kernel. However, due to its coefficients being real numbers, it cannot be fit into an integer-based calculus system, without severely sacrificing the energy compaction performance. For this reason, JPEG 2000 specifies this kernel, using floating-point calculus, making it also an irreversible transform. Extensive results on natural data\(^{17}\) show that the lossy 9x7 wavelet kernel performs better w.r.t. the rate-distortion (RD) performance than the 5x3 wavelet kernel, but only the 5x3 kernel allows for true lossless compression. After the wavelet decomposition step, the resulting sub-bands are further processed with EBCOT.
For each of the sub-bands, EBCOT Tier-1 starts out by grouping the wavelet coefficients into equally sized rectangular areas, so-called code-blocks. Then, Tier-1 performs, on each of these code-blocks, entropy coding by employing context-based binary arithmetic coding. The wavelet coefficients are scanned per bit-plane, starting at the most significant bit-plane, using three types of coding passes in alternating order, namely a significance coding pass, a magnitude refinement coding pass and a cleanup pass. As such, every coefficient bit becomes a member of exactly one of these three coding passes and is encoded into the respective code-block bit-stream. The end of every code-pass, and thus automatically also the end of every bit-plane scan, marks a potential truncation point in the resulting bit-stream. Along with each of these truncation points, Tier-1 also calculates the associated estimated mean square error (MSE) distortion reduction values that drive the Tier-2 rate-distortion optimization process.

EBCOT Tier-2 represents the rate-distortion optimization and packetization process that generates the final code-stream. Given rate and/or quality constraints, the bit-stream chunks of each of the encoded code-blocks from Tier-1 are selected and recombined into packets to form the final JPEG 2000 code-stream. This is done in a global rate-distortion optimal manner by prioritizing bit-stream chunks based on their respective rate-distortion costs over lesser important chunks, while still maintaining causality – i.e. the information dependency between chunks to allow correct decoding. The RD-optimization stops when the rate and/or quality constraints are met, or when all bit-stream chunks are included in the final code-stream. Finally, the necessary JPEG 2000 headers and markers are appended to signal decoding options.

We note that, the wavelet transform in combination with EBCOT makes JPEG 2000 a scalable image codec, both for lossless and lossy compression, with excellent rate-distortion characteristics in the lossy use-case.

### 3.2 Full Packet Decomposition with JPEG 2000

As stated before, due to the nature of off-axis holography, a significant part of important information in these image recordings are contained at high-frequency bands. This contrasts with natural images where most, if not all, of the visually meaningful information resides in the lower frequency bands. Thus, replacing the Mallat dyadic decomposition with a full packet wavelet decomposition allows for improving the compression efficiency by further decorrelating the high-pass sub-band coefficients.

By default, JPEG 2000 features only the Mallat dyadic decomposition. However, the JPEG 2000 Part 2 Arbitrary Decomposition (AD) extension enables the use of alternative decomposition styles, signaled within two additional marker segments in the code-stream. The AD extension starts by specifying the decomposition style in two parts: 1) an underlying decomposition to generate the resolution levels and 2) per resolution level, the extra sub-level decomposition of the respective high-pass sub-bands.

The resolution levels are defined in a similar fashion to Part 1 with the Mallat dyadic decomposition, with the difference being that the splitting of the low-pass sub-band at each level can be either in both horizontal and vertical directions or only one of the two directions. This sequence of resolution reduction splits is signaled in the Down-sampling Factor Styles (DFS) marker as an array, ${Ddfs}$, of two-bit symbols ("1" = both rows and
<table>
<thead>
<tr>
<th>Decomposition style</th>
<th>$N_L$</th>
<th>Ddfs</th>
<th>DOads</th>
<th>DSads</th>
<th>Signaling cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 level Full packet</td>
<td>3</td>
<td>111</td>
<td>321</td>
<td>18</td>
<td>1's</td>
</tr>
<tr>
<td>4 level Partial packet</td>
<td>4</td>
<td>1111</td>
<td>3321</td>
<td>33</td>
<td>1's</td>
</tr>
<tr>
<td>4 level Full packet</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>5 level Full packet</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>5 level FBI</td>
<td>5</td>
<td>11111</td>
<td>2321</td>
<td>111111111111</td>
<td>NA</td>
</tr>
</tbody>
</table>

Table 1. Various well known decomposition styles, using the AD extension syntax (signaling cost reflects additional required bits, excluding the cost for $N_L$).

columns, "2" = rows only and "3" = columns only). In absence of a DFS marker, the decoder assumes both-ways splitting as the default, which is the compliant case with a Part 1 code-stream.

