Adding functionalities to precomputed holograms with random mask multiplexing in holographic optical tweezers

Josep Mas,† Michelle S. Roth,† Estela Martín-Badosa, and Mario Montes-Usategui*

Optical Trapping Lab—Grup de Biofotònica, Departament de Física Aplicada i Òptica, Universitat de Barcelona (UB), Martí i Franqués 1, Barcelona 08028, Spain
*Corresponding author: mario_montes@ub.edu

Received 18 August 2010; revised 13 January 2011; accepted 21 January 2011; posted 25 January 2011 (Doc. ID 133585); published 29 March 2011

In this study, we present a method designed to generate dynamic holograms in holographic optical tweezers. The approach combines our random mask encoding method with iterative high-efficiency algorithms. This hybrid method can be used to dynamically modify precalculated holograms, giving them new functionalities—temporarily or permanently—with a low computational cost. This allows the easy addition or removal of a single trap or the independent control of groups of traps for manipulating a variety of rigid structures in real time. © 2011 Optical Society of America

OCIS codes: 350.4855, 140.7010, 090.2890, 230.6120, 100.5090.

1. Introduction

During the last decade holographic optical tweezers have become a powerful tool for the dynamic manipulation of microscopic samples in three dimensions [1]. Computer generated holograms (CGH) are displayed on a spatial light modulator (SLM) in order to modify the complex amplitude of the trapping beam. Thus, the beam can be split, tilted, or reshaped, and arbitrary three-dimensional patterns of light can be created. The modulated beam is usually focused into highly concentrated light spots by a high numerical aperture optical system. Each light spot can be used as an independent optical trap that exerts controlled forces on the sample.

Several computational algorithms are available for calculating the holograms. The efficiency, uniformity, and speed requirements [2] of the experiment determine the choice of that algorithm. However, some of the most widely used methods for computing phase holograms do not come as closed-form solutions but are iterative, e.g. the Gerchberg–Saxton (GS) algorithm [3] or the direct-search algorithm [4]. The main drawback of iterative algorithms is the long computational time needed even with modern CPUs, which are currently too slow for three-dimensional real-time applications [5]. A good way to overcome this limitation is the use of a graphics processing unit (GPU) on recent graphics cards [6], which are optimized for use with pixel matrices, allowing the real-time computation of iterative CGHs. Unfortunately, although programming environments do exist for some specific graphic cards (e.g. in Compute Unified Device Architecture from NVIDIA Corporation [7]), performing calculations using a GPU is still a difficult task.

For a much faster and more direct method, the use of extremely simple algorithms such as the prisms-and-lenses algorithm [8] is required, which consists of generating a blazed phase grating for two-dimensional positioning of each light spot, and a quadratic phase term for axial positioning. Multiple 3D trap patterns can easily be created by adding the complex amplitudes of each trap and extracting the resulting phase for the encoding of the final hologram. As the amplitude modulation term is discarded, intensity fluctuations may occur within the
trap pattern. The prisms-and-lenses algorithm is faster than iterative algorithms so it can be carried out at video rates; with some corrections [9] one can also get CGHs with good performance for a small number of traps. However, spurious light spots might appear due to the spatial harmonics in the phase-only hologram. The presence of ghost traps in unwanted locations might result in practical problems when they block particles that are meant to be trapped at designated places, but also in periodic trap patterns, where these intermodulation terms might interfere with the original traps and reduce array uniformity.

Along the same line, in [10] we presented a simple technique for producing holographic optical traps, based on the random mask (RM) encoding method. It provides a noniterative and thus very fast algorithm for calculating phase holograms that can be used for optical trapping and micromanipulation at video rates. As opposed to the prisms-and-lenses algorithm, in RM-encoded holograms the spurious light is not concentrated. Instead, the light that is not properly modulated spreads over the whole sample plane. However, the method suffers from low efficiency.

