

Visually lossless compression of digital hologram sequences

Emmanouil Darakis,^{a,b} Marcin Kowiel,^a Risto Näsänen,^a and Thomas J. Naughton^{b,a}

^a RFMedia Laboratory, University of Oulu, Oulu Southern Institute, 84100 Ylivieska, Finland;

^b Department of Computer Science, National University of Ireland - Maynooth,
County Kildare, Ireland

ABSTRACT

Digital hologram sequences have great potential for the recording of 3D scenes of moving macroscopic objects as their numerical reconstruction can yield a range of perspective views of the scene. Digital holograms inherently have large information content and lossless coding of holographic data is rather inefficient due to the speckled nature of the interference fringes they contain.

Lossy coding of still holograms and hologram sequences has shown promising results. By definition, lossy compression introduces errors in the reconstruction. In all of the previous studies, numerical metrics were used to measure the compression error and through it, the coding quality. Digital hologram reconstructions are highly speckled and the speckle pattern is very sensitive to data changes. Hence, numerical quality metrics can be misleading. For example, for low compression ratios, a numerically significant coding error can have visually negligible effects. Yet, in several cases, it is of high interest to know how much lossy compression can be achieved, while maintaining the reconstruction quality at visually lossless levels.

Using an experimental threshold estimation method, the staircase algorithm, we determined the highest compression ratio that was not perceptible to human observers for objects compressed with Dirac and MPEG-4 compression methods. This level of compression can be regarded as the point below which compression is perceptually lossless although physically the compression is lossy. It was found that up to 4 to 7.5 fold compression can be obtained with the above methods without any perceptible change in the appearance of video sequences.

Keywords: Visually lossless digital hologram coding, Digital hologram visual quality assessment, Digital holographic data coding

1. INTRODUCTION

Digital holography is a technique where whole light wavefields of scenes are digitally recorded. Digital holography has great potential for 3D imaging as numerous different reconstructions, for example focused at different depths or viewed from different viewing angles can be obtained by numerically processing a single recorded hologram.¹

Inherently, digital holograms differ from usual images in several ways. Digital holography in general, require coherent illumination and information is stored in the form of finely structured interference fringes. Digital holograms need to be processed, i.e. numerically propagated to an appropriate distance, yielding reconstructions of the scene. Furthermore, speckle noise caused by the coherent illumination contaminates the reconstructed digital holograms. For these reasons, digital holograms usually require more rigorous processing compared to ordinary images.

Compression of digital holographic data is of high interest in practice. Given that digital holograms depict highly unstructured interference fringes, lossless compression is usually inefficient.² Several, lossy digital hologram compression techniques have been investigated showing promising results.³ Some of these digital hologram coding algorithms, have been based on the use of usual image coding algorithms, whereas others have been tailored for the compression of digital holograms.

Retaining the reconstruction quality of the coded digital holograms as high as possible, is obviously of highest importance. Given the speckle noise apparent in the digital hologram reconstructions, image quality

Further author information: Thomas J. Naughton's e-mail: tomn@cs.nuim.ie.

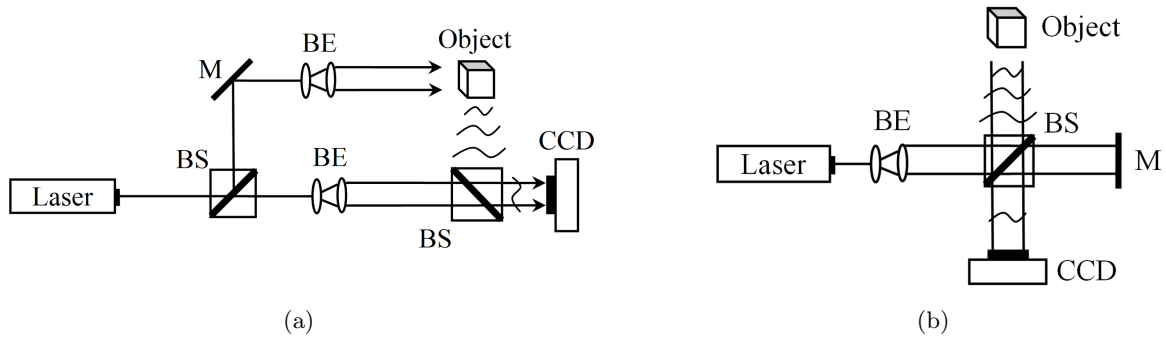


Figure 1: Digital hologram sequence recording setups: (a) Mach-Zehnder and (b) Michelson configuration. BS: Beam Splitter, BE: Beam Expander, M: Mirror.

