Fast and accurate 3D object recognition directly from digital holograms

Mozhdeh Seifi, Loic Denis*, Corinne Fournier

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

.

Corresponding author : loic.denis@univ-st-etienne.fr

Pattern recognition methods can be used in the context of digital holography to perform the task of object detection, classification and position extraction directly from the hologram rather than from the reconstructed optical field. These approaches may exploit the differences between the holographic signatures of objects coming from distinct object classes and/or different depth positions. Direct matching of diffraction patterns however becomes computationally intractable with increasing variability of objects due to the very high dimensionality of the dictionary of all reference diffraction patterns. We show that most of the diffraction pattern variability can be captured in a lower dimensional space. Good performance for object recognition and localization is demonstrated at a reduced computational cost using a low dimensional dictionary. The principle of the method is illustrated on a digit recognition problem and on a video of experimental holograms of particles.
© 2013 Optical Society of America

OCIS codes: 090.1995,100.3010,100.3190,100.5010

1. Introduction

In-line digital holography is a 3D imaging method which is being increasingly used in applications requiring to track micro-objects from fluid mechanics [1–3] to biology [4–7]. The advantage of this lensless technique compared to its alternatives is the simple imaging setup with high video frame rate resulting in accurate reconstructions of the volume of interest. During the recording step in this method, the objects are illuminated by a laser beam and the diffracted wave intensity is recorded in a video by a camera. The recorded sequence of holograms is then processed to reconstruct the volume's evolution over time.

Among the existing hologram processing techniques, model fitting methods have shown to be the most efficient and accurate for metrological applications [8–11]. The contribution of this paper is to show that it is possible to go beyond model fitting by performing diffraction-pattern matching to identify the class and 3D location of objects directly from holograms. Our approach is inspired by the method for 3D object recognition based on hologram correlation suggested in [12, 13]. However, rather than matching two experimental holograms corresponding to the two objects to be compared [14], we consider matching a hologram with a whole dictionary of diffraction patterns generated for various objects and all their possible 3D locations. As observed for the task of accurate localization of particles in holographic metrology, very good performance can be reached by searching for elements of a dictionary of diffraction patterns that best match the data. This approach contrasts with conventional hologram processing methods that analyze a coarse reconstruction obtained by hologram back-propagation using Fresnel transform or other optical propagation approximations.

In-line holography of opaque objects is restricted to mostly empty volumes (typically, the projection of all objects should cover less than 1% of the sensor, see [15]). The hologram can then be well approximated as the incoherent superimposition of the diffraction pattern of each object. Hologram reconstruction is an inverse problem that can also be seen as a linear un-mixing problem, or as a multiple-object pattern-recognition problem with overlapping patterns. In the context of particle hologram analysis, a greedy algorithm has been proposed [10, 16]. This algorithm builds incrementally a model of the hologram by fitting at each step the parametric diffraction-pattern model of a single (spherical) particle. When considering a large diversity of particle diameters and distances to the hologram, the collection of all diffraction patterns to be considered becomes huge and the computational cost rises for reconstruction of hologram videos. The extension to more general classes of objects is restricted by the explosion of the size of the dictionary of all diffraction-patterns that must be matched with the hologram.

To overcome this computational issue, we propose to approximate the dictionary in a low-dimensional space. We show that such approximation preserves enough variability between diffraction patterns to perform correct discrimination and accurate 3D localization. The dimensionality reduction is based on a singular value decomposition (SVD) of the dictionary to identify most significant modes. It leads to an efficient diffraction-pattern matching algorithm.

The structure of this paper is as follows: the pattern recognition problem is introduced in the next section. Low-dimensional approximation of the dictionary using SVD is described in section 3. We report numerical and experimental results in section 4 and we conclude the paper in Sec. 5.

2. Object recognition from a digital hologram

The majority of hologram reconstruction methods reconstruct the optical wave-field using light back-propagation techniques. After reconstruction of the wave-field at different depths, objects can be segmented and recognized in their maximum-of-focus plane. This approach is illustrated in Fig. 1, on a simulated hologram of digits. Since either the in-focus or out-of-focus signature of a given object is visible on all reconstructed planes, selection of the in-focus plane and separation of overlapping in-focus and out-of-focus images from different objects can be hard. Better reconstructions can be obtained by stating hologram reconstruction as an inverse problem. Regularized reconstruction using an oscillation-preventing term (total variation penalty) [17–22] or a sparsity constraint (ℓ_1 norm minimization) [23] have been shown to greatly enhance the quality of the reconstructed volumes. Compared to performing object recognition on the resulting volumes, it is more robust to directly recognize diffraction patterns in the hologram. In a Bayesian perspective, this can be understood as making use of a much stronger prior when solving the inverse problem of hologram reconstruction, since we specify which objects are expected to be present in the volume. Under the approximation of a hologram as a linear super-imposition of the diffraction pattern of each object, the hologram formation model (direct problem) can be formulated as

$$\mathbf{d} = \mathbf{M}\boldsymbol{\alpha} + \boldsymbol{\epsilon}\,, \tag{1}$$

where $\mathbf{d}$ is the vector of data, i.e., the N pixel values of the hologram re-ordered as a vector in $\mathbb{R}^N$. In this formulation, $\mathbf{M}$ represents a dictionary with columns containing all possible diffraction patterns of the family of objects under consideration. Two diffraction patterns may differ depending on the class / shape and 3D location of the objects (e.g., for opaque and spherical particles, diffraction patterns depend on the radius r and 3D location of the particle (x, y, z)). A given hologram $\mathbf{d}$ may be described as the weighted sum of few diffraction patterns from the dictionary $\mathbf{M}$. Let $\mathbf{m}_{i_k}$ be the diffraction pattern created by the k^{th} object ($\mathbf{m}_{i_k}$ is column i_k of dictionary $\mathbf{M}$). The corresponding component i_k of vector $\boldsymbol{\alpha}$ is then non-zero and defines the amplitude of that diffraction pattern on the hologram. Vector $\boldsymbol{\alpha}$ has mostly zero components except for the amplitude of the diffraction pattern created by each object (i.e., $\boldsymbol{\alpha}$ is a very sparse vector). Vector $\boldsymbol{\epsilon}$ accounts for modeling errors and noise in the data. In the absence of specific information on the noise statistics, it will be considered white and Gaussian.

