FAST DIFFRACTION-PATTERN MATCHING FOR OBJECT DETECTION AND RECOGNITION IN DIGITAL HOLOGRAMS

Mozhdeh Seifi , Loic Denis, Corinne Fournier

Université de Lyon, F-42023, Saint-Etienne, France,
CNRS, UMR5516, Laboratoire Hubert Curien,Université de Saint-Etienne
Telecom Saint Etienne, F-42000, Saint- Etienne, France

ABSTRACT

A digital hologram is a 2-D recording of the diffraction fringes created by 3-D objects under coherent lighting. These fringes encode the shape and 3-D location information of the objects. By simulating re-lighting of the hologram, the 3-D wave field can be reconstructed and a volumetric image of the objects recovered. Rather than performing object detection and identification in this reconstructed volume, we consider direct recognition of diffraction-patterns in in-line holograms and show that it leads to superior performance. The huge variability of diffraction patterns with object shape and 3-D location makes diffraction-pattern matching computationally expensive. We suggest the use of a dimensionality reduction technique to circumvent this limitation and show good detection and recognition performance both on simulated and experimental holograms.

Index Terms— Digital Holography, Pattern Recognition, Singular Value Decomposition, Inverse Problems.

1. INTRODUCTION

Digital holography is an imaging technique used in various domains for its wide field and large depth-of-focus. By capturing the 3-D information in a single 2-D hologram, it can be applied to track moving micro-objects in fields ranging from fluid mechanics [1, 2, 3, 4] to biology [5, 6, 7]. The advantage of in-line digital holography compared to its alternatives is the simple imaging setup (a coherent source plus a lensless imaging sensor) and high accuracy. For example, in particle imaging sub-micron accuracy is achieved for transversal location and sizing, and about 10 microns depth accuracy for a setup without magnification [8].

Following the optical reconstruction used when holograms were recorded on holographic plates and re-illuminated to reveal the 3-D image they encoded, digital holograms are still often digitally processed in two steps: 1) reconstruction of a volumetric image by diffraction simulation, 2) analysis of the obtained 3-D image to detect/recognize objects. Such an approach is however sub-optimal for several reasons. First, diffraction does not invert the recording process but leads to the superimposition of more or less out-of-focus images of the objects and strongly out-of-focus images known as twin (or virtual) images. Second, due to the small physical size of digital sensors, holograms are severely truncated. This truncation leads to artifacts in the reconstructed volume close to the sensor borders.

Direct matching of diffraction patterns on the in-line holograms dramatically improves the quality of reconstructed images [9, 10, 11] and the accuracy of particle hologram analysis [12, 8]. The extension of the latter approaches to the recognition of more general objects than spherical particles is limited by the increase in size of the dictionary of diffraction patterns when considering all possible 3-D locations of the objects. We propose in this paper to circumvent this limitation by applying a dimensionality reduction method to the dictionary and derive an algorithm similar to matching-pursuit to detect several objects directly in a digital hologram. Our proposal can be used to improve the accuracy of fast reconstructions in biological applications that deal with few objects from a known set of classes (see [13] for a potential application on detection and localization of bacteria).

The structure of the paper is as follows: the diffraction-pattern matching problem is presented next section; the dimensionality reduction of the dictionary is then described in section 3; experimental results are reported in section 4.

2. OBJECT RECOGNITION BY DIFFRACTION-PATTERN MATCHING

In Gabor holography, a hologram corresponds to the diffraction pattern of illuminated objects. To produce reconstructed images of satisfying quality, this holographic setup is restricted to small objects spread in mostly empty volumes. Under such conditions, the hologram can be approximated as the (incoherent) sum of the diffraction pattern created by each distinct object. Furthermore, the diffraction pattern of a given object is related to its transmittance through a convolution [10]. A measured n-pixels hologram $\boldsymbol{d}$ can then be described, after proper centering and normalization, using the

following linear model:

$$d = Ht + \epsilon, \qquad (1)$$

where H is a linear operator mapping a 3-D transmittance volume represented as a n'-dimensional vector $t \in \mathbb{R}^{n'}$ to a 2-D hologram $d \in \mathbb{R}^n$ in the presence of white and Gaussian noise $\epsilon \in \mathbb{R}^n$. More precisely, the application of H on t corresponds to the sum of the 2-D convolution of each transversal slice of t with a chirp function [10]:

$$d(x_f, y_f) = \sum_{x_\ell, y_\ell, z_\ell} t(x_\ell, y_\ell, z_\ell) \, h_{z_\ell}(x_f - x_\ell, y_f - y_\ell), \quad (2)$$

