



Object-based digital hologram segmentation and motion compensation

TOBIAS BIRNBAUM,^{1,2,*}  DAVID BLINDER,^{1,2}  RAEES K. MUHAMAD,^{1,2} COLAS SCHRETTTER,^{1,2} ATHANASIA SYMEONIDOU,^{1,2}  AND PETER SCHELKENS^{1,2}

¹Vrije Universiteit Brussel (VUB), Dept. of Electronics and Informatics (ETRO), Belgium

²imec, Kapeldreef 75, B-3001 Leuven, Belgium

*tbirnbaum@etrovub.be

Abstract: Digital video holography faces two main problems: 1) computer-generation of holograms is computationally very costly, even more when dynamic content is considered; 2) the transmission of many high-resolution holograms requires large bandwidths. Motion compensation algorithms leverage temporal redundancies and can be used to address both issues by predicting future frames from preceding ones. Unfortunately, existing holographic motion compensation methods can only model uniform motions of entire 3D scenes. We address this limitation by proposing both a segmentation scheme for multi-object holograms based on Gabor masks and derive a Gabor mask-based multi-object motion compensation (GMMC) method for the compensation of independently moving objects within a single hologram. The utilized Gabor masks are defined in 4D space-frequency domain (also known as time-frequency domain or optical phase-space). GMMC can segment holograms containing an arbitrary number of mutually occluding objects by means of a coarse triangulation of the scene as side information. We demonstrate high segmentation quality (down to $\leq 0.01\%$ normalized mean-squared error) with Gabor masks for scenes with spatial occlusions. The support of holographic motion compensation for arbitrary multi-object scenes can enable faster generation or improved video compression rates for dynamic digital holography.

© 2020 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

1. Introduction

The optical acquisition of digital holograms (DH) outdoors and/or of moving objects is highly impractical because of illumination constraints, detector bandwidths, and setup stability requirements. Thus, the most likely source for holographic video content is computer-generated holography based on 3D data representations. The 3D data can be either fully synthetic or acquired from alternative imaging setups, such as a set of cameras recording arbitrary scenes from multiple angles; surface reconstruction and scene stitching can recreate a virtual world from the recorded content [1]. The design of a suitable end-to-end standard framework is the scope of the JPEG Pleno efforts on plenoptic image coding systems.

Since much of multimedia content is dynamic, efficient handling of holographic video sequences is an important task. Individual hologram frames with large apertures and viewing angles require resolutions of up to 10^{12} pixels. Compounding this fact with video frame rates imposes unrealistic bandwidth requirements, if the data is not compressed. The aim of this work is to advance the use of temporal redundancies between successive hologram frames for motion compensation. By predicting subsequent frames, only the modified parts have to be computed and/or signaled rather than the entire next frame.

Motion compensation algorithms attempt to predict a target frame from one or multiple reference frames as accurately as possible by using motion information across the frames. Those designed for conventional video typically minimize the mean-squared error between

motion-compensated and reference frame [2] by subdividing the reference frame into blocks and using associated motion vectors to obtain a best estimate of its contents. Unfortunately, this block-wise approach does not apply to holography where even small motions in the 3D scene will generally affect all hologram pixels. Therefore, several techniques have been proposed recently for rigid-body motion compensation in holography. These methods can be used either for a faster generation of holographic videos [3,4] or inter-frame video encoding [5,6]. As an example, [6] could achieve a reduction of used bandwidth from 7.5 Gbit/sec to 48 Mbit/sec by using holographic motion compensation and adaptive residual coding. We will review the exact holographic motion model briefly in the following section.

However, all methods proposed so far did consider only the compensation of uniform motions of the entire scene and thus cannot be applied for multiple objects moving independently. Furthermore, no (per-object) segmentation strategies for macroscopic holograms containing multiple objects have been published so far.

In this paper, we propose two schemes to segment holograms. They are based either on spatial or on Gabor masks. We combine the presentation of both schemes with the proposition of two motion compensation methods for multiple moving objects provided only a single hologram and per-object motion vectors. The two methods are the back-propagation-based multi-object motion compensation (BPMC) and the Gabor mask-based motion compensation (GMMC). BPMC is a naive compensation method inspired by digital holographic microscopy with limited applicability and used mainly for reference, whilst GMMC is a generic method which requires a rough scene triangulation as additional side information.

The rest of the paper is organized as follows. In section 2, we review some preliminaries on global motion compensation. In section 3, we describe the hologram segmentation schemes along with the two proposed multi-object holographic motion compensation methods - namely the BPMC and the GMMC. Thereafter, we present numerical experiments in section 4 that demonstrate the quality of the segmentation schemes and the effectiveness of the motion compensation methods in the context of multiple independently moving objects for two exemplary computer-generated holographic video sequences. We close the paper with a conclusion and outlook to future work in section 5.

2. Preliminaries on global motion compensation

The analytic model of global holographic motion compensation is formalized as follows: let $H(t)$ be a sequence of holograms with time instances $t \in \mathbb{N}$ (*i.e.* frame numbers) from a scene undergoing uniform rigid-body motions. Let $\alpha(t-1)$ be the associated motion vectors, describing the object motion between frames $t-1$ and t in scene space. With global motion compensation we aim to find the prediction $\tilde{H}(t)$ of $H(t)$, provided $H(t-1)$ and $\alpha(t-1)$, such that the ℓ_2 -error $\|H(t) - \tilde{H}(t)\|_2$ is minimal. Let x, y, z denote a right-handed Cartesian coordinate system in scene space and let ξ, η parametrize the hologram plane placed parallel to the x, y plane. We choose ξ, η parallel to x, y , respectively and let z point along the optical axis. The hologram plane is placed at $z = 0$. Let us further define the numerical back-propagation of a hologram H with wavelength λ from $z = 0$ to $\tilde{z}$, in scene space, within scalar diffraction theory as

$$O[x, y] := BP(H[\xi, \eta, 0]) := \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} H[\xi, \eta] \frac{1}{r} e^{i\phi(\xi, \eta, 0; x, y, \tilde{z})} d\xi d\eta, \quad (1a)$$

$$\phi(\xi, \eta, 0; x, y, \tilde{z}) := \frac{2\pi r}{\lambda} \quad \text{with} \quad r := \sqrt{(\xi - x)^2 + (\eta - y)^2 + (0 - \tilde{z})^2}. \quad (1b)$$

The term $\frac{1}{r} e^{i\phi(\xi, \eta, 0; x, y, \tilde{z})}$ is called point-spread function (PSF) and describes the diffraction pattern in the hologram plane due to a spherical wave emitted by a single point source in the scene. Each PSF yields generally a non-zero contribution for every ξ, η in the hologram H . Hereinafter,

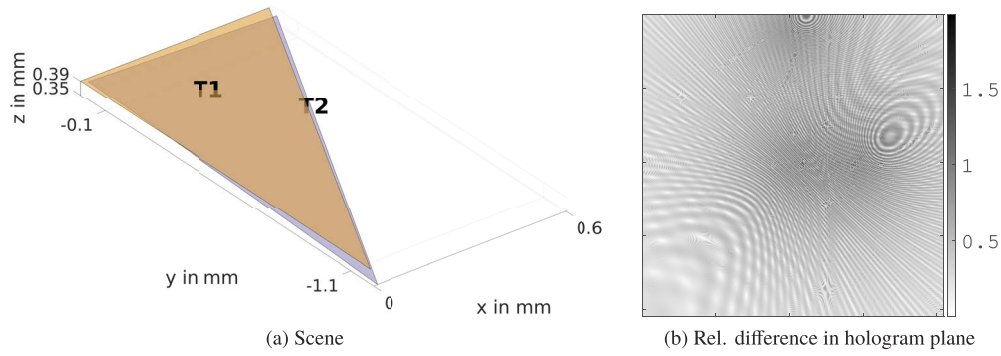


Fig. 1. (a) shows two triangles “T1” and “T2”. “T2” is obtained from “T1” by 2° rotations around the x and y -axes. (b) shows the amplitude of the relative difference of a hologram containing exactly the three point-spread functions corresponding to the vertices of either triangle.

we will shorten the notation of the (back-)propagation operation to $BP(\cdot)$ and $BP^{-1}(\cdot)$ for an adequately chosen $\tilde{z}$, respectively. For brevity, we will further write O , whenever we refer to a hologram re-focused to scene space and drop the mention of the z dependence.