In this article we present an alternative use of the RM-multiplexing method: a meta-algorithm that enhances the performance of holographic optical tweezers created in a previous step by another algorithm. The resulting hybrid method shares the advantages of both constituents: the efficiency of iterative methods in generating holograms (here used as building blocks) and the speed of RM multiplexing. Different holograms (each one generating a certain trap distribution) can be blended together to generate a more complex trap pattern, while maintaining independent control of each building block. Since the information of each constituent hologram is located at known pixel masks, the manipulation of a certain block is achieved by dynamically introducing modifications to its associated pixels. Thus, as opposed to other methods, moving a subgroup of traps within a more complex trap pattern does not involve the recomputation of the whole hologram because the remaining masks are not affected. The negative aspect of this flexibility is a decreased optical efficiency when the number of multiplexed blocks increases. This is discussed in Section 2 (theoretical) and Section 3 (experimental). To advantageously use the hybrid method, the flexibility (number of independent blocks) and trap strength (amount of light in the traps) must be balanced, as we discuss in Section 3. In that section we also show how the RM method is suitable for manipulating rigid structures in real time at a very low computational cost, using a set of precomputed GS holograms as building blocks. In Section 4 we show the practical implementation of the hybrid method where an auxiliary trap is dynamically added to a highly efficient but static trap pattern, easily loading the array with particles.

2. Efficiency of the Random Mask Encoding Method

The RM encoding method consists of splitting up the whole number of pixels of an SLM into N disjoint sets (masks) of randomly chosen pixels, $M_l (l = 1 \ldots N)$, $N$ being the number of holograms to be simultaneously encoded. The assignment of pixels to masks is performed randomly to avoid distorting the trap shapes [11]. Each mask displays a phase function $H_l (u,v)$ that can generate a single trap (as in the original method) or an arbitrary light distribution (as we will discuss below). Thus, the hologram $H(u,v)$ encoding all the holograms can be written as

$$H(u,v) = \sum_{l=1}^{N} m_l(u,v) \cdot H_l(u,v),$$  \hspace{1cm} (1)

with

$$m_l(u,v) = \begin{cases} 1 & (u,v) \in M_l, \\ 0 & (u,v) \notin M_l. \end{cases}$$  \hspace{1cm} (2)

A common problem with many CGHs is that phase-only holograms (e.g. in prisms-and-lenses algorithm) contain harmonics of the encoded information, resulting in the appearance of spurious traps. However, since the information of each hologram here is distributed in a random fashion, the light that is not going to the original traps is simply scattered over the sample.

Below, we analyze the efficiency of the RM algorithm. If we consider a hologram $H(j,k)$ of $P \times P$ pixels, a pure phase function encoding $N$ component holograms, and assume with no loss of generality that this hologram is illuminated by a plane wave of unit amplitude, $A = e^{i\phi}$, the energy at a plane immediately after the hologram is therefore

$$E_{tot} = \sum_{j=1}^{P} \sum_{k=1}^{P} |H(j,k)A|^2 \cdot H(j,k)A = \sum_{j=1}^{P} \sum_{k=1}^{P} |H(j,k)|^2 = P \cdot P = P^2.$$  \hspace{1cm} (3)

By virtue of Parseval’s theorem, this is also the total energy at the sample plane. Obviously, the amount of light devoted to each pixel mask will be proportional to the number of pixels in the mask. The field amplitude at a certain position $(x,y)$ in the sample plane (after an optical Fourier transform) can be written, in discrete notation, as

$$C(x,y) = \frac{1}{P} \sum_{j=1}^{P} \sum_{k=1}^{P} H(j,k) \cdot e^{-i2\pi jx/P} e^{-i2\pi ky/P}$$.  \hspace{1cm} (4)