with $h_z(x,y) = \sin[\pi(x^2 + y^2)/(\lambda z)]/(\lambda z)$, and where the k-th entry of vector d, corresponding to the pixel with coordinates (x_f, y_f), is written $d(x_f, y_f)$, and similarly the ℓ-th entry of vector t corresponds to spatial coordinates (x_ℓ, y_ℓ, z_ℓ). The convolution kernel h_z is known as Fresnel function. It is a 2-D circularly symmetric chirp function that varies with the depth z and the optical wavelength λ. By denoting with H_{z_i} the discrete convolution with kernel $h(\cdot, \cdot, z_i)$ that approximates diffraction at distance z_i, equation (1) can be rewritten by decomposing H as a block matrix:

$$d = \underbrace{\begin{bmatrix} H_{z_1} & \cdots & H_{z_p} \end{bmatrix}}_{H} t + \epsilon, \qquad (3)$$

where p different z-slices are considered in the 3-D transmittance t.

Since each operator H_{z_i} models diffraction at a given distance z_i, diffraction-based hologram reconstruction is readily performed by:

$$\hat{t}^{(\text{diffraction})} = \begin{bmatrix} H_{z_1} \\ \vdots \\ H_{z_p} \end{bmatrix} d. \qquad (4)$$

If the transverse dimension of the transmittance is chosen equal to that of the hologram, matrices H_{z_i} are square and equation (4) corresponds to a "back-projection" operation, i.e., a reconstruction by application of the adjoint operator[1]. Despite its wide use, reconstruction with the adjoint operator suffers from strong artifacts. When simulating the diffraction of the hologram of a planar object, instead of getting back a crisp image of the object, an in-focus image and strongly out-of-focus image (the twin/virtual image) superimpose:

$$H_z^t H_z \approx I + H_{2z}. \qquad (5)$$

Hologram reconstruction by equation (4) suffers from strong location-dependent artifacts: superimposition of the twin-images but also out-of-focus images of objects located at other distances (since $H_{z_i}^t H_{z_j} \neq 0$ for $z_i \neq z_j$) and strong

---

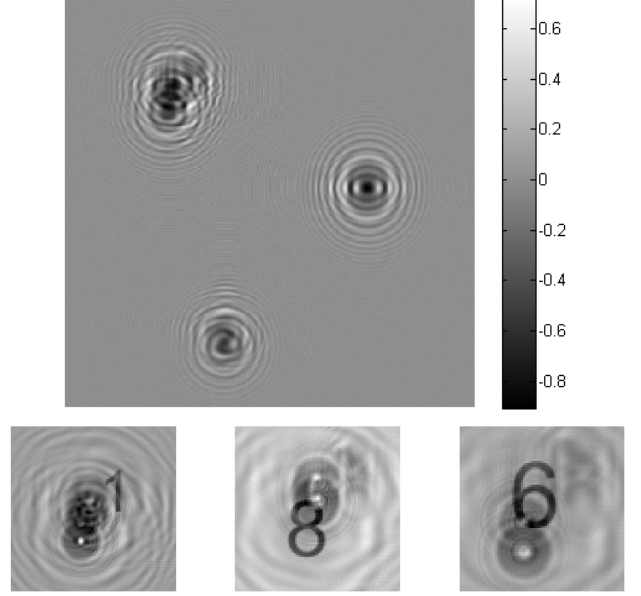[1]note that convolution kernels are symmetrical so $H_{z_i}^t = H_{z_i}$

Fig. 1. A simulated hologram of digits placed at different depth positions, and the classical reconstruction of the volume using back-projection operator for three of the digits.

border effects due to the absence of localization of convolution kernels h. Out-of-focus artifacts are visible in diffraction-based reconstructions shown in Fig. 1. Rather than performing object recognition in the coarse reconstruction $\hat{t}^{(\text{diffraction})}$, we suggest performing the detection/recognition directly on the hologram by matching diffraction-patterns.

Let M be a dictionary that collects all m object shapes that we are looking for. From this dictionary, a dictionary of diffraction-patterns P can be derived:

$$P = \begin{bmatrix} H_{z_1} \\ \vdots \\ H_{z_p} \end{bmatrix} M. \qquad (6)$$

A hologram of objects taken from the dictionary M is thus a superimposition of diffraction-patterns found in P shifted to the (x,y) location of each object. Let K be the matrix formed by collecting all possible (x,y) translates of diffraction patterns in P. A hologram of objects present in M can then be described as a weighted sum of diffraction-patterns:

$$d = CK\alpha + \epsilon, \qquad (7)$$

where C is a cropping operator that restricts the diffraction-patterns in K to the support of the n-pixels hologram d and α is an indicator vector with non-zero entries only when a given object is present at the corresponding (x,y,z) location. Vector α has dimension $m \times p \times n$.