Recognition of objects directly from the hologram $\mathbf{d}$ amounts to determining the support of vector $\boldsymbol{\alpha}$ (i.e., location of non-zero entries). Numerous methods have been developed in the literature of compressed sensing and sparse reconstruction to solve such problem. Given the very limited number of objects present in the volume compared to the size of the dictionary $\mathbf{M}$ (many different shapes and 3D locations can be considered for a given object), vector $\boldsymbol{\alpha}$ is

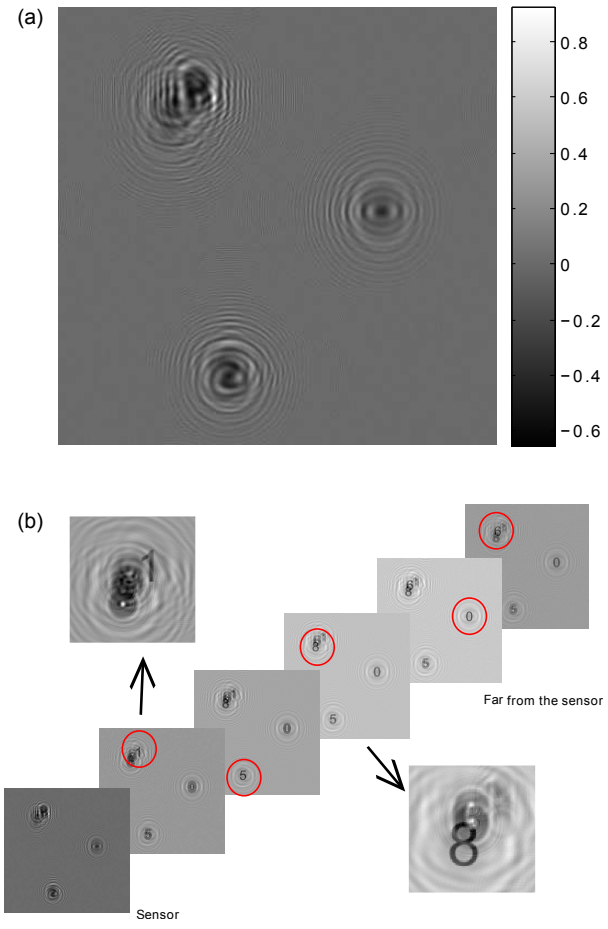


Fig. 1. Illustration of a digit recognition task from a digital hologram: (a) a simulation of a 400×400-pixels hologram of digits located at different distances (pixel size: 20 µm, fill-factor: 0.7, laser wavelength: 0.532 µm, depth range: $[0.55 \;\; 0.85]$ m); (b) Fresnel reconstruction of the volume displaying both in-focus and out-of-focus images of the digits which makes digit recognition from the reconstructed volume a difficult task. Direct recognition of the diffraction patterns using the approach described in the paper avoids dealing with these numerous image artifacts.

extremely sparse. Greedy algorithms such as matching pursuit and its variants are well suited to this problem [24]. Object identification and detection can then be performed directly from the hologram by iterating the following steps, as proposed in the case of particle holograms in [10, 16]:

1. An exhaustive search for the diffraction pattern in the discrete dictionary $\mathbf{M}$ that best improves current modeling of the data. This step corresponds to the identification of the column j of dictionary $\mathbf{M}$ that is the most correlated with residuals $\mathbf{r}$ representing the part of data that is yet to be modeled:

$$\arg\max_{j} \quad \mathbf{m}_j^{\mathrm{t}} \mathbf{W} \mathbf{r} \,, \tag{2}$$

   where $\mathbf{W}$ is a weighting matrix, often referred to as the precision matrix, that can account for noise correlation ($\mathbf{W}$ is then the inverse of noise covariance matrix), non-stationary noise variance (with a diagonal $\mathbf{W}$ matrix and non-constant diagonal values) or unmeasured areas (with zero values on the corresponding diagonal entries). In Eq. 2, the dictionary is supposed to be properly normalized ($\forall i,\ \mathbf{m}_i^{\mathrm{t}} \mathbf{W} \mathbf{m}_i = 1$). Residuals $\mathbf{r}$ are initially equal to the data $\mathbf{d}$ and iteratively updated in step 3 described bellow.

2. A local optimization step can possibly be performed to refine shape and/or location with sub-pixel/sub-sampling accuracy. This can be done by non-linear least-squares fit when a continuous model of the diffraction patterns is available (e.g., for spherical particles).

3. A cleaning step that updates the residuals by removing the identified diffraction pattern $\mathbf{m}_j$. This step plays an important role as it prevents from detecting the same object several times (its contribution being removed from the residuals once and for all) and it improves the contrast of remaining diffraction patterns.

The first step requires to perform an exhaustive search on a very large dictionary which limits the applicability of the method to strongly constrained families of objects such as spherical particles. The strong hypothesis is that the object to be identified has a representative in the dictionary. The extension to more general classes of objects requires a way to speed up this search. We describe in the next section how to approximate the dictionary in a lower dimensional space and derive a fast method to perform this diffraction-pattern matching step.

3. Low-dimensional approximation of the dictionary of diffraction patterns

To reduce the inter-pattern redundancy of the dictionary, we propose to approximate it in a low-dimensional space. We first distinguish variations due to the transversal location of the

object (i.e., its (x, y) location) from other variations. Let $\mathbf{C}$ be the restriction of dictionary $\mathbf{M}$ to diffraction patterns that are centered on the optical axis (i.e., only columns corresponding to patterns that are geometrically centered are kept). $\mathbf{C}$ captures all variations other than transversal shifts. Since diffraction patterns display some correlations, matrix $\mathbf{C}$ can be well approximated by a lower rank matrix obtained by singular value decomposition (SVD):

$$\mathbf{C} \approx \sum_{i=1}^{k} \mathbf{u}_i \mathrm{s}_i \mathbf{v}_i^t \,, \tag{3}$$

where $\mathbf{u}_i$ and $\mathbf{v}_i^t$ represent the i^{th} left and right singular vectors associated with singular value s_i from the SVD of matrix $\mathbf{C}$, and k is the rank chosen for the approximation. In this approximation, the j^{th} diffraction pattern is given by:

$$\mathbf{c}_j \approx \sum_{i=1}^{k} \mathbf{u}_i \, \beta_{i,j} \,, \text{ with } \beta_{i,j} = \mathrm{s}_i \cdot \mathrm{v}_{i,j}. \tag{4}$$

The orthonormal basis formed by $\{\mathbf{u}_i\}_{1..k}$ that spans the k-dimensional approximation subspace is called the dictionary of diffraction-pattern modes. Any diffraction pattern $\mathbf{c}_j$ is approximated as a linear combination of modes $\mathbf{u}_i$, with weights $\beta_{i,j}$. Since the array of weights $\boldsymbol{\beta}$ depends only on the dictionary $\mathbf{C}$, it can be computed off-line once for all.

Correlations between the dictionary and residuals must be evaluated in order to identify the best matching pattern in a hologram. Let κ_i denote the correlation between mode $\mathbf{u}_i$ and the residuals. Using our low-rank approximation, correlations in Eq. 2 can be expressed as a linear combination of few mode/residuals correlations κ_i:

$$\mathbf{c}_j^{\mathrm{t}} \mathbf{W} \mathbf{r} \approx \sum_{i=1}^{k} \beta_{i,j} \, \kappa_i \,, \text{ with } \kappa_i = \mathbf{u}_i^{\mathrm{t}} \mathbf{W} \mathbf{r} \,. \tag{5}$$

As mentioned earlier, for a given dictionary $\mathbf{C}$ it is necessary to compute coefficients $\beta_{i,j}$ only once. The computation of k mode/residual correlations κ_i then suffices to approximate any correlation between a dictionary element $\mathbf{c}_j$ and the residuals, resulting in a large speedup compared to direct computation of the correlation $\mathbf{c}_j^{\mathrm{t}} \mathbf{W} \mathbf{r}$ as long as the number of diffraction patterns in the dictionary $\mathbf{C}$ is much larger than the number of retained modes k.