assessment is not straightforward because the speckle patterns within the reconstructed holograms are very sensitive to holographic data changes. A slight change in the holographic data can yield a substantially different speckle pattern. Nevertheless, although the speckle pattern may change quite a lot, for small compression ratios, the reconstructed scene usually retains its visually perceived quality. As a result, a small amount of lossy compression can yield significant numerical measured compression error, whereas the visually perceived quality of the reconstructed scene might not have changed at all. For this reason, numerical quality metrics can often yield misleading result.

In order to clarify the above question of the perception of the compression errors we performed an experiment with human observers aiming to find out the highest level of compression at which a holographic video sequence cannot be discerned from an uncompressed video sequence. In the experiment the observers were shown two video sequences successively in random order, one compressed and one uncompressed. The observers were asked to indicate which one was compressed. During the threshold measurement the level of compression was varied using the staircase algorithm⁴ to find the threshold level of compression. The study reveals that compression ratios up to $\sim 4 - 7.5$ can be achieved while retaining visually lossless quality.

To the best of our knowledge, this is the first time that the quality of digital hologram reconstructions is to be assessed based on visual perception criteria. As such, the field of digital hologram compression is expected to substantially benefit from the results of this investigation.

2. RECORDING AND COMPRESSION OF DIGITAL HOLOGRAM SEQUENCES

2.1 Recording setup

In order to make our study as general as possible we have used two hologram sequences that were captured with different recording conditions (different setups, laser sources, recording cameras etc). Diagrams of the recording setups that were used are show in Fig. 1.

In both cases, the object is illuminated by a collimated laser beam. We call the light reflected by the object “Object beam” U_O . The object beam propagates towards the camera through a beam splitter, where it interferes with the undiffracted “Reference beam” U_R . The CCD camera captures the hologram which is the intensity of the interference pattern: $I(x, y) = [U_R(x, y) + U_O(x, y)]^2$.

In our case, the object was positioned on a rotation stage which was rotated a small angle ($\sim 3^\circ$) at a time. One hologram was recorded and the procedure was repeated until a full circle (360°) was completed.

The setup shown in Fig. 1 (a) uses the Mach-Zehnder configuration. The laser had a wavelength $\lambda = 785\text{nm}$ and the camera $N_x \times N_y = 1392 \times 1040$ square pixels of size $dx = 6.45\mu\text{m}$ each. The “TwoScrews” sequence, a frame of which is show in in Fig.2(a) was recorded with this setup.

The other digital hologram sequence which we call “Coil” was recorded with the setup show in Fig. 1 (b). The laser had a wavelength $\lambda = 633\text{nm}$ and the camera $N_x \times N_y = 2048 \times 2032$ square pixels of size $dx = 7.4\mu\text{m}$