Object detection and classification can be performed directly in the hologram by exploiting the sparsity of indicator vector $\boldsymbol{\alpha}$:

$$\hat{\boldsymbol{\alpha}} = \arg\min_{\boldsymbol{\alpha}} \|\boldsymbol{d} - \boldsymbol{C}\boldsymbol{K}\boldsymbol{\alpha}\|_2^2, \quad \text{s.t.} \quad \|\boldsymbol{\alpha}\|_0 \leq s, \quad (8)$$

with s the sparsity level, i.e., the number of objects in the hologram. Note that in contrast to equation (4), our aim is not to perform a reconstruction of the transmittance distribution $\boldsymbol{t}$ but to directly locate and recognize objects from $\boldsymbol{M}$ on the hologram.

Given the large dimension of $\boldsymbol{\alpha}$ and the very small number of objects s (only an order of s is known), problem (8) is best (approximately) solved using a greedy algorithm like matching pursuit. Looking for a single object at a time, the sub-problem to solve is of the form:

$$\arg\min_{i,\alpha} \|\boldsymbol{d} - \alpha \boldsymbol{C}\boldsymbol{k}_i\|_2^2, \quad \text{s.t.} \quad \alpha \geq 0, \quad (9)$$

with $\boldsymbol{k}_i$ the i-th column of matrix $\boldsymbol{K}$ and α a scalar. By introducing the 0-padded version of the data $\bar{\boldsymbol{d}} = \boldsymbol{C}^{\mathrm{t}}\boldsymbol{d}$ and the weighting matrix $\boldsymbol{W} = \boldsymbol{C}^{\mathrm{t}}\boldsymbol{C}$, sub-problem (9) is equivalent to a normalized maximum correlation search (i.e., matched filtering):

$$\arg\max_i \frac{\bar{\boldsymbol{d}}^{\mathrm{t}}\boldsymbol{W}\boldsymbol{k}_i}{\sqrt{\boldsymbol{k}_i^{\mathrm{t}}\boldsymbol{W}\boldsymbol{k}_i}}. \quad (10)$$

The maximum correlation over all (x,y) translations of a given diffraction pattern $\boldsymbol{p}_j$ can be efficiently computed using the fast Fourier transform [8]. Detection of several objects can then be performed by iteratively solving sub-problems (9) on the so-called residuals $\boldsymbol{r}$, i.e., part of the data unexplained by the currently detected diffraction-patterns. Residuals are initially equal to data $\boldsymbol{d}$. Each time a diffraction pattern $\boldsymbol{k}_i$ is found, the residuals are updated: $\boldsymbol{r} \leftarrow \boldsymbol{r} - \alpha \boldsymbol{C}\boldsymbol{k}_i$ as in matching pursuit algorithm and its variants.
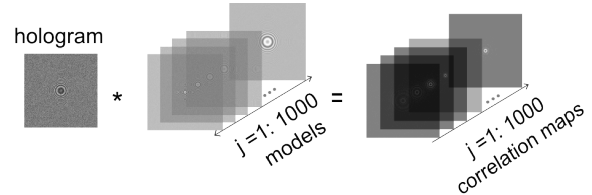
3. DIMENSIONALITY REDUCTION OF THE DICTIONARY OF DIFFRACTION PATTERNS

The huge dimensionality of the dictionary of all possible diffraction patterns $\boldsymbol{K}$ makes direct detection of objects in a hologram a very hard task. Efficient computation of correlations for all possible (x,y) translates considerably reduces the required computational effort, yet the number of diffraction-patterns to consider in the dictionary $\boldsymbol{P}$ is still very large: $n \times p$. It is therefore crucial to reduce the dimensionality of the dictionary $\boldsymbol{P}$ for general applicability of the diffraction-pattern matching method.
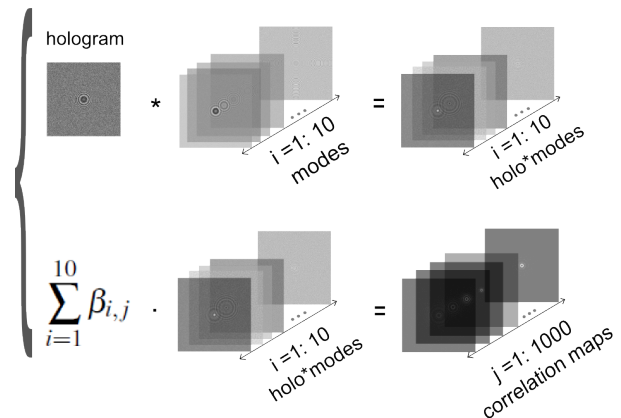
Rather than using the dictionary of geometrically-centered diffraction-patterns $\boldsymbol{P}$, we could consider a low-dimensional approximation:

$$\boldsymbol{P} \approx \sum_{i=1}^{k} \boldsymbol{u}_i \sigma_i \boldsymbol{v}_i^{\mathrm{t}}, \quad (11)$$

where $\boldsymbol{P}$ is approximated by the best rank-k matrix as obtained by the singular value decomposition (SVD) considering the singular vectors $\boldsymbol{u}_i$ and $\boldsymbol{v}_i$ associated with the k largest singular values $\{\sigma_1, \ldots, \sigma_k\}$. Within this approximation, diffraction pattern $\boldsymbol{p}_j$ is represented by the linear combination $\sum_i \beta_{i,j} \boldsymbol{u}_i$, where coefficient $\beta_{i,j}$ is equal to $\sigma_i \boldsymbol{v}_i(j)$. Vectors $\boldsymbol{u}_i$ represent the modes of the diffraction-patterns.

(a) Diffraction-pattern matching approach

(b) Diffraction-pattern matching with reduced dictionary

Fig. 2. Calculation of the X-Y criteria slices during the exhaustive search, (a) without and (b) with SVD. Operator * represents normalized weighted correlation. As shown in this figure, the reconstruction using the approximated dictionary is faster and more memory efficient than using the accurate dictionary.

Using this approximation, the computation of correlation terms in equation (10) becomes a linear combination of the correlation of each of the k modes with the data (see Fig. 2):

$$\bar{\boldsymbol{d}}^{\mathrm{t}}\boldsymbol{W}\boldsymbol{k}_j \approx \sum_{i=1}^{k} \beta_{i,j} \bar{\boldsymbol{d}}^{\mathrm{t}}\boldsymbol{W}\boldsymbol{u}_i. \quad (12)$$

In equation (12), expressions $\bar{\boldsymbol{d}}^{\mathrm{t}}\boldsymbol{W}\boldsymbol{u}_i$ do not depend on the considered diffraction-pattern $\boldsymbol{k}_j$ and can thus be computed once for all the diffraction-patterns. Using fast Fourier transforms, (x,y) correlation maps for each of the k modes can be pre-computed with a computational cost of order $k \times n \log(n)$. The correlation maps of the $m \times p$ diffraction-patterns can then be derived by performing $m \times p$ linear combinations of pre-computed correlation maps with a computational cost of

order $m \times p \times k \times n$. The denominator of (10) can either be precomputed, or can be approximated using SVD.

It shall be noted that the time complexity of the greedy algorithm presented in Sec. 2 is an order of $m\,p\,n\,log(n)$ with $m\,p >> k$. Bigger parameter ranges increase $m\,p$ but they do not affect k accordingly, which shows the significance of the time gain in the applications with big ranges for parameters like depth and Radii of the objects.
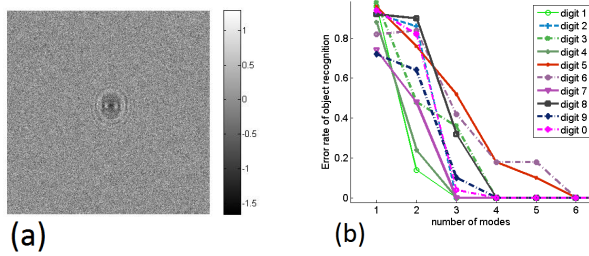
4. EXPERIMENTAL RESULTS

(a)

(b)

Fig. 3. The study of error rate on object recognition for the toy problem of digits' holograms (see Sec. 4.1). (a) contains one of the holograms of Sec. 4.1 for object 0 placed at 0.17m from the sensor, (b) shows the error rates representing the discriminating power of approximated dictionaries for a fixed level of noise with SNR ≈ 3.