With this, we can analytically model the effect of all elementary Euclidean motions in scene space on the hologram plane as follows. Let $O(t)[x, y] = \mathcal{Y}_{\delta, \epsilon} O(t-1)[x, y] = O(t-1)[x - \delta, y - \epsilon]$, where $\mathcal{Y}_{\delta, \epsilon}$ denotes a translation along x, y by δ, ϵ , respectively. Then, via a change of variables in Eq. (1a) we find:

$$H(t)[\xi, \eta] = BP^{-1}(O(t)[x, y]) = BP^{-1}(O(t-1)[x - \delta, y - \epsilon]) = H(t-1)[\xi - \delta, \eta - \epsilon]. \quad (2)$$

Thus, lateral translations in space map directly to translations along ξ, η , respectively. To avoid spatial interpolation for non-integer pixel shifts, phase shifting in the Fourier domain can be used instead [5]. Translations along z are described by Eq. (1a). Rotations of the scene space around z are described by rotations around z in the hologram plane. More involved are rotations around x and y which need to be compensated by a tilting of the hologram plane. The tilt is facilitated through a resampling of the Fourier domain and multiplication with a transfer function whose exact expression is given in [6,7]. These exact analytical models for scene space motions can be approximated as described in [5] or compensation can be performed in cylindrical or spherical coordinate systems, as done in [4,8].

However, since the analytical model relies solely on operations in the hologram plane or its Fourier domain, where contributions of every PSF are spread across the entire domain, motion compensation generally influences all pixels of H at once and can not be applied directly to the individual, independently moving objects for multi-object scenes. We show an example in Fig. 1.

3. Segmentation schemes and multi-object motion compensation methods

We propose two schemes to handle the segmentation of holograms containing multiple objects. These schemes are explained jointly with two holographic motion compensation methods, which can compensate the motion of multiple independently moving objects in a holographic video.

First, we propose a naive and fast method, referred to as back-propagation-based multi-object motion compensation (BPMC) and based on the object-based hologram segmentation in the spatial domain. Such a segmentation is not always possible and BPMC will certainly fail for deep scenes or with occlusions present. The approach of spatial segmentation is common in digital holographic microscopy, *e.g.* to refocus different specimen [9] or to track particle motions [10],

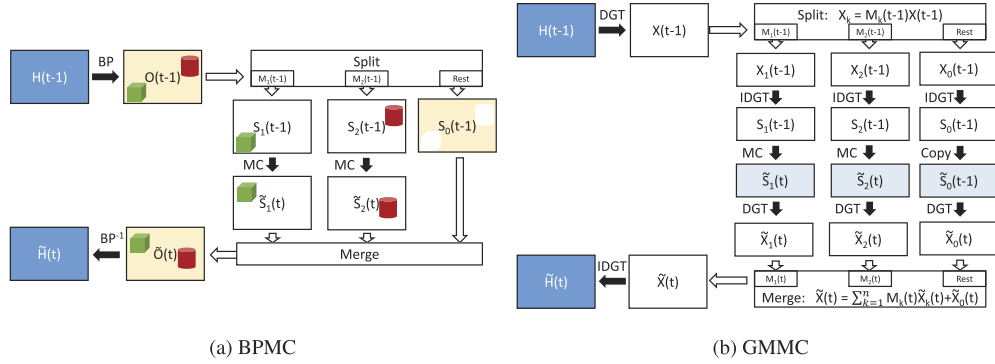


Fig. 2. Overview of naive (BPMC) and generic (GMMC) holographic motion compensation for scenes with multiple independently moving objects. Example for $K = 2$ objects plus background. BP: back-propagation, (I)DGT: (inverse) discrete Gabor transform, MC: motion compensation.

but it has, to our knowledge, thus far not been used for motion compensation in macroscopic holography.

Second, we propose a generic method: the Gabor mask-based motion compensation (GMMC). It is based on hologram segmentation in 4D phase-space using per-object masks defined in Gabor space which are generated from coarse object triangulations. The four dimensions of phase-space arise from the two spatial dimensions of the complex-valued hologram and two associated frequency dimensions, which correspond to the lateral viewing angles. An overview of both motion compensation methods is shown side by side in Fig. 2.

3.1. Back-propagation-based multi-object motion compensation (BPMC)

If a hologram contains a scene in which all objects are sufficiently shallow and placed at similar depths, the entire scene can be brought approximately into focus through numerical back-propagation Eq. (1a). Whenever objects are laterally well-separated throughout two subsequent frames, a spatial hologram segmentation and per-object motion compensation is plausible.

An example scene is shown in Fig. 3 by means of reconstructions as well as computed spatial object masks. We will refer to the motion compensation method based on this segmentation as back-propagation-based multi-object motion compensation (BPMC) in this work, see also Fig. 2(a). BPMC will be chosen as the naive reference method.

Given a hologram $H(t-1)$ and the motion vectors α_k , which describe the 3D motion of each moving object, $k \in \{1, \dots, K\}$, from $t-1$ to t , BPMC proceeds as follows:

1. **Back-propagating** the hologram $H(t-1)$, without using any aperture, from its hologram plane to the common focal plane of all objects in the scene. That is $O(t-1) := BP(H(t-1))$.
2. **Splitting** $O(t-1)$ using per-object spatial masks derived from any spatial scene segmentation scheme for natural images applied to $|O(t-1)|$. This yields one sub-hologram $S_k(t-1)$, $k \in \{1, \dots, K\}$ per object, plus $S_0(t-1)$ representing the residual of $O(t-1)$.
3. **Compensating** the global rigid-body motion per sub-hologram $S_k(t-1)$, $k \in \{1, \dots, K\}$, using the provided motion vectors $\alpha_k(t-1)$ and any global holographic motion compensation scheme yields the predicted individual object holograms $\tilde{S}_k(t)$. We set $\tilde{S}_0(t) = S_0(t-1)$.
4. **Merging** the predicted sub-holograms $\tilde{S}_k(t)$, $k \in \{0, 1, \dots, K\}$ through summation into the predicted hologram $\tilde{O}(t)$ within the chosen focal plane.