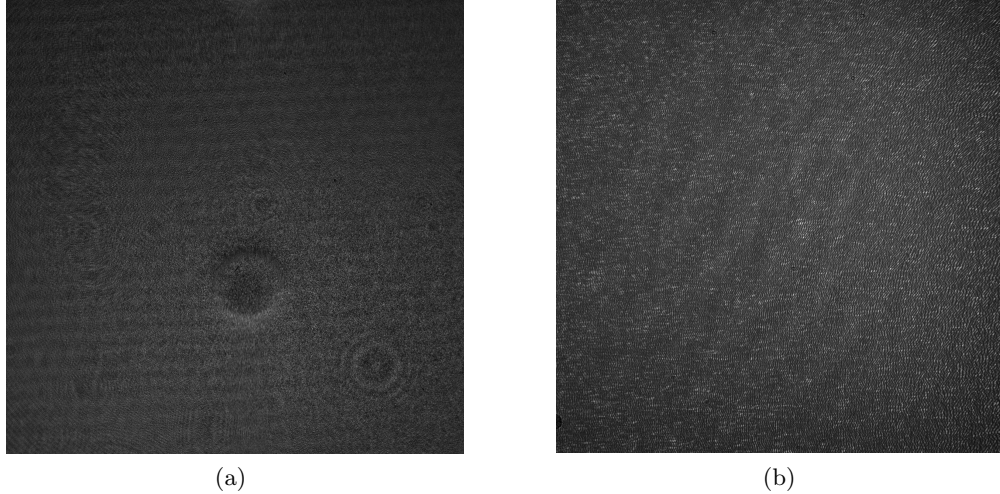


Figure 2: A frame from (a) the “Coil” and (b) the “TwoScrews” hologram sequences.

each. A captured frame of the sequence is shown in Fig. 2(b) and a sequence of the reconstructed frames at Fig. 3(b).

2.2 Compression of hologram sequences

Following recording the captured holograms were fed to the compression algorithms. In this work we have chosen two video standards: MPEG-4 Part 10,⁵ also known as H.264 or Advanced Video Coding (AVC) and the Dirac video format.⁶

Compared to earlier versions of MPEG-4 (like MPEG-4 Part 2), Part 10 offers more optimization options targeting low compression (high quality) applications such as streaming of High Definition (HD) videos. Since we are targeting visually lossless compression we selected MPEG-4 Part 10. Dirac⁶ is an open video compression format that uses wavelet for the transform coding. More details on compression algorithms can be found in the literature.⁵⁻⁷

For the implementation of the compression algorithms we used the “ffmpeg” software.⁸ In particular, we used the libraries “libx264”⁸ and “libschrödinger”⁶ for the implementation of the MPEG-4 Part 10 and the Dirac algorithm respectively. The “normal”⁸ compression preset values were used for libx264 and the default options for “libschrödinger”.

2.3 Reconstruction and numerical evaluation of the compression quality

To reconstruct each hologram, $I(x, y)$ is multiplied by the reference wave $U_R(x, y)$ and the result is propagated to the desired distance d using the Fresnel-Kirchhoff integral:⁹

$$U_d(x', y') = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} I(x, y) U_R(x, y) \frac{\exp \left[-i \frac{2\pi}{\lambda} \sqrt{(x-x')^2 + (y-y')^2 + d^2} \right]}{\sqrt{(x-x')^2 + (y-y')^2 + d^2}} dx dy \quad (1)$$

This procedure is repeated for each frame of the hologram sequence. Figs. 3(a) and (b) show the amplitude of the reconstruction ($U_d(x', y')$) from the hologram sequences “Coil” and “TwoScrews” respectively, from the original (uncompressed) data.

To obtain the holograms from the compressed sequences, the individual decompressed frames $\tilde{I}(x, y)$ are obtained using the video decoder. Following this, Eq. 1 is used to reconstruct the decompressed holograms $\tilde{I}(x, y)$ to the desired distance d .