Subsequently, one or more Arbitrary Decomposition Styles (ADS) markers can be used to signal the sub-level decomposition of high-pass sub-bands. Unfortunately, according to the standard, the ADS syntax only allows for two additional decompositions of the high-pass sub-bands, slightly limiting the possible wavelet packet analysis styles. The ADS marker contains two arrays: $D\text{Ddfs}$ specifies the maximum number of split levels per resolution using entries of two-bit values, and $D\text{DSads}$ specifies the type of extra split using two-bit symbols (0 = no extra split, 1 = both rows and columns, 2 = rows only and 3 = columns only). $D\text{Ddfs}$ entries with a value of 1 indicate that no extra high-pass decompositions are required, while values 2 and 3 indicate one and two extra decompositions respectively. The $D\text{DSads}$ array, on the other hand, describes the depth-first traversal of the decomposition tree, with the sub-bands ordered, from the highest resolution to the lowest, and within each level as $HH$-$LH$-$HL$-$LL$, $HX$-$LX$ or $XH$-$XL$, depending on the applied split type.

Hence, the Arbitrary Decomposition extension of JPEG 2000 limits full packet decompositions to $N_L = 3$ levels. As such, the application of four or more levels in such a full packet decomposition style, as shown by figure 1(c), is not possible without modification of the standard. Figure 1(b) visualizes the closest matching decomposition style to full packet with four levels that can be described by the AD extension. Finally, to illustrate how the AD extension works, we show in table 1 some of the more commonly known decomposition styles.

### 4. PROPOSED EXTENSIONS

#### 4.1 Truly Arbitrary Decomposition Styles

To overcome the limitation of the AD extension, we designed our codec to employ an alternative syntax that is able to describe truly arbitrary wavelet decompositions, including full packet decompositions containing more than three levels.

The proposed syntax describes a decomposition style as an ordered array of split-operations that work on a stack of available sub-bands. The split-operations are each represented as a tuple $(s, m, r)$ in the array. Symbol $s$ represents the split type ($XY$ = both rows and columns, $X$ = rows only, $Y$ = columns only and $-$ = termination). A binary pattern mask $m$ indicates which of the resulting sub-bands after splitting will be added on the stack for further processing. And finally, a positive integer value $r \in \mathbb{N}$ for which the meaning depends on the actual split type in $s$ and the value of mask $m$. When $s$ is not a termination symbol and $m$ is non-zero, then $r$ specifies the number of times to recursively repeat the split operation onto all of the respectively generated sub-bands. When $m = 0$, the value of $r$ is not signaled, because it has no purpose. On the other hand, when $s$ is equal to $-$, then $r$ specifies the number of times to repeat this termination operation to the stack of sub-bands (thus one termination is applied when $r = 0$). Applying the termination operation on the sub-band stack is identical to removing the top element. Furthermore, the selected split type $s$ also determines the number of bits in $m$ and their relation with the generated sub-bands, as given in table 2. A value of $m_n = 1$ marks the respective sub-band for further processing, while $m_n = 0$ signals termination. When $m$ contains no bits, its value is considered to be $m = 0$, and as such $m$ is simply not signaled (as stated before).

The decomposition process starts with $LL_0$ (i.e. the original image) as the only available sub-band on the stack. Subsequently, the process iterates over the list of tuples and with each tuple the according split-operation
Table 2. This table shows the relation between the applied split type, signaled by $s$, and the definition of mask bits $m_n$ indicating termination of the decomposition ($m_n = 1$ marks the respective sub-band for further processing, while $m_n = 0$ signals termination). It also specifies the bit-stream encoding for $r$, with $L_{sbs}$ equal to the sub-band stack size just before processing the respective tuple.