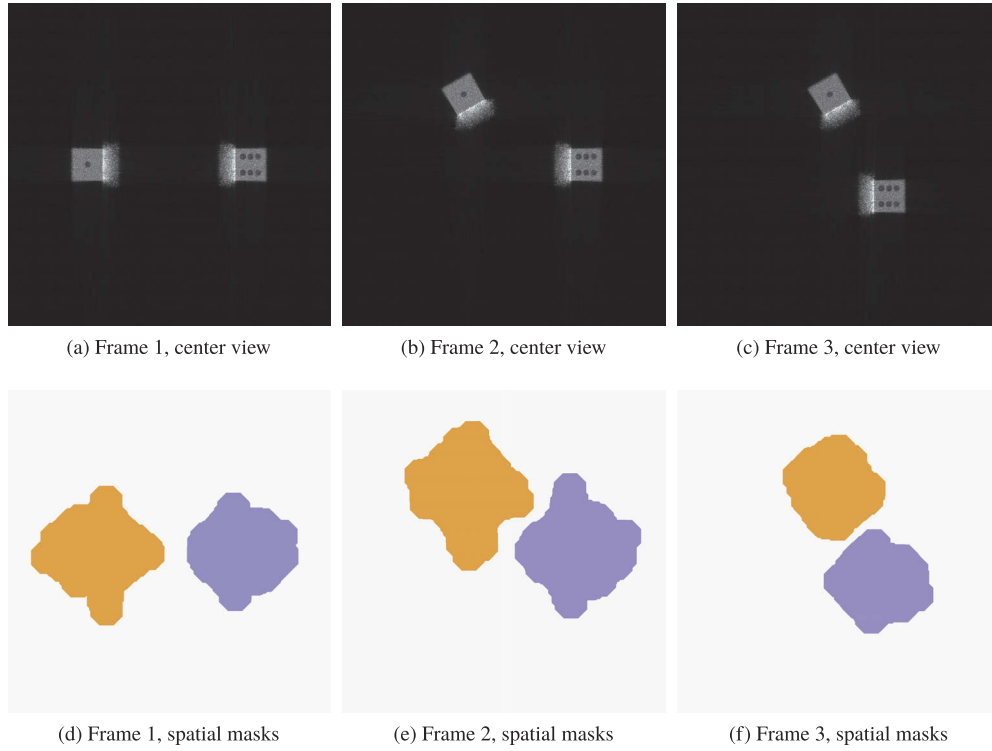


Fig. 3. The holographic sequence, called “split dices”, can be compensated with BPMC. Top: Reconstructions of frames 1-3 at central focal distance. Bottom: Object masks used in BPMC.

5. **Propagating** $\tilde{O}(t)$ to the original hologram plane, without using any aperture, finally returns the predicted master hologram $\tilde{H}(t) := BP^{-1}(\tilde{O}(t))$ at time instance t .

BPMC is a comparatively simple method involving two propagations of the entire hologram, K motion compensations, and one hologram segmentation step. It will fail, once objects are moving at vastly different depths, as the diffractive footprint of the objects with larger distance, along the optical axis, will bleed into closer objects due to the diffractive nature of holography and the associated spreading of information with increasing distance.

Another problem that cannot be handled well by considering only the spatial domain are occlusions. In the limit, occlusions that occur at time instance t can be approximated with object-wise shielding, similar to [11], where the occluded regions of the rear object(s) are masked before summation of the sub-holograms to yield $\tilde{O}(t)$. However, this masking will lead to artifacts on the rear object under off-axis viewing angles, due to the masking being done entirely in the spatial domain. For these reasons, we will study in the following a more generic framework based on segmentation in the space-frequency domain.

3.2. Gabor mask-based multi-object motion compensation (GMMC)

In this section, we are presenting some essential theory on the space-frequency domain first, before we elaborate on the Gabor mask segmentation scheme and the Gabor mask-based multi-object motion compensation (GMMC) method.

3.2.1. Motivating space-frequency domain segmentation for DH

In order to segment holograms of scenes in general arrangements, phase-space representations are highly advantageous. Figure 4 shows the example of two occluding objects placed in some out of focus plane. Neither in spatial nor in frequency domain two distinct objects are visible. The visible difference in the frequency domain stems merely from a difference in illumination intensity of the objects and is not visible in general. However, in space-frequency domain two band-limited signals can be seen, whose unequal slope is an indication for different object depths in 3D space.

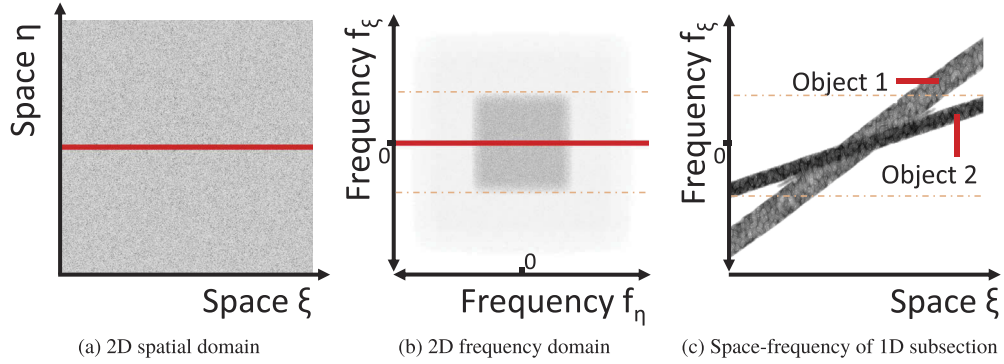


Fig. 4. The 2D amplitude of a hologram containing two occluding objects is depicted in (a) spatial, and (b) frequency domain. A 1D cross-section is highlighted in both domains and its phase-space is shown in (c). Otherwise inseparable objects appear well separated in phase-space.

For any PSF, and therefore any point-source in any hologram, the horizontal and vertical (instantaneous) spatial frequencies f_ξ, f_η can be computed within the assumption of stationary phase [12]. The latter states that the (instantaneous) phase φ , Eq. (1b), is approximately sinusoidal while varying ξ, η over several λ . f_ξ and f_η are given as

$$f_\xi(\xi, \eta, 0; x, y, z) := \frac{1}{2\pi} \frac{\partial \varphi}{\partial \xi} = \frac{(\xi - x)}{\lambda \sqrt{(\xi - x)^2 + (\eta - y)^2 + z^2}}, \quad (3a)$$

$$f_\eta(\xi, \eta, 0; x, y, z) := \frac{1}{2\pi} \frac{\partial \varphi}{\partial \eta} = \frac{(\eta - y)}{\lambda \sqrt{(\xi - x)^2 + (\eta - y)^2 + z^2}}. \quad (3b)$$

Provided a suitable space-frequency representation allowing access to well-localized areas of the space-frequency domain, we can thus derive a mapping between 3D volumes in space to 4D phase-space volumes and subsequently leverage it for hologram segmentation.

3.2.2. GMMC method - overview

The GMMC method, shown in Fig. 2(b), can be used to compensate for the motions of multiple independently moving objects captured by a single digital ground-truth hologram – subsequently called “master hologram”. GMMC relies on a space-frequency domain segmentation of the master hologram facilitated by Gabor masks and is described in the following. The segmentation is based on a coarse scene triangulation and works for holograms of deep scenes and, in principle, for arbitrarily many objects of arbitrary size, shape, and positions. Occlusions can be handled and GMMC works irrespective if the objects are voluminous or hollow shells. Note, that GMMC does not account for illumination in its present form. An example of such a scene is given as

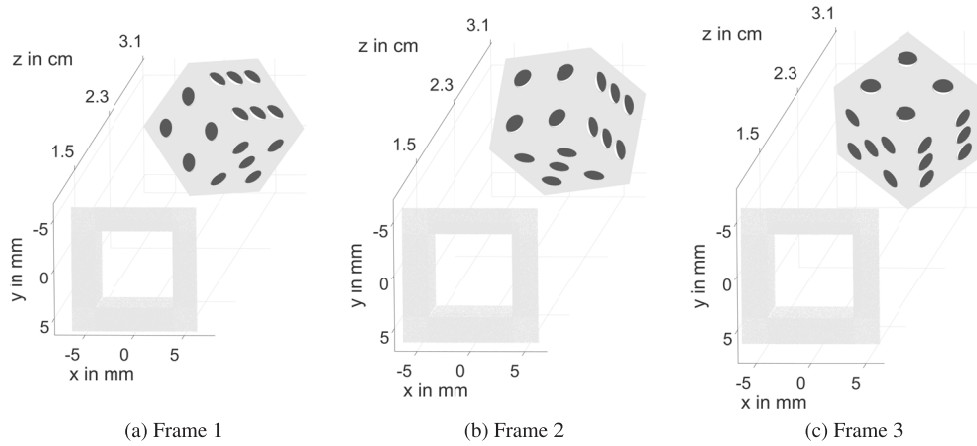


Fig. 5. Point cloud models for the 3 frames of the “spyhole” hologram sequence are shown.