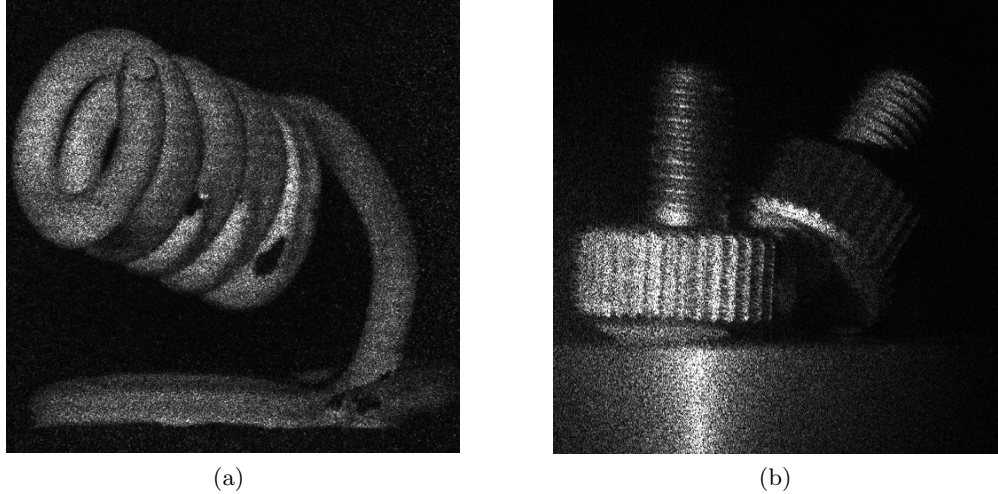


Figure 3: Reconstructions from the uncompressed data of (a) the “Coil” and (b) the “TwoScrews” hologram sequences.

A numerical metric that is usually used to evaluate quality degradation of lossy hologram compression algorithms is the Normalized Root Mean Square (NRMS) error defined as:^{3,10}

$$NRMS = \left[\frac{\sum_{N_x} \sum_{N_y} (|U_d(x, y)|^2 - |\tilde{U}_d(x, y)|^2)^2}{\sum_{N_x} \sum_{N_y} (|U_d(x, y)|^2)^2} \right]^{1/2} \quad (2)$$

where $U_d(x, y)$ and $\tilde{U}_d(x, y)$ are the object plane complex wavefronts from the original and the compressed data, respectively. NRMS is calculated for each frame of the sequence using Eq. 2 and its mean value over the length of the sequence is used. Also, the obtained compression ratio is calculated as $c = S/s$, where S is the size of the original files and s is the size of the compressed video file in bytes.

Fig. 4 shows the numerical results obtained for the two hologram sequences for the two coding algorithms we have used. Note in Fig. 4(a) how nonmonotonic the response of the MPEG-4 curve becomes at high compression ratios. This property was noticed previously for k-means quantisation¹¹ and wavelet compression.¹²

As it can be seen, even for the smallest compression ratio ($\sim 1.5 : 1$) the NRMS error is quite high (~ 0.15) for both cases and for the higher compression ratios ($\sim 25 : 1$) the NRMS metric is below (~ 0.6). It is apparent that it is difficult to evaluate the visually obtained quality from these numerical results.

3. EXPERIMENTAL METHODS

In order to evaluate how observers perceive the obtained compression errors the following experiment was performed.

Video sequences were shown on a 22 inch HP LP2275w LCD display with a resolution of 1650×1050 and a frame ratio of 60Hz. The pixel size was 0.282mm. The video frame ratio used was 20 Hz. The size of the video frame was 512×512 pixels. The viewing distance was 70 cm. Thus, the spatial extent of the video frame was $14.4 \times 14.4\text{cm}^2$, which corresponds to $11.6 \times 11.6\text{deg}^2$ of visual angle.

At each trial, the observer was shown two video sequences, one of which was the original uncompressed sequence and the other was a compressed sequence. For each trial the order of the sequences was chosen at random. The duration of the video sequences was 1.05 s. The interval between the two videos was 0.95 s. Below the video picture there was a graphical response panel with buttons labelled First and Second. The observer was required to indicate which one of the sequences was compressed by clicking one of the two buttons with the mouse.