Identification of the best-matching diffraction pattern in Eq. 2 not only involves geometrically-centered patterns $\mathbf{c}_j$ but also all possible (x, y) shifts. Correlation with all shifted versions of a given geometrically centered pattern can be efficiently computed using fast Fourier transforms (FFT) with zero-padding to prevent from periodization artifacts. Two shifted versions $\mathbf{m}_1$ and $\mathbf{m}_2$ of a pattern $\mathbf{c}$ generally differ in norm ($\mathbf{m}_1^{\mathrm{t}} \mathbf{W} \mathbf{m}_1 \neq \mathbf{m}_2^{\mathrm{t}} \mathbf{W} \mathbf{m}_2$), so that a proper normalization must be computed. The pattern matching step given in Eq. 2 is then replaced by:

$$\arg\max_j \quad \frac{\mathbf{m}_j^{\mathrm{t}} \mathbf{W} \mathbf{r}}{\mathbf{m}_j^{\mathrm{t}} \mathbf{W} \mathbf{m}_j} \,, \tag{6}$$

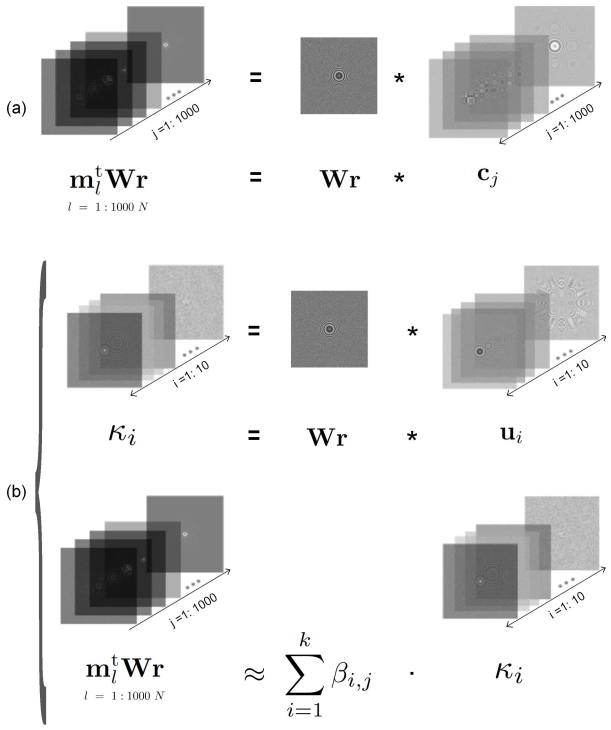


Fig. 2. Identification of the best-matching pattern in the dictionary: (a) by correlation with each element of the dictionary, straightforward but costly; or (b) by linear combination of the correlation with few diffraction-pattern modes, much faster. Computation of correlation maps is denoted using the symbol '$*$'.

where normalization of the dictionary $\mathbf{M}$ is no longer required.

Fast computation of the numerator in Eq. 6 for all integer pixel shifts of a pattern $\mathbf{c}$ can be performed with FFT by computing correlations between $\mathbf{c}$ and $\mathbf{Wr}$ in Fourier domain. The normalization term corresponding to the denominator depends only on the dictionary and can either be computed off-line or be approximated using another SVD. When weighting matrix $\mathbf{W}$ is diagonal ($\mathbf{W} = \mathrm{diag}(\mathbf{w})$), the denominator can be computed for all integer shifts by computing with FFT the correlations between $\mathbf{w}$ and $\mathrm{diag}(\mathbf{m})\mathbf{m}$, as done in [16].

Using the approximation derived in Eq. 5, the correlation between any diffraction-pattern from the dictionary $\mathbf{M}$ and the residuals $\mathbf{r}$ can be approximated as a linear combination of mode/residuals correlations. It is necessary to compute those mode/residuals correlation maps for all integer shifts only once for each mode. Each of these k maps can benefit from the fast FFT-based implementation described.

Fig. 2 compares the computationally expansive matching of diffraction patterns taken from

the whole dictionary $\mathbf{M}$ in (a) with the use of the low-dimensional approximation based on diffraction-pattern modes in (b). For a dictionary made of n different geometrically-centered patterns (i.e., $\mathbf{C}$ has n columns), computation of the correlations for all integer shifts has a complexity of the order of $n \times N \log(N)$ whereas using the proposed approximation reduces the complexity to $k \times N \log(N)$ where the number of modes k can be generally chosen much smaller than n, as illustrated next section.

Note that the size of matrix $\mathbf{C}$ can be very large. Rather than computing its full SVD, it is preferable to use (approximate) iterative methods to compute a truncated SVD. We used a method [25] based on Lanczos bi-diagonalization [26] of $\mathbf{C}$ to simplify the calculation of the SVD of a large sparse matrix.

4. Experimental results

In this section we report the results of our method on both simulated and real holograms. The implemented method uses FFTW [27] library and OpenMP [28] to exploit multi-threading on a six-core Xeon CPU for the calculation of the forward and backward Fourier transforms and the models. PROPACK [29] is used to calculate the truncated SVD of the dictionary employing an iterative Lanczos method.

4.A. Object discrimination and location: illustration with holograms of digits

We illustrate, using simulations, the capability of the proposed method to correctly identify and locate several digits by matching diffraction patterns directly on digit holograms. We aim to find the minimum rank of dictionary resulting in correct recognition and location of all digits.

First, we illustrate our reconstruction results of the hologram of Fig. 1-a which contains overlapping and non-overlapping digit patterns. Inspired by the decay rate of the singular values of the dictionary (see the green curve in Fig. 3), we verified that a rank 5 approximation of the dictionary results in correct object recognition and depth estimation for non-overlapping objects (i.e., the digits 5 and 0 in Fig. 1-b were recognized and located within pixel accuracy). However, the overlapping patterns (i.e., patterns of objects 1,8 and 6 in Fig. 1-b) are not completely recognizable using only the first 5 modes. Our experiment showed that at least 16 modes are necessary to obtain an approximation of the dictionary that provides enough information for all the digits to be recognized and located in the 3D volume. In this case, given the relatively small size of the dictionary, a time gain of a factor 3 was observed compared to matching the full dictionary. The setup parameters are chosen to simulate a hologram captured by a 400×400-pixel camera with the pixel size of 20 μm and the fill-factor of 0.7. the laser wavelength is set to 0.532 μm. The depth position of digits is chosen to be in the range of $[0.55 \;\; 0.85]$ m. The dictionary contained 300 patterns, 30