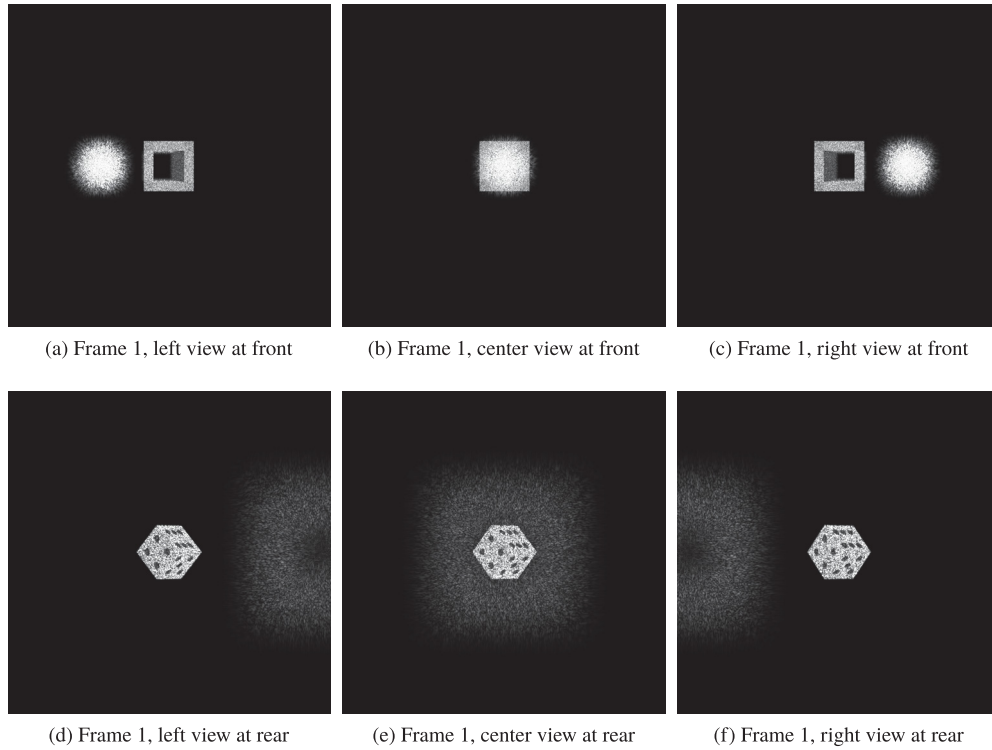


Fig. 6. Perspective front, (a)-(c), and rear, (d)-(f), reconstructions of frame 1 of the “spyhole” hologram sequence containing occlusions are shown.

point cloud models in Fig. 5, along with the reconstructions from the corresponding first frame of the holographic video in Fig. 6. The phase-space representation of frame 2 is shown in Fig. 4(c).

GMMC consists of several algorithmic blocks, of which the most vital contribution certainly is the generation of the per-object masks for Gabor coefficients provided some triangulation. The GMMC procedure is outlined as follows:

1. **Forward discrete Gabor transform** used to render the master-hologram $H(t-1)$ at time instance $t-1$ accessible for manipulations in 4D space-frequency domain.
2. **Mask generation** of M_k used to retain a sub-selection of all Gabor coefficients X_k belonging to one object by leveraging a rough triangulation of the scene.
3. **Splitting** the master hologram by application of the mask $M_k(t-1)$ to its Gabor coefficients $X(t-1)$ and using scene awareness to account for occlusions. A subsequent inverse Gabor transform yields one sub-hologram $S_k(t-1), k \in \{1, \dots, K\}$ per object, plus $S_0(t-1)$ representing the residual of $H(t-1)$.
4. **Compensating** for the global rigid-body motion per sub-hologram $S_k(t-1), k \in \{1, \dots, K\}$, using the provided motion vectors $\alpha_k(t-1)$ and any global holographic motion compensation yields the predicted object sub-holograms $\tilde{S}_k(t)$. We set furthermore $\tilde{S}_0(t) = S_0(t-1)$.
5. **Merging** the predicted sub-holograms $\tilde{S}_k(t)$ into the predicted master hologram $\tilde{H}(t)$ is done using another forward Gabor transform and newly generated Gabor coefficient masks $M_k(t)$ to address occlusions apparent after motion compensation.
6. **Inverse discrete Gabor transform** used to retrieve a hologram from the manipulated 4D space-frequency domain after occlusion-aware merger of the compensated sub-holograms.

In the following, we shall elaborate on each of these points in a separate subsection.

3.2.3. Forward / inverse discrete Gabor transform

As explained DHs are easily understood and manipulated in space-frequency domain [13]. We select the Gabor transform to yield an intermediate space-frequency representation for the splitting of the master holograms and merging operations of the predicted sub-holograms, before retrieving back the signal in the spatial domain of the hologram plane. The Gabor transform is an excellent candidate because it tiles phase-space uniformly by employing frequency analysis of a signal over a small region of space, called a “window” g . It has typically a small, bounded support and is apodized to resemble a Gaussian, such as Hamming windows. The Gabor transform $\mathcal{G}(H; g)$ is facilitated by scalar products of the analyzed signal with a “Gabor system” consisting of translations and frequency modulations of that base window. Figure 7(b) showcases the real

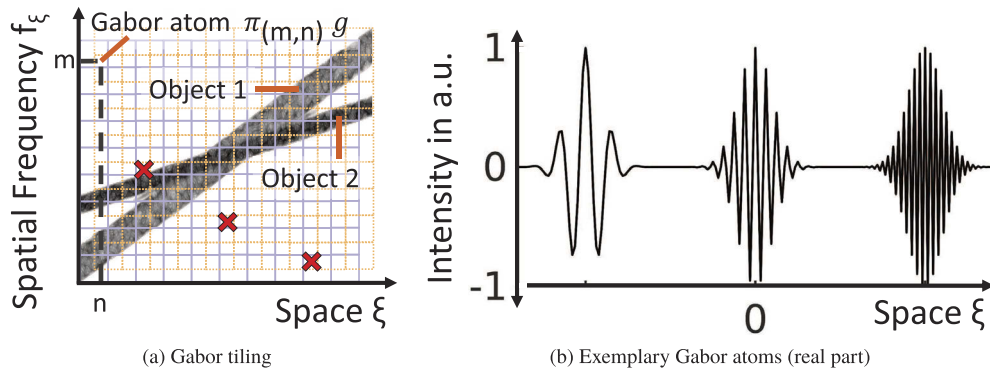


Fig. 7. A redundant 2D phase-space tiling ($r = 2$) shown in (a) atop the phase-space of a 1D signal, see Fig. 4(c). Space and frequency are discretized by integers $n \in \{0, \dots, N-1\}$, $m \in \{0, \dots, M-1\}$. (b) shows exemplary unit amplitude Gabor atoms at the grid positions indicated with crosses.

part of three unit amplitude Gabor atoms in 1D with varying spatial position and frequency modulations.

The Gabor system $\mathcal{G}(\cdot; g)[m, n]$, with $m \in \{0, 1, \dots, M-1\}$ frequency modulations and $n \in \{0, 1, \dots, N-1\}$ spatial translations, covers the full space-frequency bandwidth of the analyzed signal. For a 1D input signal $H \in \mathbb{C}^L$ the Gabor transform yields a matrix $X \in \mathbb{C}^{M,N}$ of Gabor coefficients. $|X[m, n]|$ is proportional to the energy of H near the frequency mb at the spatial position of na , see Fig. 7(a), whereas na and mb are the Gabor atom offsets in space and frequency in pixel. The factor $r := \frac{MN}{L}$ is called redundancy of the Gabor system; it is equal to the ratio of Gabor atoms to input samples (*i.e.* hologram pixels). The transform encodes all the information found in the signal if $r \geq 1$ and is thereby invertible. To guarantee stability of the inverse discrete Gabor transform (IDGT) $\mathcal{G}^{-1}(X; \gamma) = H$, it is required that $r > 1$ due to the Balian-Low theorem [14] and that γ is a dual window to g [15].