For simplicity we now consider the particular case of $N$ multiplexed holograms in which each RM encodes a constituent hologram $H_l(j,k)$ generating one single shifted trap (blazed phase grating). Thus, in a certain trap $t$ located at $(x_t, y_t)$, with the hologram
encoded in mask \( m_t \) with \( n_t \) pixels, the field amplitude will be

\[
C(x_t, y_t) = \frac{1}{P} \sum_{j=1}^{P} \sum_{k=1}^{P} m_t(j, k) \cdot e^{-\frac{i \pi}{P^2} ((x - x_t)x + (y - y_t)y)}
\]

\[
\approx \frac{1}{P} \sum_{j=1}^{P} \sum_{k=1}^{P} m_t(j, k) = \frac{n_t}{P},
\]

considering that the light coming from mask \( m_t \) will only contribute significantly to the spot at position \((x_t, y_t)\) in the sample plane. Thus, the total energy focused in the traps will be

\[
E_{\text{traps}} = \sum_{t=1}^{N} \left( \frac{n_t}{P} \right)^2.
\]

The efficiency \( \eta \) i.e. the relationship between the energy in the traps and the total energy, can be easily calculated in the case of \( N \) equally weighted traps (\( n_t = P^2/N \) for every mask):

\[
\eta = \frac{E_{\text{traps}}}{E_{\text{total}}} = \frac{\frac{1}{P^2} \sum_{t=1}^{N} n_t^2}{P^2} = \frac{\sum_{t=1}^{N} \frac{1}{N^2}}{N} = \frac{1}{N}.
\]

This means that with an increasing number of constituent holograms the efficiency decreases, which is the main limitation of this technique. Thus, as a standalone procedure the RM algorithm is suitable only for generating a small number of traps; however, as we will show, it lends itself to be applied in combination with other, more efficient algorithms while keeping its advantages.

3. Hybrid Holograms Combining the Random Mask Encoding with an Iterative Algorithm

Holographic optical tweezers can be used as a tool to assemble small objects such as microbeads, cells, nano objects, and combinations of them [12,13]. In experiments involving the real-time manipulation of microscopic samples, a dynamic CGH calculation method is needed. Several micromanipulation experiments involve grabbing, moving, and rotating relatively large objects, which must be held from different points to achieve stable trapping with a controlled orientation. For example, in [14] a cell is suspended with several optical traps for immobilizing and performing controlled positional changes during in vivo imaging with Raman spectroscopy. In some cases more than one single object must be controlled independently e.g. in [15] where nontrivial structures are built from several semiconductor nanowires by holding, cutting, translating, rotating, and fusing them with holographic optical tweezers. This requires a dynamic control over the position and orientation of each part. In fact, the grouping of holographic optical traps into rigid assemblages that are then operated as a single object is a basic functionality included in many advanced interfaces, such as those in [16–18]. However, these group operations are not optimized at the algorithm level and always involve the recomputation of the whole hologram.

The hybrid holograms we suggest here make the manipulation of rigid groups of traps easier, taking advantage of the high efficiency of an iterative algorithm and the high speed of the RM technique, allowing real-time control of an experiment. A set of precomputed holograms (generated with a high-efficiency method such as GS) can be spatially multiplexed using random pixel masks. Each individual hologram is considered a building block and may encode a single trap, a group of traps, or even an arbitrarily complex light distribution. The computation of each individual high-efficiency hologram (which might be slow if iterative methods are used) is carried out before the experiment, when time is not a limitation. Then, during an experiment requiring dynamic control of the traps, the precomputed holograms are modified and multiplexed with the RM method. Any algorithm can be used in the first step as we only need to manipulate the end holograms. The resulting multiplexed mosaic will simultaneously reconstruct all the building blocks.