In this section we report the results of our method on both simulated and real holograms. The goal of this section is to show how the low rank approximation of the dictionary affects: (i) the error rate of object recognition, (ii) the estimation accuracy, (iii) the computational costs. The implemented method uses FFTW library and OpenMP to exploit multithreading on a six-core CPU for the calculations of the forward and backward Fourier transforms and the models. GPU acceleration can be used to decrease the time costs further more. PROPACK [14] is used to calculate truncated SVD of the dictionary using an iterative Lanczos method.

4.1. Illustration: recognition of digits in a hologram

The first Monte Carlo study was performed to determine the pattern discrimination power of the greedy method as a function of the rank of the approximated dictionary. Considering holograms containing white Gaussian noise with SNR a ≈ 3 (see Fig. 3-a), Fig. 3-b shows the number of modes required for every digit to be classified and located. As shown in this figure, all digits are classified and localized using only first 6 modes (i.e., error rate is 0). In this study, 50 holograms were simulated for every digit. The setup parameters were chosen to simulate the holograms captured by a 400×400-pixel camera with pixel size of $20\,\mu m$ and fill-factor of 0.7. The laser

wavelength was set to $0.532\,\mu m$. The depth position of digits was chosen to be in the range of $[0.15\ \ 0.2]$ meters.

It should be noted that when diffraction-patterns overlap, diffraction-pattern matching is more difficult and more modes are required to perform object identification. Therefore for the hologram of Fig. 1 16 modes were necessary for correct object recognition. The time gain for this hologram was 3, regarding the relatively small dictionary.

4.2. Study of bias and standard deviation of estimated parameters
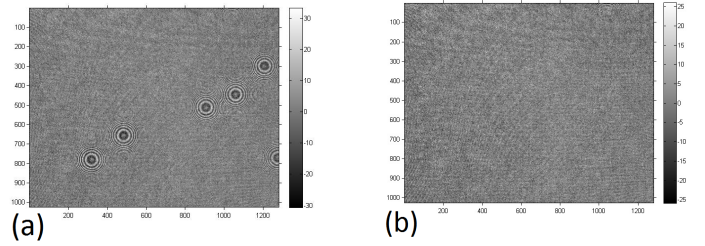

(a)

(b)

Fig. 4. (a): The experimental hologram of Sec. 4.3 captured in the department of fluid mechanics and acoustics of Lyon (LMFA), (b): same experimental hologram cleaned from the in-field particles using the 5 first modes. The residual's magnitude is high due to the signature of the out-of-field particles placed close to the borders.

Simulations were performed for a set of Monte Carlo studies to check the accuracy of parameter estimation for different ranks of the approximated dictionary of opaque spherical particles (i.e., droplets or bubbles). Comparison between the bias and standard deviation of the results with the theoretical Cramer-Rao lower bounds(CRLBs) [15] showed that the first 50 modes are enough to assure that the greedy algorithm results in minimum variance unbiased parameter estimators. In this study the camera was considered to have 512×512 pixels with the pixel size of $7\,\mu m$ and fill-factor of 0.7. The laser wavelength was set to $0.532\,\mu m$. For the calculation of SVD, the range of parameters Z was set to $[0.1\ \ 0.3]$ m and the range of radii was set to $[60\ \ 100]\,\mu m$. To check the bias and standard deviation of the estimated parameters, 50 holograms were simulated for a particle placed at the center of the hologram and 50 holograms for a particle at the right bottom corner of the sensor. The SNR was set to approximately 19.

4.3. Experimental Holograms of spherical particles

Another set of experiments involved object detection and parameter estimation from a video of captured holograms (see Fig. 4-a). These holograms were captured of injected water droplets at the department of fluid mechanics and acoustics of Lyon (LMFA). The goal was to detect all spherical particles

and estimate their 3D position and radius. The dictionary contained 410 diffraction patterns and our experiments showed that a rank 5 dictionary was accurate enough to detect and locate all the particles in the field of view of camera and clean them successfully. The time gain using the approximated dictionary (46 s) over the normal approach (342 s) was a factor of 8. A sample cleaned hologram is shown in Fig. 4-b. The droplets were generated by a piezoelectric jetting device manufactured by MicroFab Technologies. This injector produces mono-dispersed droplets with radii of $31\,\mu\text{m} \pm 0.5\,\mu\text{m}$. The droplets were produced to be at distances ranging from $30\,\text{cm}$ to $48\,\text{cm}$ of a 1024×1280 pixel camera with pixel size of $21.7\,\mu\text{m}$ and fill-factor of 0.84. This imaging setup had a magnification of 1.42 which resulted in the holograms with an SNR of ≈ 16.