The Gabor coefficients of a 2D hologram $H \in \mathbb{C}^{L \times L}$ form a 4D set of Gabor coefficients denoted as $X[m_1, m_2, n_1, n_2] \in \mathbb{C}^{M \times M \times N \times N}$. To keep the notation simple, we will consider only square holograms $H \in \mathbb{C}^{L \times L}$ with equal Gabor systems along each dimension. Specifically, provided L and a desired redundancy $r > 1$, we used the following values for N, M, a, b :

$$\begin{aligned} [M', N'] &:= \text{factor}(L), \quad N := r_{\text{spat}} \cdot N', \quad a := \frac{L}{N}, \\ [r_{\text{spat}}, r_{\text{freq}}] &:= \text{factor}(r), \quad M := r_{\text{freq}} \cdot M', \quad b := \frac{L}{M}, \end{aligned} \quad (4)$$

with $\text{factor}(\cdot)$ being any function that factors natural numbers into two integers $[p, q] \in \mathbb{N}$, such that $q - p$ is minimal and $0 < p \leq q$. For the employed discrete Gabor transform (DGT), we choose as windowing function

$$l_1, l_2 \in \{0, 1, \dots, L-1\} : g[l_1, l_2] := \exp\left(-\frac{(l_1 - \lfloor \frac{L}{2} \rfloor)^2 + (l_2 - \lfloor \frac{L}{2} \rfloor)^2}{2\sigma^2}\right), \quad (5)$$

with some variance $\sigma > 0$ in this work set as $\sigma = \frac{aM}{L} = \frac{a}{b}$. The symbol $\lfloor \cdot \rfloor$ denotes the flooring operation (rounding down). We use $r = 2$ unless stated otherwise, as it is sufficient to guarantee a stable numerical reconstruction without imposing a large calculation overhead. More degrees of redundancy could be introduced, by adding a scaling dimension to the Gabor systems, such as Gabor wavelets which have been applied to DH in [16].

3.2.4. Mask generation

GMMC relies on a time-frequency segmentation scheme, which segments a single two-dimensional hologram by application of four-dimensional binary masks to its Gabor coefficients, leading to a sub-holograms for each of K objects — moving independently in 3D space. GMMC subsequently compensates the individual object motions in the hologram plane. The masks can be used to handle occlusions upon segmentation into and merger of the sub-holograms. They are calculated from rough triangulations of the scene space. Per triangulation one mask is obtained.

To better visualize the problem, we will rearrange the four-dimensional array of Gabor coefficients into a two-dimensional matrix. For example, we can choose an arrangement, where the coefficients corresponding to all the possible spatial frequencies (viewing angles) (m_1, m_2) form a sub-image per lateral spatial position in the hologram plane (n_1, n_2) and all sub-images are placed next to each other. Each sub-image will show the scene as it would be observed through a pinhole at (n_1, n_2) . The spatial frequencies within each sub-image are related to the lateral

viewing angles θ_i per dimension $i \in \{1, 2\}$ by

$$\forall i \in \{1, 2\} : \theta_i = \arcsin\left(\frac{\lambda}{2} m_i f_c\right) \quad \text{with} \quad f_c := \frac{1}{\Delta_i}, \quad (6)$$

where m_i is a normalized frequency, with range $[-1, 1]$, λ is the wavelength of the monochromatic light used to record the hologram, and Δ_i is its pixel pitch in meters along dimension $i \in \{1, 2\}$. f_c is also called “critical frequency” and is the largest frequency that can be sampled by any DH of the specified pixel pitch according to the Nyquist-Shannon bound.

An excerpt of the obtained arrangement showing 2×2 of 128×128 clusters, is shown in Fig. 8. Figure 8(a) shows only the amplitudes of the coefficients. Figure 8(b) shows the amplitudes and the predicted masks applied and color coded per object. For reasons that will become clear in section 3.2.4.2 the direct prediction may be insufficient. Thus Fig. 8(c) depicts the same masks, as finally used after applying a dilation operator.

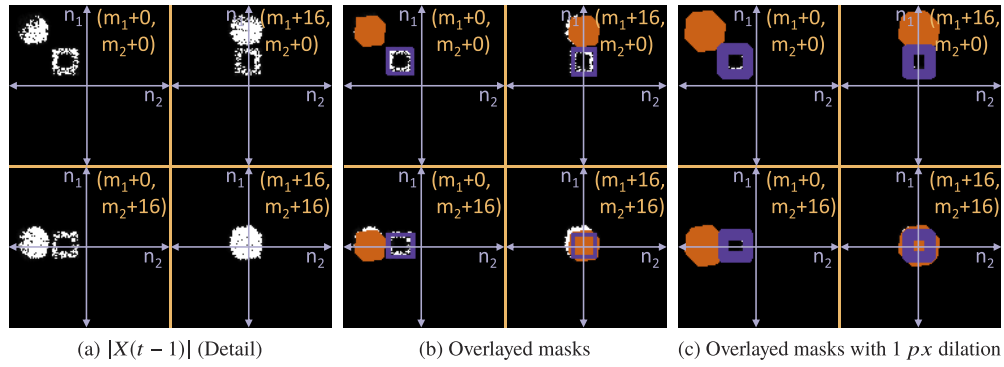


Fig. 8. Detail (2×2 out of 128×128 sub-images) from the two-dimensional re-arrangement, clustered by spatial frequencies $[m_1, m_2]$, of the amplitude the of four-dimensional Gabor coefficients $|X[m_1, m_2, n_1, n_2]|$ (a) as well as overlaid masks (b) without and with additional dilation (c) for the individual objects.

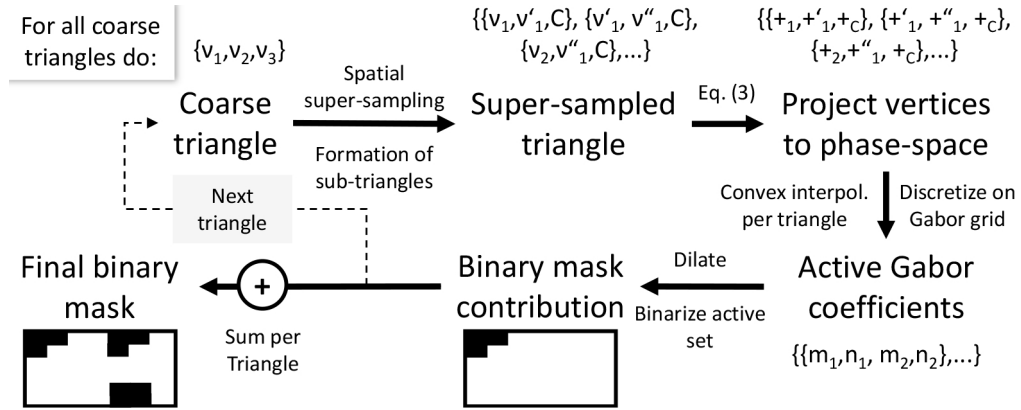


Fig. 9. An overview of the binary Gabor mask generation procedure is sketched.

In the following, we will describe the synthesis of the binary masks, for isolated points and thereafter for entire objects.

Space-spatial frequency relationship for individual points Using the mappings Eq. (3), we now deduce which Gabor coefficients (n_1, n_2) of the hologram plane, mapping to the frequencies f_ξ, f_η , will be affected for any given 3D point source at (x', y', z') . Given (x', y', z') and a target hologram of size $L \times L$, we evaluate Eq. (3) for each spatial grid position $(\xi, \eta) = (\xi[n_1], \eta[n_2])$

$$n_1 \in \{0, 1, \dots, N-1\} : \xi[n_1] := \frac{L\Delta_1}{2N} (2n_1 - N + 1), \quad (7a)$$

$$n_2 \in \{0, 1, \dots, N-1\} : \eta[n_2] := \frac{L\Delta_2}{2N} (2n_2 - N + 1). \quad (7b)$$

Due to the use of the exact expression for the instantaneous frequency, there are no restrictions on the diffraction regime for the mask generation — that is the scheme will work for all $\Delta, \lambda, z \gg \lambda$.