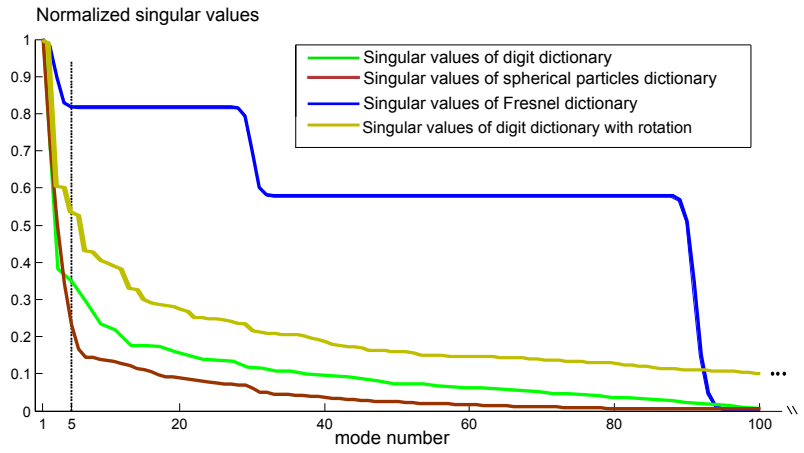
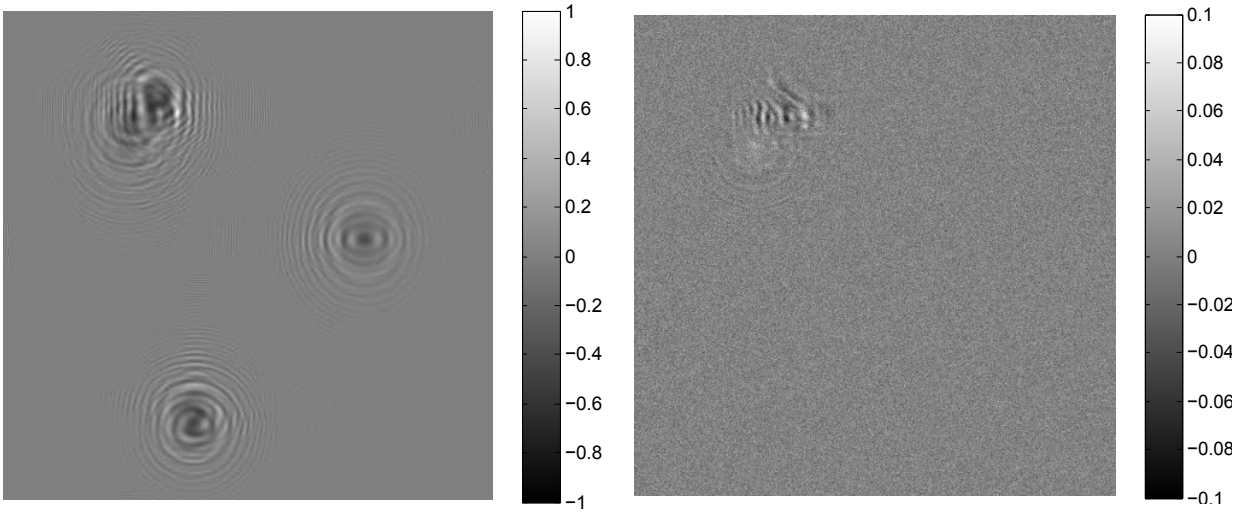


Fig. 3. First 100 singular values of the dictionary in descending order. The dictionary of digits (in green) contains 300 centered patterns for digits placed at different depth positions with the application parameters in Sec. 4.A. The dictionary of spherical particles (in red) consists of the diffraction models for spherical objects changing depth and radius with the application parameters in Sec. 4.B. The Fresnel dictionary (in blue) corresponds to Fresnel functions with changing depth values. The dictionary of rotated digits (in mustard) contains 900 centered patterns (10 (digits)$\times$90 (angle samples)). In the case of spherical objects, 5 first modes can approximate the dictionary accurately. The vanishing rate of the singular values of the Fresnel dictionary is very low which implies that it is not possible to deduce an accurate low rank approximation of Fresnel dictionary.

patterns for different depth values per digit class. The dimensions of this relatively small dictionary is therefore 160000×300.

Fig. 4-a shows the hologram of digits and Fig. 4-b shows the residuals after cleaning the signature of the objects. The model that is used to clean the patterns contains the second order term of the object signature. However, the inter-pattern second order terms (i.e., the interferences between the object diffractions) are not known, since all the objects are not yet detected when the signature of one object is being removed (these interferences belong to the signatures of 1,8 and 6). These interferences (with very low intensities) are not cleaned from the residuals and are visible in Fig. 4-b.

Next, we study the discrimination power of the method as a function of the noise level and the approximation rank. The goal is to (i) show the effect of noise magnitude on the required rank of dictionary for correct object recognition and location, and (ii) to estimate



(a) (b)

Fig. 4. Digit recognition from a digital hologram: (a) the hologram of digits (b) the residuals after cleaning all the signatures from the hologram.

the minimum required rank for successful discrimination of every digit from the others. The time costs calculated in these experiments are finally used to evaluate an order of magnitude for the time gain of the proposed algorithm in this application.

Our simulations contain a set of 10000 holograms of digits. 10 noise level are selected and for each noise level and for each digit 100 holograms are simulated. Changing the rank of the approximation, object recognition is performed to calculate the rate of error of classification. One hologram and the results for one digit (a "2") are shown in Fig. 5. As shown in Fig. 5-(a), a rank 4 approximation of dictionary results in correct recognition of "2" for every level of noise. The object is placed at 0.17 m and one of the holograms with noise of $\sigma = 0.01$ is shown inFig. 5-(b). The setup parameters are the same as the previous study except for the depth range which is changed to [0.15 0.2] m.

Using the same configuration, the second goal is to illustrate the discrimination error of every digit as a function of approximation rank. Fig. 5-(c) shows the error rate of object recognition for all 10 digits for the noise with the standard deviation of 0.31. This plot shows that a rank 6 approximation of the dictionary can correctly classify all digits without error. This plot can be used in conjunction with the plot of singular values (see the green curve in Fig. 3) to choose the approximation rank of dictionary for an assured object recognition task. One of the holograms containing a '2' is shown in Fig. 5-(d).