<table>
<thead>
<tr>
<th>Split type $s$</th>
<th>Code-words for $s$</th>
<th>Generated sub-bands (in order)</th>
<th>Associated mask bits</th>
<th>Bit coding for $r$</th>
<th>No. bits for $r$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$XY$</td>
<td>$11_b$</td>
<td>$HH, LH, HL, LL$</td>
<td>$m_3, m_2, m_1, m_0$</td>
<td>Golomb</td>
<td>$r + 1$</td>
</tr>
<tr>
<td>$X-$</td>
<td>$10_b$</td>
<td>$HX, LX$</td>
<td>$m_1, m_0$</td>
<td>Golomb</td>
<td>$r + 1$</td>
</tr>
<tr>
<td>$-Y$</td>
<td>$01_b$</td>
<td>$XH, XL$</td>
<td>$m_1, m_0$</td>
<td>Golomb</td>
<td>$r + 1$</td>
</tr>
<tr>
<td>--</td>
<td>$00_b$</td>
<td>none</td>
<td>none</td>
<td>Raw</td>
<td>$\lceil \log_2 L_{sbs} \rceil$</td>
</tr>
</tbody>
</table>

Table 3. Various well known decomposition styles, using our proposed syntax.

<table>
<thead>
<tr>
<th>Decomposition style</th>
<th>Operation tuples</th>
<th>Signaling cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 level Full packet</td>
<td>$(XY, 1111_b, 2)$</td>
<td>9 bits</td>
</tr>
<tr>
<td>4 level Partial packet</td>
<td>$(XY, 1111_b, 2)$</td>
<td>5 bits</td>
</tr>
<tr>
<td></td>
<td>$(XY, 0000_b, 0)$</td>
<td></td>
</tr>
<tr>
<td>4 level Full packet</td>
<td>$(XY, 1111_b, 3)$</td>
<td>10 bits</td>
</tr>
<tr>
<td>5 level Full packet</td>
<td>$(XY, 1111_b, 4)$</td>
<td>11 bits</td>
</tr>
<tr>
<td>5 level FBI</td>
<td>$(XY, 1111_b, 1)$</td>
<td>65 bits</td>
</tr>
</tbody>
</table>

is executed. Generated sub-bands that result from a split operation are immediately pushed onto the sub-band stack, following $HH-LH-HL-LL$, $HX-LX$ or $XH-XL$ ordering. The process ends when all tuples on the list are processed (left-over sub-bands on the stack are not further processed). It is an error when the sub-band stack becomes empty, while not all tuples are consumed. Table 3 shows how to specify some commonly used decomposition styles. Please note that, although possible, in practice the extended syntax will not be used to specify a Mallat dyadic decomposition style, as this is the JPEG 2000 default anyways.

Signaling of the decomposition style in the final code-stream happens via a newly proposed XAD marker that encapsulates the binary representation of the array of split-operation tuples, padded to a byte boundary with 0-bits. The marker length field $L_{\text{XAD}}$ allows to determine the number of elements in array. It is valid in the main header and also in tile-component headers to allow defining different decomposition styles per tile-component.

### 4.2 Directional Transforms

A salient feature of off-axis digital holographic image data are the strongly oriented interference fringes. This hints that the use of directional wavelet transforms can further improve the compression performance, as they are able to align with the directional features of the data.\textsuperscript{13} As such, we show how the JPEG 2000 architecture can be easily extended to include the block-based Directional Adaptive DWT (DA-DWT).

We employ a separable lifting scheme, similar to that of JPEG 2000 DWT, but with modified prediction and update functions that are no longer confined to only the horizontal direction (1, 0) for row-based splits and the vertical direction (0, 1) for column-based splits. Doing so enables the directional DWT to adapt to local geometric features by adjusting its operational direction. However, all of the applied direction vectors are also required at the decoder side in order to perform the inverse DA-DWT operation. From a compression performance point of view, it is evident that the unavoidable increment in rate for signaling these directions to the decoder should not jeopardize the rate reduction caused by the improved energy compaction of the transform. Thus, in practice, the adaptability of the directional DWT is restricted by 1) allowing the selection of only one vector per block of samples for the row and column splits, by 2) limiting the directions to a discrete set of vectors, and by 3) only performing the DA-DWT on low-pass sub-bands (i.e. LL, XL or LX bands).
With the first restriction, directional adaptive (DA) blocks are defined in an identical way as JPEG 2000's code-blocks and precincts. Per decomposition level, they represent a grid of equally sized rectangles that anchor at (0,0) with their dimensions restricted to powers of two. The width and height parameters are signaled in a dedicated marker segment, which we label XDA, for the DA-DWT. As such, the smallest possible DA-block is 4x4.