The obtained values (f_ξ, f_η) for each (n_1, n_2) are then discretized onto the discrete spatial frequency grid $(f_\xi[m_1], f_\eta[m_2])$ provided by the Gabor transform.

$$m_1 \in \{0, 1, \dots, M-1\} : f_\xi[m_1] := \frac{1}{2\Delta_1 M} (2m_1 - M + 1), \quad (8a)$$

$$m_2 \in \{0, 1, \dots, M-1\} : f_\eta[m_2] := \frac{1}{2\Delta_2 M} (2m_2 - M + 1). \quad (8b)$$

The phase-space volume accessed in the hologram plane by a point source located at (x', y', z') is the union of all interpolated $[m_1, m_2]$ for all $[n_1, n_2]$.

From triangulation of objects to 4D masks Now, we discuss the mapping between triangles in scene space to 4D phase-space volumes before we state the algorithm to map triangulated 3D volumes to phase-space— an overview of which is presented in Fig. 9. Let us consider the phase-space footprint of a triangle placed in scene space for a fixed position on the hologram plane (ξ, η) . The question is, what shape does this triangle take in the (f_ξ, f_η) plane?

To understand this, imagine the hologram being completely opaque, except for in pixel (ξ, η) , and imagine observing the illuminated scene through the transparent pixel. Then depending on the position of this single-pixel aperture, we will observe I) the scene under different perspectives and II), depending on the propagation distance z , the scene will appear with a barrel distortion centered at the optical axis, see Fig. 10(b). I) can be rephrased as: rays emitted by the same scene points will be perceived as stemming from different directions for each fixed (ξ, η) . And since in diffractive optics directions are mapped to spatial frequencies via Eq. (6), the mask of active coefficients in (f_ξ, f_η) will take again the shape of a (perspectively distorted) triangle.

The effect of II) is illustrated in the top row of Fig. 10, where the scene space triangle shown in Fig. 10(a) expands around the optical axis to the shape shown in light blue in Fig. 10(b). This is due to the spherically expanding wavefronts mapping ν onto the corresponding points in phase-space, indicated as +, upon propagation. These points are eventually mapped onto the discrete Gabor grid to find the active Gabor coefficients $X(m_1, m_2, n_1, n_2)$ per triangle, see Fig. 10(c). To account for II), without requiring more triangles to be signaled, we employ the following super-sampling technique showcased in the bottom of Fig. 10:

1. Uniform spatial super-sampling of every edge of each triangle, thereby dividing each into $h+1$ segments of equal length. See $\nu'_j, \nu''_j, \forall j \in \{1, 2, 3\}$ in Fig. 10(d) with super-sampling $h=3$.
2. Refine each signaled triangle by forming triangles from all two neighboring vertices ν along the edges with the center of the initial triangle (C) as the third vertex.

Algorithm 1 Gabor mask generation - GMMC

```

1: procedure TRI_TO_4Dvol(Convex triangulation  $T \in \mathbb{R}^{3\Omega \times 3}$ , Gabor parameters  $M, N$ ,
   Hologram parameters  $\lambda, \Delta_1, \Delta_2$ )
2:   Initialize  $M \in \mathbb{R}^{M \times M \times N \times N}$ 
3:   for Triangle  $\omega \in \{0, 1, \dots, \Omega - 1\}$  do
4:     for  $n_1 \in \{0, 1, \dots, N - 1\}$  do
5:       for  $n_2 \in \{0, 1, \dots, N - 1\}$  do
6:         Initialize  $tri \in \mathbb{R}^{3 \times 2}$  with 0.
7:         for Vertex  $v \in \{1, 2, 3\}$  do
8:            $r = \sqrt{(\xi[n_1] - T[3\omega + v, 0])^2 + (\eta[n_2] - T[3\omega + v, 1])^2 + T[3\omega + v, 2]^2}$ 
9:            $f_\xi = (\xi[n_1] - T[3\omega + v, 0]) / (\lambda r)$  ▷ See Eq. (3).
10:           $f_\eta = (\eta[n_2] - T[3\omega + v, 1]) / (\lambda r)$ 
11:           $tri[v, 0] = f_\xi$  and  $tri[v, 1] = f_\eta$ 
12:         Perform convex interpolation in  $[m_1, m_2]$  domain using vertices in  $tri$ .
13:         ▷ See Eq. (8) for spatial frequency grid.
14:         Set  $M[m_1, m_2, n_1, n_2] = 1$  for all  $[n_1, n_2]$  described by the convex interpolation.
15:   return Binary mask  $M$  of active Gabor coefficients

```

Because Eq. (3) is exact, all v will be mapped onto their correct phase-space projections (+) and the mismatch of the activated volume (dark-blue in Fig. 10(e)) to the precise shape (light-blue) is minimized and eventually zero after discretization on the Gabor grid, if the super-sampling factor is large enough. By increasing the granularity of the refined triangulation below the Gabor transform's space-frequency resolution any distorted shape is reproduced exactly and the barrel effect remains unresolved. With the refined triangles at hand, we obtain a mask for the 4D phase-space volume occupied of any triangle by

1. **Evaluating** for each fixed (ξ, η) from the spatial Gabor grid Eq. (7), the impacted frequencies (f_ξ, f_η) for each of the 3 corner vertices of the refined scene space triangle.
2. **Forming** the convex hull of the three points in (f_ξ, f_η) plane yields the perspectively distorted triangle in the (f_ξ, f_η) plane.
3. **Discretizing** on the spatial frequency Gabor grid, using Eq. (8) to markdown explicit entries $[m_1, m_2, n_1, n_2]$ in the mask.

Therefore, by knowing the footprints of three corner vertices of a triangle alone, all interior points of the triangles will be mapped out in phase-space, thus tremendously reducing the complexity of the mask generation.

Finally, the 3D scene space volume of any object k is mapped onto a 4D phase-space volume by forming the union of the 4D coefficients activated by each triangle of a convex, coarse triangulation $T_k(t-1)$ of the surface of the object. The convexity of the triangulation ensures that the union covers interior points as well. We are therefore only required to repeat the mapping for all refined triangles in $T_k(t-1)$ per object k to learn which Gabor coefficients $X(t-1)$ will carry the signal of the entire object. The convexity of the triangulation is a weak limitation. A non-convex triangulation can be split either into several convex sub-triangulations or can be approximated by a convex encapsulating triangulation. In the latter case, we may obtain a more detailed mask of a non-convex object by subtracting from the mask of a convex encapsulating triangulation, one or multiple masks corresponding to convex sub-triangulations of "holes" in the object. See, for example the mask of the spyhole in Fig. 8(c). If all parts are compensated in the same way, no difference will be apparent.