Unlike other methods, in RM the information of each building block is accessible because it is encoded in separate masks. Thus they can be manipulated independently at a very low computational cost, since the manipulation of a single block does not affect the other masks. A linear phase, a spherical phase, or an in-plane rotation can be added to any mask in real time, resulting in independent movements of a single block of traps, with no need to reconstruct the whole hologram. In particular, as the number of mathematical operations required in our algorithm scales as the number of pixels to be updated \( (U) \), whereas in most iterative algorithms the number scales as the product of the number of iterations \( (K) \), the number of traps \( (M) \), and the total number of pixels \( (P^2) \) [2], the relative gain of using our solution compared to recomputation of the hologram would be

\[
R = \frac{K \cdot M \cdot P^2}{U} = K \cdot M \cdot B,
\]

where the last identity holds for \( U = P^2/B \), with \( B \) being the number of building blocks, which is the case illustrated in Fig. 1 (equally weighted masks). In that example, the improvement would amount to 160 times fewer operations as \( K = 10, M = 8, \) and \( B = 2 \).

Figure 1 and Media 1 show how two rigid groups of traps are moved independently in real time using a custom-made LabView interface. In this experiment, two precomputed GS holograms were loaded and multiplexed into two random pixel masks. Adding linear phases to the original hologram causes a translation of the associated group of traps. Also rotating one of the multiplexed holograms in the screen
makes the associated group of traps rotate in 2D, as shown in Media 1. The center of rotation can be tuned by adding an initial linear phase to the hologram in a previous step. The manipulation of assembled groups of trapped objects is shown in Fig. 2 (frame sequence from Media 2), in which three independent blocks of four polystyrene spheres (1 μm in diameter) are moved along complex trajectories by means of three multiplexed GS holograms.

The holographic tweezers setup we used for the experiments presented in this paper was built around a commercial inverted microscope (Nikon Eclipse TE-2000E). The beam of an ytterbium fiber laser (IPG YLM-5-1064-LP) is expanded by a telescope to overfill a phase-only SLM (Hamamatsu X10468-03), and it is then directed through a second (inverted) telescope entering the microscope through the epifluorescence port, where a dichroic mirror sends the beam to the sample through an oil-immersion objective (Nikon Plan Flur 100×, 1.30 NA). The CGHs are sent to the SLM, and the encoded phase distribution is imaged onto the back focal plane of the microscope objective, where the beam slightly overfills the entrance pupil of the objective. Further details can be found in [19].

As the efficiency of the RM encoding method decreases with the number of masks, the performance of the final hologram is improved by using these groups of several traps as building blocks (compared to multiplexing single-trap holograms as in the original RM method). Meanwhile the practical advantages of the RM encoding method (speed, localized information, and lack of spurious light concentrations) are maintained. The downside is that the shape of each independent group of traps (each one associated to one precomputed hologram) must remain constant during the experiment.

In order to evaluate the performance of the hybrid method in practice, we measured the stiffness of several holographically generated optical traps as a function of the number of multiplexed holograms. For this experiment, we used an array of 4 × 3 optical traps, with approximately 4 μm separation between traps, and we trapped 12 polystyrene beads (Fluka 72938, 1.16 ± 0.04 μm in diameter) suspended in water. The array was first created using a single GS hologram. Then, four hybrid configurations generating the same trap pattern were constructed using a different number of multiplexed holograms (N) in every case. For each hybrid hologram, an elementary GS hologram representing a small group of 12/N traps was cloned N times using the RM encoding method.

Table 1 summarizes the construction process. Appropriate linear phases were added to each clone in order to place the traps exactly at the same locations as in the original array. In Fig. 3 the theoretical efficiencies of the final holograms are plotted, showing agreement with the expected efficiency behavior derived in Section 2 (efficiency proportional to the inverse of N), which was achieved even when the
individual holograms were not just linear phases but more complex GS holograms. Efficiency of individual holograms varies between 0.79 and 0.95.