Secondly, the proposed extension uses the following set of \((x,y)\) vectors for row-based splits \{\((1,0), (3,1), (3,2), (1,1), (1,2), (1,3), (-1,3), (-1,2), (-1,1), (-3,2), (-3,1)\}\} and \{\((0,1), (1,3), (2,3), (1,1), (2,1), (3,1), (3,1), (3,-1), (2,-1), (1,-1), (2,-3), (1,-3)\}\} for column-based splits. Note that inclusion of the vectors \((1,0)\) and \((0,1)\) allow the DA-DWT to fall back to the classic DWT in the case that no dominant direction is present. Thus, each vector can be represented as an index in the set of available vectors. Per DA-DWT level, we use the JPEG 2000 tag-tree system to encode the two grids of direction indexes (one for the row-based split and one for the column-based split). The actual tag-tree values (two for every DA-block) are coded in synchronization with the first instance of any of the possibly associated code-blocks – i.e. depending on the chosen dimensions of code-blocks and DA-blocks, each DA-block can relate to one or more code-blocks. This also means that, at very low bit-rate constraints, it can happen that no code-block contributions exist whatsoever for a specific DA-block. In such case, the associate direction indexes are simply skipped and not encoded in the tag-trees. The resulting encoded bit-stream is signaled in the XDA marker segment.

Thirdly, the restriction to only allow the DA-DWT on the low-pass sub-bands does not negatively impact the compression performance capabilities of the coding framework. As demonstrated in figure 2, an intrinsic property of the directional DWT causes resulting high-pass coefficients to be already horizontally or vertically aligned after the directional wavelet prediction. As such, a single parameter (i.e. one byte) in the XDA marker segment signals the number of decomposition levels that use the DA-DWT, starting at \(LL_0\).

Technically, an encoder implementation is free to use any type of direction vector selection mechanism to drive the forward DA-DWT. Our framework takes a somewhat naive approach by trying all directions and selecting the one that minimizes the \(L_1\)-norm of the high-pass coefficients.

5. EXPERIMENTS

5.1 Off-axis holographic image data

Figure 3 lists the five off-axis holographic captures that were used for the experiments in this work (data courtesy of Lyncée Tec SA). All images were captured at 8 bits per pixel (bpp). They were selected to represent a wide sample of typical applications in the field of off-axis microscopic holography.
5.2 Objective quality metric

The lossy compression results, presented in this paper, were generated using a standard JPEG 2000 rate-distortion optimizer, which is based on mean-square-error (MSE) estimation to measure distortion. As such, we use Peak-Signal-to-Noise-Ratio (PSNR) results as the objective quality metric, sampled at various bit-rates, ranging between 0.125 and 2 bpp. PSNR is defined in equation 4, with $I$ the original image, $\hat{I}$ the reconstructed image, $A_{\text{max}} = 2^8 - 1 = 255$ the maximal sample value, and:

$$\text{PSNR} = 10 \cdot \log \left( \frac{A_{\text{max}}^2}{\text{MSE}} \right), \quad \text{with MSE} = \frac{1}{n} \sum_{i=0}^{n} (I_i^2 - \hat{I}_i^2)$$

Subsequently, we apply the Bjøntegaard delta metric,\textsuperscript{18} which is a commonly accepted metric for image codec evaluations, to summarize and present the results per image and per tested setup.

On the other hand, in the case of lossless coding results, we only present rate reductions relative to the rate of the 4 level Mallat decomposition style to measure the compression performance, as reconstruction is perfect.

Finally, we would like to remark that the use of a MSE-based distortion metric, such as PSNR, taken globally over the full image, is not necessarily optimal to assess the visual quality of holographic images. Depending on the holography use case (e.g. depth-map extraction, multi-focal planes, ...) certain image features become more important than others. This means that a global MSE approach does not always relate well to the visual quality in more specific use-cases. In fact, this is a topic for further investigation and beyond the scope of this paper.