Finally, we performed a primitive experiment to test the robustness of object identifica-

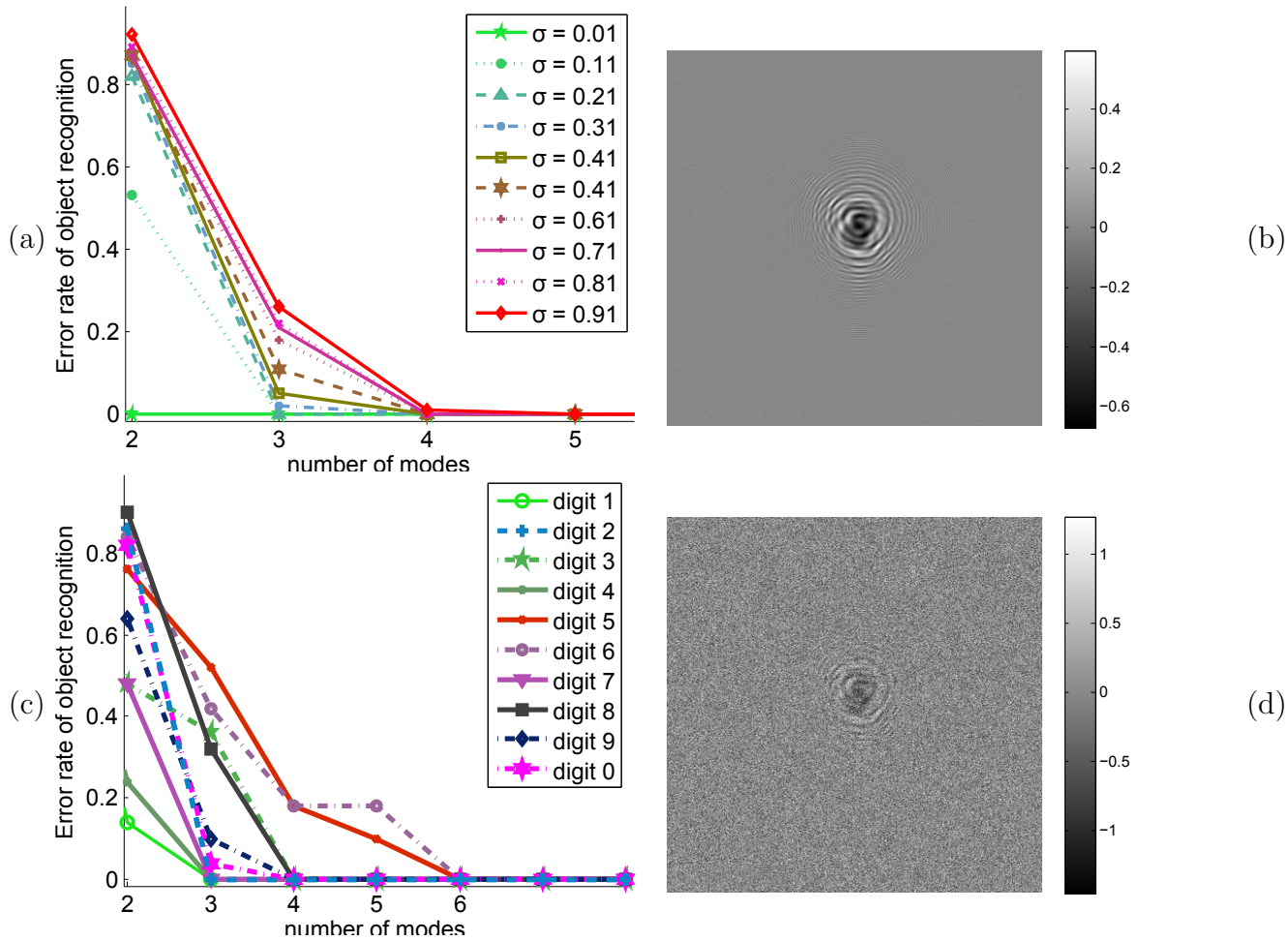


Fig. 5. Study of the error rate on object recognition for the toy problem of holograms of digits (see Sec. 4.A). (a) shows the error rates of object recognition for a '2' placed at 0.17m from the sensor. The curves show the error rate for different approximation ranks and different noise levels, (b) illustrates one hologram of 2 degraded by white Gaussian noise with $\sigma = 0.01$, (c) plots the error rates representing the discriminating power of approximated dictionaries for a fixed level of noise resulting in a SNR ≈ 3, (d) illustrates one of the holograms of Sec. 4.A for object '2' placed at 0.17m from the sensor.

tion to 2D orientation changes of the objects. Our dictionary contained 90 orientations (i.e., sample step size of 4 degrees) and 10 digits. The resulted dictionary contained 900 diffraction patterns. Truncated SVD method was then employed to obtain the closest rank 500 approximation of this dictionary. The first 100 singular values of this dictionary are plotted in Fig. 3. For the Monte Carlo study, 100 holograms of a digit "2" with an arbitrary orientation angle were simulated and degraded by white and Gaussian noise with standard deviation 0.3. Our results showed that a rank 12 approximation of the dictionary was accurate enough to estimate the orientation angle and recognize the object accurately. Fig. 6-a shows the object 2 that is rotated counter-clock wise by 210.5 degrees. This rotation value falls between two dictionary samples of 208 and 212 degrees. Fig. 6-b shows one hologram of such an object placed at depth of 0.75 m. This hologram is degraded by white and Gaussian noise with the standard deviation of 0.31. Fig. 6-c shows the error rates of the object identification and orientation estimation for this object. A full study of robustness to rotation and scale variations is left as future work.

4.B. Simulations of spherical particle holograms

In this section, we show the results of our proposed method on the reconstruction of spherical object holograms. Monte Carlo simulations are performed to illustrate the accuracy of object sizing and location, and the time gain. This experiment is performed to check the biases and standard deviations of the estimated radii and 3D position using the low rank approximation of the dictionary. The results show standard deviations of the same order as the Cramer-Rao lower bounds (CRLBs) on the estimation of the radii and 3D position (see [30] for details on the calculation of CRLBs) and negligible bias using the first 50 modes, which validates the use of an approximated dictionary in the exhaustive search step of the greedy algorithm (see Sec. 2).

The details of this experiment is as followed: 50 holograms are simulated for a particle placed at the center of the hologram and another set of 50 holograms are simulated for a particle at the bottom right corner of the sensor, with the depth distance of 0.2 m and radii of 80 μm. The SNR of the signal is set to approximately 19. The camera is considered to have 512×512 pixels with the pixel size of 7 μm and fill-factor of 0.7. The laser wavelength is 0.532 μm.