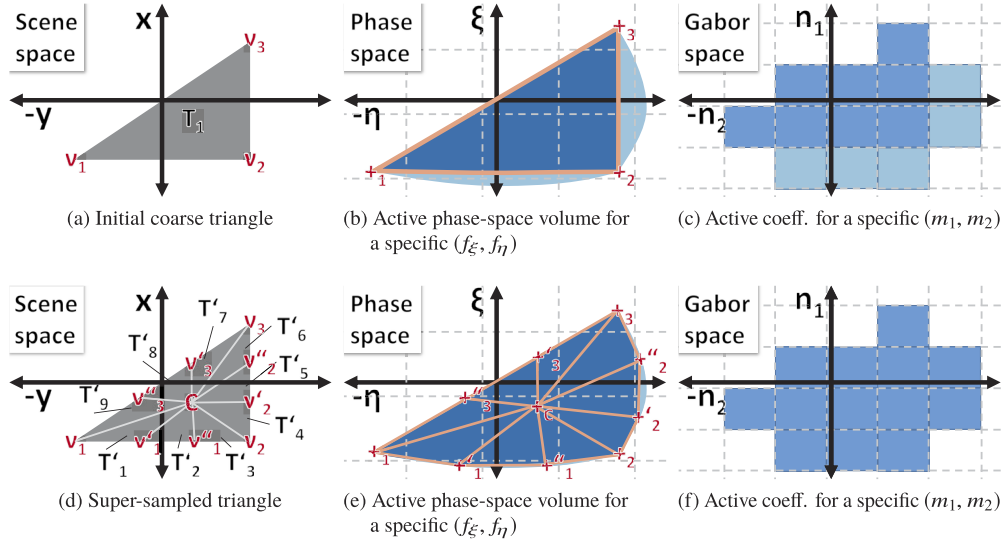


Fig. 10. Example on the necessity of super-sampling of coarse triangulations for accurate Gabor mask creation. The left column shows the spatial vertices of a given triangle (coarse on top, super-sampled to $h = 3 \times$ the number of vertices). The center column shows the area defined by linear interpolation and convex interpolation between the vertices of each triangle projected with Eq. (3) into phase-space in dark-blue for a specific (f_ξ, f_η) . The exact shape of the triangle is shown in light-blue underneath and the phase-space volumes occupied by individual Gabor coefficients are indicated by dashed lines. The right column shows the active volumes after discretization onto the Gabor grid and binarization. (f) shows that the super-sampling is sufficient.

Note, that the size of the side-channel information required by GMMC in the form of T and α is much smaller than the actual data. For example, Ω triangles in a triangulation T of K objects require per frame overhead of at most $9\Omega K$ real-valued single-precision entries encoding the vertex coordinates and edges. In the simple cases of a tetrahedral and cuboidal triangulation, Ω is 4 and 12 respectively. The motion vectors α_k per object can be encoded in $6K$ entries.

We summarized the procedure of the mask generation in Alg. 1 for a refined triangulation containing the coordinates of Ω triangles stored as row (x', y', z') per vertex.

The result of Alg. 1 applied to a super-sampled triangulation was shown in Fig. 8(b). As can be seen, the obtained mask may still not cover the entire set of activated Gabor coefficients, *e.g.* due to the rounding of the calculated 4D projected coordinates of the triangle vertices when mapping them onto the Gabor grids in space and frequency. The resolutions in space and spatial frequency are given by Eq. (7) and Eq. (8). In a final step, one may therefore perform a dilation on the generated masks, obtaining for example Fig. 8(c) with a dilation by a ball of 1 px radius in discretized space (n_1, n_2) and spatial frequency (m_1, m_2) . Empirically, we found that a radius of 2 px was sufficient in all considered cases. Detailed results will be presented in the section 4.

3.2.5. Splitting of the master hologram

The per-object masks $M_k(t-1)$ obtained in the previous section require a minor modification before they can be used to split up the master hologram $H(t-1)$. To account for occlusions in scene space, the order in which the K sub-holograms are extracted matters as 4D volumes of different objects can overlap when rays tracing from a rear object are occluded. To facilitate the hologram segmentation, sub-holograms of the front-most objects are extracted first and before proceeding towards the rear while ignoring already extracted content. In the simplest case, one

can define a processing order by sorting the K objects by their proximity to the hologram plane obtained via sorting the centers of the provided triangulations $T_k(t-1)$ by their z coordinates. Let the resulting permutation of the K objects be denoted as $\Pi(\{1, \dots, K\})$. We thus modify the sorted masks $M_p(t-1)$, $p \in \Pi(\{1, \dots, K\})$ for any $p' > p$ by zeroing out mask coefficients in $M_{p'}$ that were already extracted earlier on. We define new masks $\widehat{M}_p$ as

$$\widehat{M}_0 := \text{BIN}\left(1 - \sum_{q=1}^K M_q\right), \quad \forall p \in \Pi(\{1, \dots, K\}) : \widehat{M}_p := \text{BIN}\left(M_p - \sum_{q=p+1}^K M_q\right), \quad (9)$$

where

$$\forall x \in \mathbb{R} : \text{BIN}(x) := \begin{cases} 1 & \text{if } x \geq 1 \\ 0 & \text{if } x < 1 \end{cases}. \quad (10)$$

Thereby, $\widehat{M}_0$ contains all static scene parts leftover after the extraction process of the K objects. We show, exemplary a detail of the mask M_2 of the rear dice of the second frame of the “spyhole” hologram sequence before (Fig. 11(a)) and after (Fig. 11(b)) the modification described in Eq. (9). The mask used was generated by Alg. 1, adding a 1 px dilation. The combination of $\widehat{M}_1$ and $\widehat{M}_2$ is shown in Fig. 8. With the modified, binary masks $\widehat{M}_p$ at hand, we split $H(t-1)$ up as follows:

$$X(t-1) := \mathcal{G}(H(t-1); g), \quad (11a)$$

$$\forall p \in (\Pi(\{1, \dots, K\}) \cup \{0\}) : X_p(t-1) := \widehat{M}_p(t-1) \odot X(t-1), \quad (11b)$$

$$\forall p \in (\Pi(\{1, \dots, K\}) \cup \{0\}) : S_p(t-1) := \mathcal{G}^{-1}(X_p(t-1); \gamma). \quad (11c)$$

where $\odot$ is the Hadamard product and g, γ is a pair of dual Gabor transform windows, as specified in section 3.2.3.

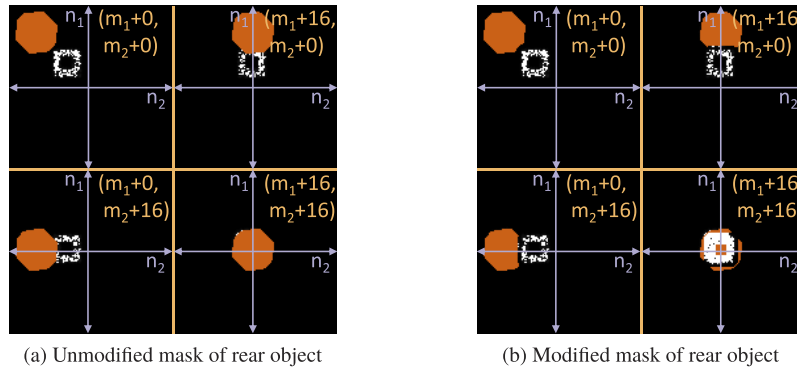


Fig. 11. Using the same phase-space subset as in Fig. 8, (a) shows the unmodified mask M_2 , generated by Alg. 1 and 1 px dilation, of the rear object. (b) shows the mask $\widehat{M}_2$, after subtracting the mask of the front object. $\widehat{M}_2$ is used for the extraction of the rear object.

3.2.6. Motion compensation

In order to compensate for the motion of each of the independently moving objects $k \in \{1, \dots, K\}$ in $S_k(t-1)$, we may apply any global holographic motion compensation method MC to the K

sub-holograms of which each is transformed by a single (global) motion vector $\alpha_k(t-1)$.

$$\forall k \in \{1, \dots, K\} : \tilde{S}_k(t) = MC(S_k(t-1), \alpha_k(t-1)) . \quad (12)$$

Note, since $S_0(t-1)$ contains only static scene parts, we can set $\tilde{S}_0(t) = S_0(t-1)$. Several holographic motion compensation strategies have been proposed in literature [5,6,8] and can be used in Eq. (12) without the need for any adaptations.