5.3 Decomposition styles and settings

The purpose of these experiments is to assess whether JPEG 2000 or more generically a JPEG 2000 compatible coding architecture can be used to efficiently compress off-axis holographic image data. As such, it is evident
Table 4. Compression results for lossless compression, using the 5x3 kernel. Column-notations use abbreviations for Mallat (M), Partial Packet(PP), Full Packet (FP) and DA-DWT enabled (+DA), preceded by the number of levels. The numbers represent bit-rate reductions (in bpp), when compared to the standard 4 level Mallat decomposition.

<table>
<thead>
<tr>
<th>Image</th>
<th>3FP</th>
<th>4PP</th>
<th>4FP</th>
<th>3FP+DA</th>
<th>4M+DA</th>
<th>4PP+DA</th>
<th>4FP+DA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neuron</td>
<td>0.31</td>
<td>0.34</td>
<td>0.37</td>
<td>0.41</td>
<td>0.45</td>
<td>0.42</td>
<td>0.39</td>
</tr>
<tr>
<td>Erythocryte</td>
<td>0.32</td>
<td>0.38</td>
<td>0.40</td>
<td>0.55</td>
<td>0.61</td>
<td>0.54</td>
<td>0.50</td>
</tr>
<tr>
<td>Microlenses</td>
<td>0.03</td>
<td>0.02</td>
<td>-0.04</td>
<td>0.08</td>
<td>0.17</td>
<td>0.07</td>
<td>0.00</td>
</tr>
<tr>
<td>Ball</td>
<td>0.25</td>
<td>0.24</td>
<td>0.21</td>
<td>0.50</td>
<td>0.51</td>
<td>0.49</td>
<td>0.45</td>
</tr>
<tr>
<td>Scratch</td>
<td>0.11</td>
<td>0.12</td>
<td>0.10</td>
<td>0.10</td>
<td>0.11</td>
<td>0.11</td>
<td>0.08</td>
</tr>
</tbody>
</table>

Table 5. Compression results for lossy compression, using the 9x7 kernel. Column-notations use abbreviations for Mallat (M), Partial Packet(PP), Full Packet (FP) and DA-DWT enabled (+DA), preceded by the number of levels. The numbers represent BD-PSNR improvements (in dB), when compared to the standard 4 level Mallat decomposition.

<table>
<thead>
<tr>
<th>Image</th>
<th>3FP</th>
<th>4PP</th>
<th>4FP</th>
<th>3FP+DA</th>
<th>4M+DA</th>
<th>4PP+DA</th>
<th>4FP+DA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neuron</td>
<td>4.44</td>
<td>5.15</td>
<td>5.77</td>
<td>5.66</td>
<td>5.72</td>
<td>5.96</td>
<td>5.96</td>
</tr>
<tr>
<td>Erythocryte</td>
<td>5.45</td>
<td>6.35</td>
<td>6.78</td>
<td>7.03</td>
<td>7.18</td>
<td>7.19</td>
<td>7.05</td>
</tr>
<tr>
<td>Microlenses</td>
<td>2.77</td>
<td>3.19</td>
<td>3.20</td>
<td>3.13</td>
<td>2.86</td>
<td>3.22</td>
<td>3.11</td>
</tr>
<tr>
<td>Ball</td>
<td>4.36</td>
<td>4.50</td>
<td>4.86</td>
<td>5.72</td>
<td>5.09</td>
<td>5.73</td>
<td>5.72</td>
</tr>
<tr>
<td>Scratch</td>
<td>2.36</td>
<td>2.53</td>
<td>2.65</td>
<td>2.27</td>
<td>1.99</td>
<td>2.37</td>
<td>2.35</td>
</tr>
</tbody>
</table>

to include a setup that is fully compliant with the JPEG 2000 specification. We categorize settings here in two parts, in order to provide results that are compliant with JPEG 2000 Part 1 and settings that additionally rely on the AD extension of Part 2. More specifically, we test the ”Mallat dyadic”, the ”3-level Full packet” and the ”4-level Partial packet” decomposition styles as listed in table 1.