To estimate the bias and standard deviation of the proposed method, we compared our results with the reference location and size obtained using a huge dictionary with very fine sampling. To build that reference dictionary, we sampled the parameters with step sizes equal to the CRLBs (which are an order of 15 μm for z and 50 nm for radii), obtaining the dictionary $\bar{\boldsymbol{C}}_1$ corresponding to 616 (z,r) pairs. The search space with this dictionary is limited in our experiments to $[0.1999 \;\; 0.2001]$ m for z and $[79 \;\; 81]$ μm for the radii to avoid

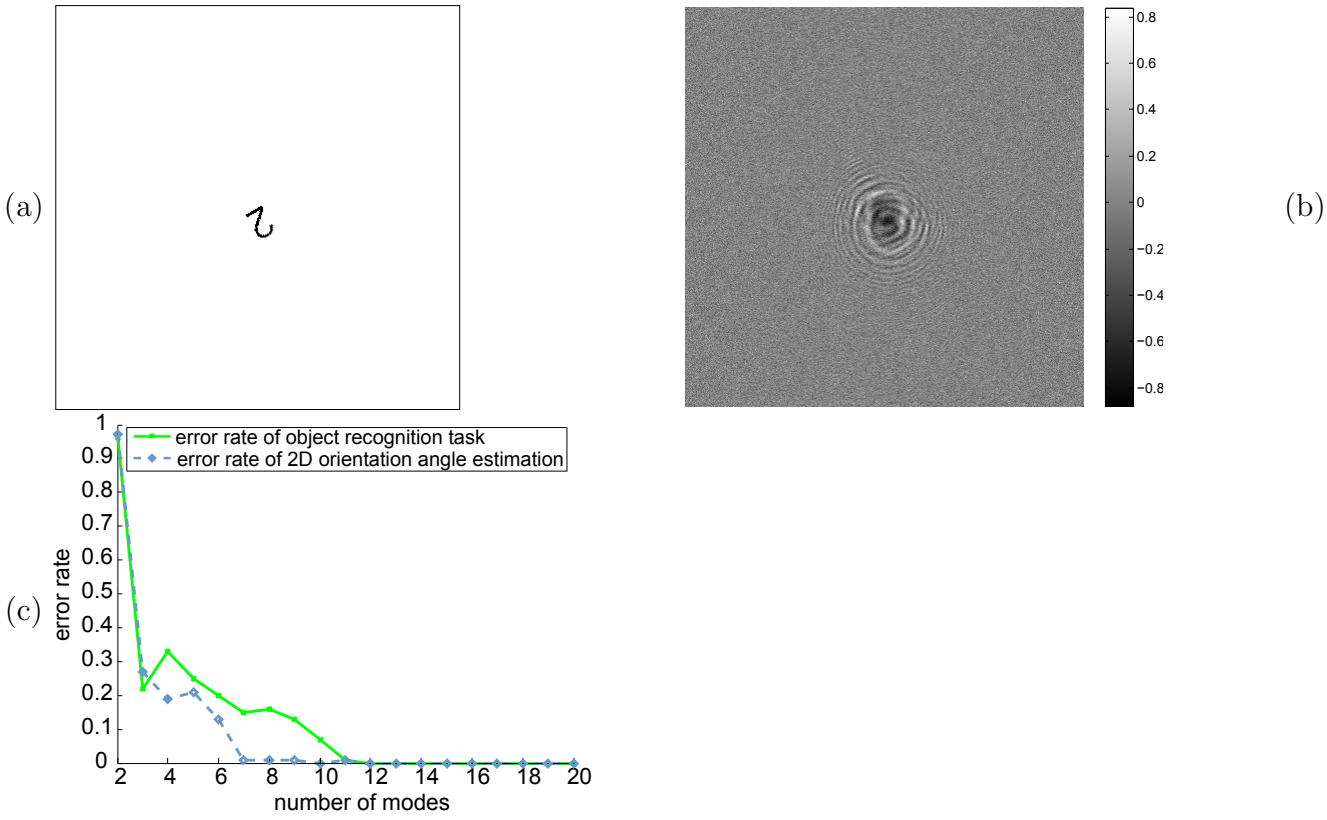


Fig. 6. Study of the robustness of our method to the orientation factor (see Sec. 4.A). (a) shows an object '2' rotated counter-clock wise on the object plane by 210.5 degrees, (b) illustrates one of the simulated holograms for object '2' placed at depth of 0.75m. This hologram is degraded by white and Gaussian noise with $\sigma = 0.31$, (c) shows the error rates of object recognition and orientation estimation for this object.

having a huge number of samples.

In calculating the mode dictionary for one setup, wide ranges of parameters can be considered. Such mode dictionary can be used in different experiments . In our experiment, the mode dictionary is calculated for the range of parameters z as $[0.1 \;\; 0.3]$ m and the range of radii as $[60 \;\; 100]$ µm. To sample these ranges, a trade-off must be found between (i) the of the coarseness of the sampling and (ii) the dimensions of the dictionary. In our experiments, we chose the sampling steps as 500 times the CRLBs for z and 50 times the CRLBs for r, that are experimentally shown to satisfy both limits. The general dictionary $\bar{\boldsymbol{C}}_2$ has 600 (z, r) pairs. The first 100 modes of the dictionary are calculated using the truncated SVD algorithm implemented in PROPACK. The obtained mode dictionary is then used to approximate the application dictionary $\bar{\boldsymbol{C}}_1$. It has been verified that the first 50 modes are enough to have estimations with standard deviations of the same order as the CRLBs and negligible bias.

Fig. 7 illustrates the evolution of the mean square difference between diffraction patterns, (i.e., the opposite of their correlation) around its global minimum increasing the number of modes. According to this illustration, the curvature of the main basin with about 10 modes is accurate enough for a coarse detection of the object's coordinates.

To show the efficiency of using the proposed method, we compare the time costs of the exhaustive search step of the original greedy method to the proposed method. The exhaustive search takes more than one minute and half to obtain the results using the full dictionary. Applying the proposed method, the first 5 modes are showed to be enough for accurate estimation of parameters. The time costs of the exhaustive search using such an approximated dictionary takes less than 10 seconds. The second step of the greedy algorithm (i.e., the local optimization step) is not changed in the proposed method and it takes in both cases less than 10 seconds. The time gain of using the proposed method is then a factor of 8.

It shall be noted that in the case of light back-propagation methods using the Fresnel dictionary [31] (i.e., a dictionary which contains only the Fresnel transforms and not the diffraction patterns), SVD does not necessarily speedup the reconstruction. The blue curve in Fig. 3 shows the singular values of the Fresnel dictionary in a descending order. As shown in this illustration, the vanishing rate of the singular values is very low, resulting in very inaccurate low rank approximations of the dictionary. Eckart-Young theorem states that the best rank-k approximation of the matrix is the one found by SVD. One can therefore conclude that it is not possible to have an accurate low rank approximate of the Fresnel dictionary. This is not surprising as the 3D capability of holography comes from the difference between diffraction patterns produced by point-like objects located at different depths (i.e., Fresnel transform is close to invertible). Extended objects affect the Fresnel transform as relative low-pass filters that decrease the significance of high frequencies. These filtering operations