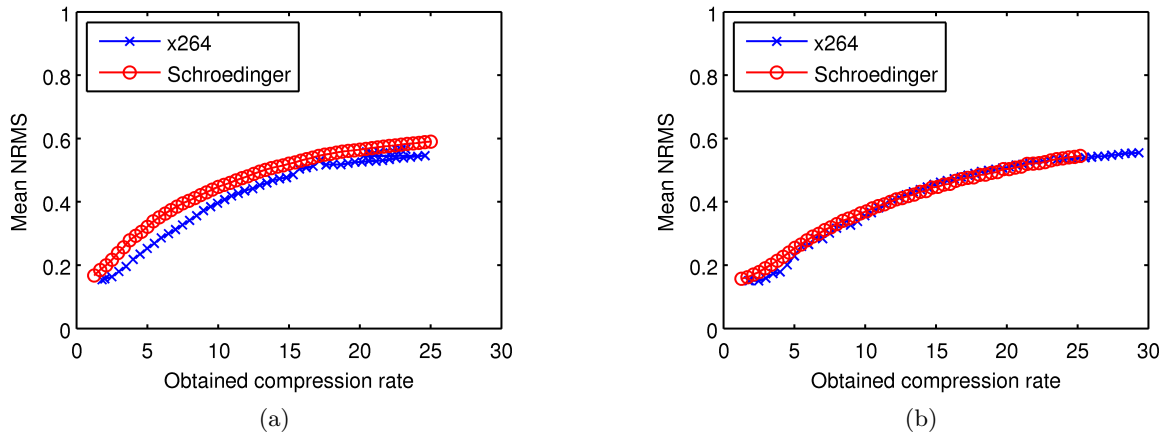


Figure 4: NRMS numerical results (a) for the “Coil” sequence and (b) for the “TwoScrews” sequence, respectively.

A variant of the staircase algorithm⁴ was used to determine the threshold compression ratio. At the beginning, the compression ratio was about 20, and the compressed sequence could easily be discriminated from the uncompressed sequence. Depending on the correctness of the observer’s responses, the compression level was decreased or increased by an amount equal to $3/n + 0.5$, where n is the number of trials done so far. Thus, with increasing number of trials the change in compression ratio gradually approached 0.5. The algorithm proceeded in two stages. During the first stage each correct response was followed by a decrease of the compression ratio. The second stage of the algorithm started after the first incorrect response. During the second stage, four consecutive correct responses led to a decrease of compression ratio and each incorrect response led to an increase of compression ratio. The algorithm automatically adjusted the compression ratio close to a level at which the probability of responding correctly (P) four times consecutively (P^4) was approximately equal to not responding correctly four times ($1 - P^4$). It follows from this that at threshold the probability of responding correctly four times consecutively is equal to 0.5. Thus, the probability of correct responses is $P = 0.5^{1/4} = 0.84$. For each observer in each condition, the threshold estimate for the compression ratio at the probability level of 0.84 of correct responses was obtained as the mean of four reversals. The counting of reversals started after the observer had made one error.

Five observers participated in the experiment. Compression thresholds were measured for two different compression methods, “Dirac” and “MPEG-4”, using two different objects, a coil (“Coil”) and an object comprising two screws (“TwoScrews”). Of these TwoScrews was clearly noisier. Thus, for each observer four different compression thresholds were measured. In order to prevent any systematic learning and fatigue effects during the measurements, the order of measurement of each condition was different for each observer.

4. RESULTS

The results are shown in Table 1. The results show that the median threshold compression ratio varied from about 4 to nearly 7.5. There is variability that seems to depend on compression method and object. In order to evaluate whether the differences between the conditions (i.e. “Coil & Dirac,” “Coil & MPEG-4,” “TwoScrews & Dirac,” “TwoScrews & MPEG-4”) were statistically significant we performed the Friedman one dimensional non-parametric analysis of variance on the data.¹³ The analysis was performed using the values of reversal points of each threshold measurement. The statistical analysis showed that the differences were statistically highly significant ($p < .0001$; $\chi_r^2 = 39.4$; $df = 3$, where χ_r^2 is the test statistic obeying the chi-squared distribution, df is the degrees of freedom, and p is the significance level value, i.e. the probability that the conclusion made is false).

The threshold seem to be affected by object type (Coil vs. TwoScrews) with both compression methods as well as by methods (Dirac vs. MPEG-4) with both object types. Post hoc analysis with the Mann-Whitney

Table 1: Median, 25th and 75th percentiles for the four experimental conditions.

Sequence	Coil		TwoScrews	
Codec	Dirac	MPEG-4	Dirac	MPEG-4
Median	5.42	7.48	6.52	4.00
75th percentile	5.84	7.98	10.83	4.99
25th percentile	5.15	7.08	9.16	4.96

Table 2: Median, 25th and 75th percentiles for normalised root-mean-square errors of the compressed video relative to uncompressed videos at threshold.