5. CONCLUSION

Signal processing approaches can be used in digital holography for object recognition and volume reconstruction directly from the hologram. These approaches exploit the diffraction pattern dependencies on shape and position of the objects to perform direct diffraction-pattern matching on the hologram. In this paper we showed that reducing the rank of the diffraction-pattern dictionary (using a dimensionality reduction method) can speedup the digital hologram reconstruction. The calculation of the algorithm complexity shows that the time gains of such methods depend on the application parameters, where bigger search spaces increase the time gain.

Our Monte Carlo studies on the application of spherical objects have shown that the parameter estimation using a low rank approximation of the dictionary is unbiased with a standard deviation comparable to the Cramer-Rao lower bounds. We have also studied the discrimination power of the pattern matching algorithm in terms of the approximation rank for non-parametric objects, and we have shown that low rank approximations give accurate object recognition and location extraction. In the last study, we achieved a speedup rate of 8 on a video of experimental holograms of water droplets.

6. REFERENCES

[1] J. Katz and J. Sheng, "Applications of holography in fluid mechanics and particle dynamics," *Annual Review of Fluid Mechanics*, vol. 42, pp. 531–555, Jan. 2010.

[2] Yong-Seok Choi and Sang-Joon Lee, "Holographic analysis of three-dimensional inertial migration of spherical particles in micro-scale pipe flow," *Microfluidics and Nanofluidics*, vol. 9, pp. 819–829, Mar. 2010.

[3] E. Malkiel, J. Sheng, J. Katz, and J. R. Strickler, "The three-dimensional flow field generated by a feeding calanoid copepod measured using digital holography," *J Exp Biol*, vol. 206, pp. 3657 –3666, Oct. 2003.

[4] S. L. Pu, D. Allano, B. Patte-Rouland, M. Malek, D. Lebrun, and K. F. Cen, "Particle field characterization by digital in-line holography: 3D location and sizing," *Experiments in Fluids*, vol. 39, pp. 1–9, June 2005.

[5] F. Dubois, C. Schockaert, N. Callens, and C. Yourassowsky, "Focus plane detection criteria in digital holography microscopy by amplitude analysis," *Optics Express*, vol. 14, pp. 5895–5908, June 2006.

[6] I. Moon, M. Daneshpanah, B. Javidi, and A. Stern, "Automated three-dimensional identification and tracking of Micro/Nanobiological organisms by computational holographic microscopy," *Proceedings of the IEEE*, vol. 97, no. 6, pp. 990 –1010, June 2009.

[7] W. Xu, M. H. Jericho, I. A. Meinertzhagen, and H. J. Kreuzer, "Digital in-line holography for biological applications," *PNAS*, vol. 98, no. 20, pp. 11301–11305, Sept. 2001.

[8] F. Soulez, L. Denis, E. Thiébaut, C. Fournier, and Ch. Goepfert, "Inverse problem approach in particle digital holography: out-of-field particle detection made possible," *JOSA. A*, vol. 24, pp. 3708–3716, Dec. 2007.

[9] DJ Brady, K. Choi, DL Marks, R. Horisaki, S. Lim, et al., "Compressive holography.," *Optics express*, vol. 17, no. 15, pp. 13040, 2009.

[10] L. Denis, D. Lorenz, E. Thiébaut, C. Fournier, and D. Trede, "Inline hologram reconstruction with sparsity constraints.," *Optics Letters*, vol. 34, pp. 3475–3477, 2009.

[11] S. Sotthivirat and J. A Fessler, "Penalized-likelihood image reconstruction for digital holography," *JOSA A*, vol. 21, no. 5, pp. 737–750, 2004.

[12] S.H. Lee, Y. Roichman, G.R. Yi, Sh.H. Kim, S.M. Yang, A. Van Blaaderen, P. Van Oostrum, and D.G. Grier, "Characterizing and tracking single colloidal particles with video holographic microscopy," *Optics Express*, vol. 15, no. 26, pp. 18275–18282, 2007.

[13] C. P. Allier, G. Hiernard, V. Poher, and J. M. Dinten, "Bacteria detection with thin wetting film lensless imaging," *Biomedical optics express*, vol. 1, no. 3, pp. 762–770, 2010.

[14] "PROPACK," http://soi.stanford.edu/ rmunk/PROPACK/, [Online; accessed 23-November-2012].

[15] C. Fournier, L. Denis, and Th. Fournel, "On the single point resolution of on-axis digital holography," *JOSA. A*, vol. 27, pp. 1856–1862, Aug. 2010.