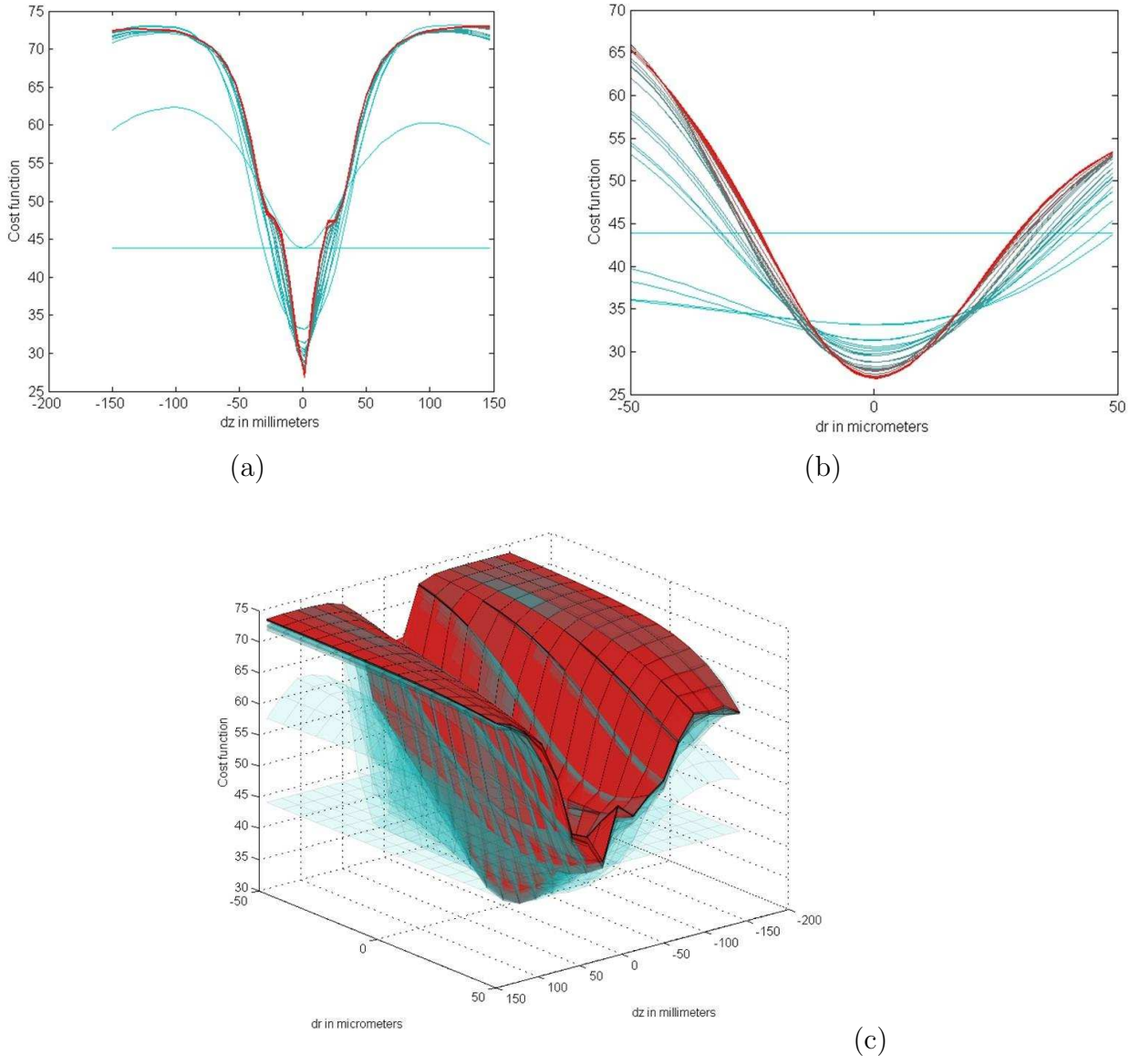


Fig. 7. With an increased number of modes, the approximation of the dictionary improves and matching a reference diffraction pattern against the approximated dictionary produces a sharper correlation peak leading to more accurate 3D location. Plotted curves represent the mean square difference between diffraction patterns, i.e., the opposite of their correlation. Reddest curves correspond to the most accurate approximations of the dictionary (using up to 50 modes). The pattern of the dictionary minimizing the mean square difference is in best match with the reference diffraction-pattern and gives the 3D location and diameter of the detected particle.

introduce correlations between the columns of the dictionary which result in fast vanishing singular values (see the red curve in Fig. 3). Accurate low-rank approximations can be obtained in these cases.

4.C. Experimental Holograms of spherical particles

The last set of experiments involves the reconstruction of captured holograms using an approximated dictionary. Our experiments show that a rank 5 dictionary is accurate enough to detect and successfully clean all the particles placed in the field of view of camera (see Fig. 8 for a captured hologram and the cleaned one). The time gain of the proposed method is then a factor of 8. These holograms (see Fig. 8-a) were captured of injected water droplets at the Department of Fluid Mechanics and Acoustics of Lyon (LMFA). The droplets were generated by a piezoelectric jetting device manufactured by MicroFab Technologies. This injector produces mono-dispersed droplets with radii of 31 μm $\pm$ 0.5 μm. The droplets were produced to be at distances ranging from 30 cm to 48 cm of a 1024×1280 pixel camera with pixel size of 21.7 μm and fill-factor of 0.84. This imaging setup had a magnification of 1.42 and the holograms were captured with an SNR close to 16. The search space contained 410 pairs of (z, r).

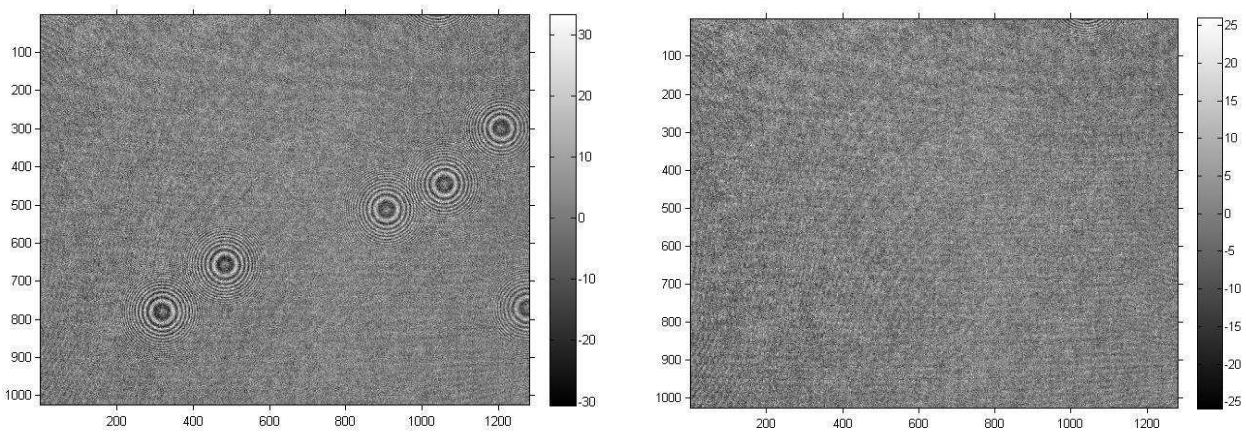
It should be noted that the detection of out of field particles [16] is very limited using the first modes of dictionary. The signatures of the out-of-the-field objects contain only the highest frequencies and these high frequencies are not present in the first modes of the dictionary. A future work can be directed towards finding a rigorous method of mode selection for out-of-field detection.