Sequence	Coil		TwoScrews	
Codec	Dirac	MPEG-4	Dirac	MPEG-4
Median	0.34	0.33	0.18	0.28
75th percentile	0.35	0.34	0.23	0.38
25th percentile	0.30	0.28	0.17	0.21

test¹⁴ using Bonferroni correction¹⁵ shows that these effects were statistically significant. However, there seems to be an interaction between method and object type. MPEG-4 was better for the “Coil” object than for the “TwoScrews” object, and Dirac was better for the TwoScrews object than for the Coil object. When the median is computed for the results of both objects there was little difference between the two methods: the total median for Dirac was 5.84 and for MPEG-4 5.24.

If the the visibility of compression errors would only depend on the the NRMS error, it should be approximately similar in all experimental conditions. Table 2 shows the average NRMS compression errors corresponding to threshold compression ratios. The NRMS values are rather similar for the Coil objects compressed with Dirac and MPEG-4 and for the TwoScrews object compressed with MPEG-4. However, the NRMS error for the TwoScrews object compressed with Dirac seems to be different. Statistical analysis with the Friedman non-parametric variance analysis showed that the difference is highly significant ($p < .0001$; $\chi_r^2 = 52.3$; $df = 3$). Thus, NRMS error can not be regarded as a reliable predictor of the visibility of compression errors.

5. DISCUSSION AND CONCLUSIONS

The numerical results discussed above show that the NRMS error between uncompressed and compressed video sequences cannot predict reliably the visibility of compression errors.

In addition, the visual experiments showed that the holographic video sequences can be compressed up to a ratio of $\sim 4 - 7.5$ before humans are able see compression errors. Therefore, up to these compression ratios compression is visually lossless. Fig. 5 show reconstructions from data compressed to a ratio of ~ 5.5 where the visually lossless quality in comparisson to Fig 3 can be seen.

Finally, the experimental results showed that the two compression methods “MPEG-4” and “Dirac” seem to be roughly equally good, in terms of visually lossless compression, for the compression of holographic videos.

Lossy compression of holographic video produce changes in reconstruction. However, in this paper it was shown that if the changes are small enough, they may not have any perceptual significance being invisible to the viewer. One reason for this may be that human observers most probably cannot do any point-wise comparison of the compressed and uncompressed video sequences. Instead, they may rely on some sort of overall similarity or dissimilarity of the sequences possibly based on image structure.¹⁶

It follows from this, that the quality of compressed sequences should be evaluated by human observers. However, this may sometimes be too time consuming. Therefore, a more practical alternative would be to apply a computational model that resembles the way human observers compare images. One such method could be the structural similarity index suggested by Wang et al.¹⁶

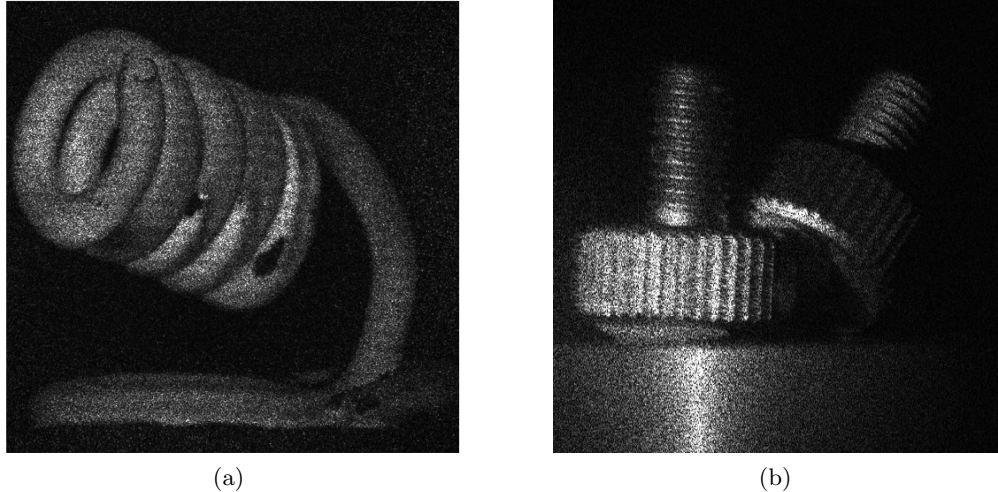


Figure 5: Reconstructions from (a) the “Coil” and (b) the “TwoScrews” hologram sequences compressed to compression ratio of $\sim 5.5 : 1$ using Dirac. According to the analysis the compression is visually lossless at this ratio.