The transverse stiffness \( k_x \) and \( k_y \) of the traps in the array were also estimated for each hybrid hologram as well as for the original GS hologram. Since there were several optical traps, the stiffness calibration was performed by video analysis. The Brownian motion of the trapped beads was registered with a CCD camera (QICAM 12 bit-mono Fast) for \( 10^5 \) frames. Particle trajectories were obtained with nanometer resolution using the Video Spot Tracker software from the University of North Carolina (CISMM, UNC-CH) [20]. Because the temporal resolution of the camera was not high enough to allow us to calibrate the trap via a spectral analysis, trap stiffness was estimated using the equipartition theorem:

\[
k_x = \frac{k_B T}{\sigma_x^2}, \quad k_y = \frac{k_B T}{\sigma_y^2},
\]

where \( \sigma_x \) and \( \sigma_y \) are the standard deviations of the measured particle path, \( T \) is the absolute temperature of the sample, and \( k_B \) is the Boltzmann constant. Interestingly, Fig. 4 shows that the mean transverse stiffness of the optical traps (averaged over the entire array) decreases as a power law of the number of multiplexed holograms in a similar way to efficiency behavior analyzed in Section 2, but with a faster decay.

Thus, when using the suggested hybrid method, it is important to carefully consider the balance between the number of independent parts and the strength of the optical traps. Splitting a trap pattern into separate pieces gives flexibility, allowing the easy manipulation of different objects, but at the expense of weaker optical forces. Because the amount of laser power in practice is limited, it is important to keep the efficiency of the holograms high enough to allow trapping. Thus, we recommend the use of the hybrid method with a reasonably low number of independent blocks.

![Fig. 3.](image1.png)  
Fig. 3. The efficiencies of the holograms used for the experiment in Fig. 4 were evaluated numerically by Fourier transforming each phase hologram to obtain the energy distribution at the sample plane. Efficiency was calculated as the ratio between the amount of energy at trap positions (computed by adding the square of the image values in the trap centers and its immediate neighbor pixels) and the total energy of the image. The spots with error bars show the efficiency of the holograms as a function of the number of multiplexed holograms \( N \). Error bars are obtained from trap intensity variation in the simulated image of the trap pattern. The solid line shows the fitted power law (efficiency = \( a \times N b \)) in the logarithmic graph, with \( a = 0.61 \pm 0.01 \) and \( b = -1.01 \pm 0.04 \).

![Fig. 4.](image2.png)  
Fig. 4. Experimental performance of the hybrid method for generating a \( 4 \times 3 \) array of optical traps. Data show the mean experimental transverse trap stiffness \( \langle k \rangle \) (averaged for the whole array) as a function of the number of multiplexed holograms \( N \). The values are relative to a reference stiffness \( \langle k \rangle_{ref} = 27 \pm 7 \text{ pN}/\mu\text{m} \), which corresponds to that of traps in the pure GS hologram. Error bars arise from the variability of stiffness in the array, which may come from the small differences in particle size, or possible power fluctuations from trap to trap. The solid line shows the fitting to an exponential function (negative) in the logarithmic graph, resulting in a power law (relative stiffness = \( a \times N b \)) with \( a = 1.00 \pm 0.08 \) and \( b = -1.6 \pm 0.2 \).
4. Loading Trap Patterns with an Auxiliary Trap

A particularly interesting implementation of the above method involves adding a free-moving auxiliary trap over a static trap pattern to either load it with samples, remove them to create vacant sites, or to make an existing trap vanish by superimposing a trap with the opposite phase.

Holographic optical tweezers can be used to create ordered trap arrays for research in colloid science, for example. However, only a few publications have addressed the problem of loading a large array of optical traps, which can cause particles to get trapped in the outer regions, leaving unoccupied traps at the center that do not get filled unless manipulated. This is particularly threatening in closely packed patterns. In [21] they suggest a method that involves temporarily extinguishing part of the traps in a large array. They place a knife edge in a conjugate of the object plane, blocking part of the beam so that only one row of traps remains. Once there is a particle trapped in each trap of that row, the knife edge is retracted, revealing the next row of traps. By repeating this process the entire array is filled. An active and thus more selective method was presented in [22]. It involves the use of an additional, steerable trap to fill an array of traps (in this example formed from the interference of two pairs of coherent laser beams), which also enables the generation of lattices with designed defects. The disadvantage of this technique is that it requires an extra laser to create the auxiliary trap.