However, using our proposed extended syntax for the decomposition styles, we also test the ”4-level Full packet” decompositions, as given in table 3. Given the limited resolution of the image data, a decomposition tree of 4 levels will be sufficient to reach optimal energy compaction. In addition to that, our framework also provides directional wavelets, for which we also present the results. In combination with the described decomposition styles, we include results using a DA-DWT on the first three levels, applied only on the low-pass sub-bands with DA-blocks of 32x32. Please note that the results include the extra cost for coding the direction vectors, using the described tag-tree encoding methodology.

For the lossless compression of the set of images, we make use of the standard integer based 5x3 wavelet kernel. For the lossy compression results, we rely on the more efficient, but inherently lossy, 9x7 kernel. All the experiments use code-blocks of 32x32, and precincts and tiling are disabled.

5.4 Results

Table 4 summarizes the lossless compression results presented as the bit-rate gains relative to the lossless rate when using the classic 5x3 wavelet kernel making use of the default 4 level Mallat wavelet decomposition. These results clearly show that the largest compression efficiency gain is obtained by enabling the DA-DWT transform in combination with the 4-level Mallat decomposition style. The results also show that the various packet decomposition styles applied alone, still improve the compression efficiency significantly, but less so than for the DA-DWT enabled tests.

Table 5 shows the Bjøntegaard delta PSNR results relative to the default 4-level Mallat configuration when using the 9x7 wavelet kernel. These results indicate that the 9x7 kernel with lossy coding provides the largest compression performance gain when applying the 4-level Full Packet decomposition style, in combination with the DA-DWT transform. Again, similar to the lossless coding results, the packet decomposition styles alone also offer a significant compression performance boost compared to the conventional JPEG 2000 system.

Taking the results from both table 4 and 5 shows that even in a JPEG 2000 constrained application, the compression efficiency for off-axis holographic images can seriously benefit from the use of a packet decomposition style, such as the 3-level Full Packet or 4-level Partial Packet styles. However, applying the two proposed extensions, allows for enhancing the compression efficiency even more. When computational resources are restricted at the encoder, it might be useful to only apply a 4-level packet decomposition, and leave the DA-DWT behind.
But, when these restrictions are not an issue, the DA-DWT proves to be a very powerful tool in significantly enhancing the compression performance for off-axis microscopic holographic data.

6. CONCLUSIONS

In this work, we show how JPEG 2000 can be efficiently used to compress off-axis microscopic holographic data. Moreover, we present two extensions for JPEG 2000 that overcome some of its limitations when handling holographic imagery.

The first proposed extension enhances the existing Arbitrary Decomposition Styles feature such that any decomposition style becomes possible. Along with this extension, we also provide the means to efficiently signal these arbitrary decomposition styles in the code-stream. Our proposed code-stream syntax for the XAD marker requires up to 10 times less rate for the header syntax than that of JPEG 2000’s AD syntax (ADS and DFS markers). Despite the fact that the header syntax signaling is negligible in comparison to the overall size of the complete code-stream, our XAD implementation also has a lower complexity. Switching from the Mallat decomposition style to the Full Packet decomposition style yields gains in lossless compression of up to 0.4 bpp with the 5x3 kernel and BD-PSNR improvements of over 6 dB in lossy compression.

Our second extension introduces a practical implementation of a block-based directional adaptive DWT for JPEG 2000 (DA-DWT). From the results, it is clear that the compression performance for off-axis microscopic holography data benefits from employing the DA-DWT, even if the direction vectors need to be signaled in the code-stream. With the application of the DA-DWT, for the lossy coding, the compression performance additionally improves with up to 1dB in comparison with the non-DA-DWT results.

Finally, we observe that both the packet decompositions and the DA-DWT are able to provide a comparable initial compression performance gain of up to 6 dB over the original JPEG 2000. The combination of both can still deliver additional, though smaller, gains, accumulating to a total of more than 7 dB. Thus, combining both extensions, offers a significant compression boost that justifies both extensions.

ACKNOWLEDGMENTS

We would like to thank the Lynce Tec SA (Lausanne, Switzerland) for providing the digital holographic recordings used in these experiments. This research was supported by the Research Foundation Flanders (FWO) with project No. G014610N.

REFERENCES