To conclude, in this paper visually lossless compression of hologram sequences has been studied. It was shown that numerical quality metrics such as the NRMS error, that has been extensively used in similar studies, cannot predict reliably how visible the compression errors are. With appropriate experiments and statistical analysis it was shown that when MPEG-4 and Dirac compression algorithms are used to compress digital hologram sequences, compression ratios of $\sim 4 - 7.5$ can be achieved.

Considering that MPEG-4 and Dirac are compression algorithms that have been optimized for the compression of ordinary (non-hologram) video sequences these results are very promising. It is expected that coding algorithms specially tailored for the compression of hologram sequences are expected to be able to achieve even higher compression ratios facilitating the storage and transmission of hologram sequences.

ACKNOWLEDGMENTS

The research leading to these results has received funding from the European Community’s Seventh Framework Programme FP7/2007-2013 under grant agreement no. 216105 (Real 3D), from the Irish Research Council for Science, Engineering and Technology, from Science Foundation Ireland, under the National Development Plan, and from a European Commission Marie Curie Intra-European Fellowship. The authors would also like to thank Nitesh Pandey for assisting with the recording setup.

REFERENCES

- [1] Frauel, Y., Naughton, T. J., Matoba, O., Tajahuerce, E., and Javidi, B., “Three-dimensional imaging and processing using computational holographic imaging,” *Proc. of the IEEE* **94**(3), 636–653 (2006).
- [2] Naughton, T. J., Frauel, Y., Javidi, B., and Tajahuerce, E., “Compression of digital holograms for three-dimensional object reconstruction and recognition,” *Applied Optics* **41**(20), 4124–4132 (2002).
- [3] Darakis, E., Naughton, T. J., Soraghan, J. J., and Javidi, B., “Measurement of compression defects in phase-shifting digital holographic data,” **6311**, 63110B, SPIE (2006).
- [4] Wetherill, G. and Levitt, H., “Sequential estimation of points on a psychometric function,” *J. Math. Stat. Psych.* **18**, 1–10 (1965).
- [5] Richardson, I. E. G., [*H.264 and MPEG-4 Video Compression: Video Coding for Next Generation Multimedia*], John Wiley & Sons (2003).
- [6] “Dirac video compression, <http://www.diracvideo.org/>,” (2009).

- [7] Pereira, F. and Ebrahimi, T., [*The MPEG-4 Book*], Prentice Hall, Upper Saddle River, NJ (2002).
- [8] "<http://www.ffmpeg.org/>," (2009).
- [9] Schnars, U. and Jueptner, W., [*Digital Holography: Digital Hologram Recording, Numerical Reconstruction, and Related Techniques*], Springer, Berlin (2005).
- [10] Darakis, E. and Naughton, T. J., "Compression of digital hologram sequences using MPEG-4," *Holography: Advances and Modern Trends* **7358**(1), 735811, SPIE (2009).
- [11] Shortt, A. E., Naughton, T. J., and Javidi, B., "Compression of optically encrypted digital holograms using artificial neural networks," *Journal of Display Technology* **2**(4), 401–410 (2006).
- [12] Shortt, A. E., Naughton, T. J., and Javidi, B., "Compression of digital holograms of three-dimensional objects using wavelets," *Optics Express* **14**(7), 2625–2630 (2006).
- [13] Friedman, M., "The use of ranks to avoid the assumption of normality implicit in the analysis of variance," *Journal of Statistics Of the American Statistical Association* **32**, 675 – 701 (1937).
- [14] Mann, H. B. and Whitney, D. R., "On a test of whether one of two random variables is stochastically larger than the other," *The Annals of Mathematical Statistics* **18**(1), 50–60 (1947).
- [15] Miller, R. G., [*Simultaneous Statistical Inference*], Springer-Verlag (1981).
- [16] Wang, Z., Bovik, A., Sheikh, H., and Simoncelli, E., "Image quality assessment: from error visibility to structural similarity," *Image Processing, IEEE Transactions on* **13**(4), 600–612 (2004).