However, using our hybrid method, a steerable helper-trap can be easily generated temporarily, using the same laser as was used for the trap array, simply by replacing the values of a number of randomly chosen pixels with a phase grating, producing a single off-axis trap. The new trap is then superimposed on to the original trap pattern while maintaining independent control of it using the associated pixel mask. The strength of the auxiliary trap can be controlled by manipulating the number of pixels dedicated to its hologram accordingly.

For example, we filled a square array of 41 traps with polystyrene microbeads of three different sizes, which were 1, 2, and 5 μm in diameter. We use a modified version of the holographic tweezers control software presented in [23] to add an auxiliary trap to an array of traps (in this example formed from the interference of two pairs of coherent laser beams), which also enables the generation of lattices with designed defects. The disadvantage of this technique is that it requires an extra laser to create the auxiliary trap.

![Fig. 5](image1.png)  
Fig. 5. (Color online) An array of 41 traps created with a Gerchberg–Saxton hologram was loaded with beads of three different sizes using an auxiliary trap. (a) An auxiliary trap (here marked with a black circle) is used to place the beads in the desired positions. (b) After the pattern is finished the extra trap can be removed (see Media 3).

![Fig. 6](image2.png)  
Fig. 6. (Color online) Elimination of a trap from an existing pattern. (a) The trap marked with a yellow circle is superimposed with an auxiliary trap. (b) If the phase and percentage are chosen correctly, the trap is extinguished.
the existing trap pattern. The auxiliary trap that is used to place the beads in the desired positions (loading the array or reshuffling it) is generated with a random mask made from 30% of the pixels of the GS hologram that encodes the trap array. Figure 5 (or Media 3) illustrates the trap loading process. Once all the traps are filled, the auxiliary trap is deleted and the original GS hologram is restored. In this case the helper-trap is moved by hand. However, one could also consider adding a feature recognition algorithm to create an automated process similar to the one presented in [24]. In this way, the highly efficient but static GS hologram is given a simpler, although much-needed, interactive capability.

There may also be situations where a particle must be removed from a trap. If the auxiliary trap is strong enough (i.e. if the number of pixels used to generate that trap is large enough), the particle can simply be dragged away or pushed out of the focus plane by moving the auxiliary trap in three dimensions. Alternatively, one of the existing traps could be eliminated by superimposing it with an auxiliary trap at the same coordinates but with a π phase difference, as shown in Fig. 6. To achieve the maximum extinction their intensities must be the same, which can be achieved by adjusting the percentage of pixels in the hologram dedicated to the auxiliary trap.

5. Conclusions

We have presented a hybrid method combining an iterative algorithm with the RM encoding method for manipulating groups of holographic optical traps in real time. The individual hologram for each block was previously computed using the highly efficient iterative algorithm, and the RM encoding method was used to quickly and accurately join the holograms. We have shown that with the proposed method, different rigid parts (i.e. groups of traps with a fixed spatial distribution according to the specific object shape and dimensions) can be independently manipulated without time-consuming computations. The main advantages of this method compared to others are its simplicity, its speed, and the lack of spurious light spots. We have shown how an existing trap pattern can be easily modified (e.g. by adding an auxiliary trap or extinguishing an existing trap) with minimal computational cost. The main drawback of the method is an inherent efficiency reduction, given by the inverse relationship of efficiency with the number of multiplexed holograms. A reduction in efficiency also reduces the trap stiffness in practice. Thus, there is an important trade-off between the desired hologram efficiency (the lower the number of masks, the better) and its flexibility, or number of independent pieces.

This research was funded by the Spanish Ministry of Education and Science project FIS2007-65880. M. S. Roth is grateful for a postdoctoral fellowship from the Swiss National Science Foundation (grant PBBE2-117006). J. Mas acknowledges the Spanish Ministry of Education for an FPU grant. The authors also thank E. Pleguezuelos and J. Andilla for their early contribution to this research.

J. Mas and M. S. Roth contributed equally to this work.

References