5. Conclusion

Signal processing methods can perform the object recognition task directly from the digital holograms. To reduce the computational costs of such approaches, we have proposed to obtain a lower dimensional space that captures the inter-pattern variability of the objects' holographic signatures. The calculation of the algorithm complexity shows that the time gain of this approach depends on the application parameters and is large when considering huge dictionaries of diffraction-patterns.

Our Monte Carlo studies on the application to 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 shown that very low rank approximations give results which are accurate enough for the greedy object reconstruction method to successfully find the accurate object parameters. To this end, we could achieve a speedup rate of 8 on our video of experimental holograms of water droplets.

We have shown in a primitive experiment the robustness of our proposed method to



(a) (b)

Fig. 8. (a): Experimental hologram of droplets from LMFA Lyon used in Sec. 4.C, (b): same experimental hologram cleaned from the in-field particles using the 5 first modes. The magnitude of the residuals is high due to the signature of the out-of-field particles placed close to the borders.

orientation changes. A more comprehensive study on the potential of the proposed approach to detect the orientation, scale and 3D location of objects from their diffraction patterns is a perspective of our work with potential interesting applications.

References

1. J. Katz and J. Sheng, "Applications of holography in fluid mechanics and particle dynamics," Annual Review of Fluid Mechanics **42**, 531–555 (2010).
2. J. Coupland and J. Lobera, "Optical tomography and digital holography," Measurement Science and Technology **19**, 070101 (2008).
3. K. D. Hinsch and S. F. Herrmann, "Holographic particle image velocimetry," Measurement Science and Technology **15** (2004).
4. J. Garcia-Sucerquia, W. Xu, S. K. Jericho, P. Klages, M. H. Jericho, and H. J. Kreuzer, "Digital in-line holographic microscopy," Applied Optics **45**, 836–850 (2006).
5. C. P. Allier, G. Hiernard, V. Poher, and J. M. Dinten, "Bacteria detection with thin wetting film lensless imaging," Biomedical Optics Express **1**, 762–770 (2010). PMID: 21258507 PMCID: PMC3018049.
6. X. Zhang, I. Khimji, U. A. Gurkan, H. Safaee, P. N. Catalano, H. O. Keles, E. Kayaalp, and U. Demirci, "Lensless imaging for simultaneous microfluidic sperm monitoring and

sorting," Lab on a chip **11**, 2535–2540 (2011). PMID: 21677993.

7. O. Mudanyali, E. McLeod, W. Luo, A. Greenbaum, A. F. Coskun, Y. Hennequin, C. P. Allier, and A. Ozcan, "Wide-field optical detection of nanoparticles using on-chip microscopy and self-assembled nanolenses," Nature Photonics **7**, 247–254 (2013).
8. S. H. Lee, Y. Roichman, G. R. Yi, S. 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 **15**, 1827518282 (2007).
9. F. C. Cheong, B. J. Krishnatreya, and D. G. Grier, "Strategies for three-dimensional particle tracking with holographic video microscopy," Optics Express **18**, 13563–13573 (2010).
10. F. Soulez, L. Denis, C. Fournier, É. Thiébaut, and C. Goepfert, "Inverse-problem approach for particle digital holography: accurate location based on local optimization," Journal of the Optical Society of America. A **24**, 1164–1171 (2007).
11. J. Fung, K. E. Martin, R. W. Perry, D. M. Kaz, R. McGorty, and V. N. Manoharan, "Measuring translational, rotational, and vibrational dynamics in colloids with digital holographic microscopy," Optics Express **19**, 8051–8065 (2011).
12. B. Javidi and E. Tajahuerce, "Three-dimensional object recognition by use of digital holography," Optics Letters **25**, 610–612 (2000).
13. T. Kim and T.-C. Poon, "Extraction of 3-d location of matched 3-d object using power fringe-adjusted filtering and wigner analysis," Optical Engineering **38**, 2176–2183 (1999).
14. T. Kim and T.-C. Poon, "Three-dimensional matching by use of phase-only holographic information and the wigner distribution," Journal of the Optical Society of America A **17**, 2520–2528 (2000).
15. "An application of high-speed microholography: the metrology of fogs," Nouvelle Revue d' Optique **5**, 87–93 (1974).
16. F. Soulez, L. Denis, E. Thiébaut, C. Fournier, and C. Goepfert, "Inverse problem approach in particle digital holography: out-of-field particle detection made possible," Journal of the Optical Society of America. A **24**, 3708–3716 (2007).
17. A. Bourquard, N. Pavillon, E. Bostan, C. Depeursinge, and M. Unser, "A practical inverse-problem approach to digital holographic reconstruction," Optics Express **21**, 3417–3433 (2013).
18. M. M. Marim, M. Atlan, E. Angelini, and J.-C. Olivo-Marin, "Compressed sensing with off-axis frequency-shifting holography," Optics Letters **35**, 871–873 (2010).
19. S. Lim, D. L. Marks, and D. J. Brady, "Sampling and processing for compressive holography [Invited]," Applied Optics **50**, H75–H86 (2011).
20. Y. Rivenson, A. Stern, and B. Javidi, "Compressive fresnel holography," Journal of Display Technology **6**, 506–509 (2010).

21. K. Choi, R. Horisaki, J. Hahn, S. Lim, D. L. Marks, T. J. Schulz, and D. J. Brady, "Compressive holography of diffuse objects," Applied Optics **49**, H1–H10 (2010).
22. X. Zhang and E. Y. Lam, "Edge-preserving sectional image reconstruction in optical scanning holography," Journal of the Optical Society of America A **27**, 1630–1637 (2010).
23. L. Denis, D. Lorenz, E. Thiébaut, C. Fournier, and D. Trede, "Inline hologram reconstruction with sparsity constraints." Optics Letters **34**, 3475–3477 (2009).
24. J. A. Tropp, "Greed is good: Algorithmic results for sparse approximation," Information Theory, IEEE Transactions on **50**, 2231–2242 (2004).
25. H. D. Simon and H. Zha, "Low rank matrix approximation using the lanczos bidiagonalization process with applications," SIAM J. Sci. Comput **21**, 22572274 (2000).
26. R. M. Larsen, *Lanczos Bidiagonalization with Partial Reorthogonalization* (Computer Science Department, Aarhus University, 1998).
27. "FFTW," http://www.fftw.org/. [Online; accessed 23-November-2012].
28. "OpenMP," http://openmp.org/wp/. [Online; accessed 23-November-2012].
29. "PROPACK," http://soi.stanford.edu/~rmunk/PROPACK/. [Online; accessed 23-November-2012].
30. C. Fournier, L. Denis, and T. Fournel, "On the single point resolution of on-axis digital holography," Journal of the Optical Society of America. A, Optics, Image Science, and Vision **27**, 1856–1862 (2010).
31. T. Kreis, *Handbook of Holographic Interferometry: Optical and Digital Methods* (Wiley-VCH, 2005), 1st ed.