3.2.7. Merging of predicted sub-holograms

The merger of the $K+1$ predicted sub-holograms $\tilde{S}(t)$ can be done with proper occlusion handling within the motion-compensated scene as follows:

1. Forward Gabor transform of all predicted sub-holograms, yielding $\tilde{X}_k(t), k \in \{0, 1, \dots, K\}$.
2. Generating masks $M_k(t), k \in \{1, \dots, K\}$ at time instance t for the motion-compensated holograms, see section 3.2.4.
3. Permute the object indices k such that they are sorted from rear to the front and modify the masks to account for the occlusions as discussed in section 3.2.5 and Eq. (9). The required scene information can be obtained from the triangulations $T_k(t)$, which can be precisely obtained from $\alpha_k(t-1)$ and $T_k(t-1)$. We denote the required permutation as $\Lambda(\{1, \dots, K\})$.
4. To merge, start with $\tilde{X}_0(t)$ and for each $k \in \Lambda(\{1, \dots, K\})$ overwrite all coefficients that are contained within the unmodified masks $M_k(t)$, while summing as well all contributions that might be present outside of any mask in any $\tilde{X}_k(t)$ due to artifacts from motion compensation operations.
5. Finally, a single inverse Gabor transform yields $\tilde{H}(t)$.

3.2.8. Limitation and computational complexity of GMMC

The applicability of GMMC to arbitrary scenes is limited by the granularity of the Gabor frame, which is a fundamental property of phase-space analysis. In case that multiple objects with independent motion vectors occupy the same volume of 4D phase-space associated with a Gabor atom, artifacts will arise as the entire cell is being attributed solely to one object. This can be addressed by using additional time-frequency filters on atoms located at those edges in phase-space at the expense of higher computational costs. As the distortion affect only a few Gabor cells, it can easily be accounted for by re-computing the atoms fully in computer-generated holographic videos or it can be encoded as residual in a video compression scheme.

The computational complexity of the GMMC method can be estimated per predicted frame as:

$$R + K \cdot MC + (K+1) \cdot DGT + (K+1) \cdot IDGT + 2K \cdot Alg. 1 \approx K \cdot MC + 2(K+1) \cdot DGT . \quad (13)$$

R describes any additional overhead, such as from the mask manipulations. Since the main work thereby is the rasterization and filling of binary triangles as well as the coordinate projection, R can be neglected when implemented on GPU. The main computational complexity of GMMC stems from the global motion compensation methods “MC” and the Gabor transforms “(I)DGT”. The cost of “MC” varies and has to be considered a fix cost. The computational complexity of IDGT is essentially the same as the complexity of the DGT and depends highly on the chosen window length, required accuracy, amount of active coefficients, the redundancy r , and the size of the hologram L . Often, windows of length <512 suffice and their generation can be done once for all frames. Detailed overviews over the computational complexity of Gabor transforms can be found in [14,17–19]. In brief, one can state that the computational effort of a discrete

Gabor transform is typically slightly above that of a Short-term Fourier transform with the same redundancy, i.e. $O(M^2 N^2 \log(M^2))$. In practice, the (I)DGT of a single hologram with $L = 4096$, $r = 2$, window lengths 512 or 4096, takes 5s and 8s, respectively, with the C implementation provided by the LTFAT toolbox [20], executed on a single core of a Intel Xeon E5-2687W v4. Using Matlab code executed on a single CPU core, the mask generation (with dilation = 2 px) and motion compensation each took on average 5s per object. Despite the use of non-optimized code, the each frame of the “spyhole” sequence could be compensated in $\sim 102s$.

4. Experiments

First, we describe our tested hologram scenes and provide some details on the implementations of BPMC and GMMC. Next, we analyze the quality of the segmenting masks numerically and visually and close by showcasing the motion compensation methods on holographic sequences containing scenes with multiple independently moving objects.

4.1. Experimental settings

4.1.1. Test data

Two computer generated hologram sequences (CGH) “split dices” (Fig. 3) and “spyhole” (Fig. 5), were used in the experiments. The holograms were generated from dense point clouds via PSF splattering Eq. (14), which is simple but also physically highly accurate. It is denoted as

$$H[\xi, \eta] = \sum_{j=1}^J \frac{A_j}{r_j} e^{\frac{2\pi i}{\lambda} r_j + \phi_j}, \quad (14)$$

with point source amplitude A_j and phase ϕ_j . The distance r_j is given by Eq. (1b). The objects were set to be diffusely reflecting by assigning Gaussian random phases ϕ_j to the individual points. A simple occlusion handling was implemented, by modifying the original point clouds ($>4 \times 10^6$ points) through the removal of occluded points with the help of the hidden point removal operator proposed in [21]. The experimental parameters can be found in Table 1.

The scene of the “split dices” hologram sequence contains two dices, one of which depicts 1 eye on the front face (object 1) and one with the 6 eyed face in front (object 2). The motions in between frames are a 60° rotation around z followed by a translation along x for object 1. Object 2 only experiences a translation along $-x$ and y towards the third frame.

Table 1. Hologram generation parameters

	resolution in px	pixel pitch in μm	wavelength in nm	depth in mm
split dices	4096×4096	0.32	635	2.8 – 3.0
spyhole	8192×8192	1	633	1.6 – 3.7

In the scene of the “spyhole” hologram sequence a dice (object 2), placed behind a spyhole (object 1), rotates around the optical axis by 45° per frame. The spyhole stays fixed. In on-axis views the dice is partially occluded. The big motion demonstrates how GMMC can predict information in the center view, which was previously occluded, due to its phase-space segmentation which considers all information present in the hologram.

Back-propagation of the holograms was done using the angular spectrum method and zero-padding in the hologram plane, which avoids aliasing artifacts at any distance. All point clouds are placed such that we can operate in the aliasing free cone, see [22].

4.1.2. Implementation details

For the BPMC method, the segmentation was facilitated by a simple binarization of the back-propagated scene via thresholding hologram amplitudes. The binary masks were dilated and hole

filling techniques were employed to smooth their shapes. Finally, a labeling technique, based on pixel connectivity with the `bwlabel` Matlab command implementing [23], was used.

The GMMC method was implemented in Matlab R2019a. The convex hulls were computed with Matlab's `convhulln` an interface to `qhull` [24]. (I)DGT's were calculated with the LTFAT toolbox [20] using Gaussian windows with equal space-frequency resolution. All remaining parameters were chosen as stated in section 3.2.3 or left at their default values. The triangulations were obtained from initial point cloud models of the objects via application of `convhulln`. The spyhole was explicitly parametrized. Alternative schemes such as forming the triangulation of an enclosing cube are possible. The triangulation super-sampling factor h in phase-space was M per triangle edge. The object ordering in the split and merger steps was determined by the evaluation of the mean depth of the corresponding objects. The global motion compensation proposed in [5] was used and the DGT redundancy and mask dilation were set to 2 per dimension and 3 for the spatial and frequency domains, respectively unless stated otherwise.

4.2. Hologram segmentation

To verify the quality of the object masks employed in BPMC and GMMC, we treat all objects in the scene as a joint object and measure the error introduced through masking. For this, we compare a hologram H with a version H' containing only the information retained by the union of all object masks M_k . That is for BPMC, we back-propagate H to obtain O , and propagate anything contained in the collection of the spatial masks M_k , to obtain H'_{BPMC} . For GMMC, we apply the appropriate union of masks M_k to H in the Gabor domain and obtain H'_{GMMC} .

$$H'_{\text{BPMC}} := BP^{-1} \left(\sum_{k=1}^K M_k \odot BP(H) \right) \text{ and } H'_{\text{GMMC}} := \mathcal{G}^{-1} \left(\sum_{k=1}^K M_k \odot \mathcal{G}(H; g_1, g_2); \gamma_1, \gamma_2 \right). \quad (15)$$

We then evaluated the normalized mean-square error (NMSE) in percent as

$$NMSE(H, H') := 100 \frac{\|H - H'\|_2}{\|H\|_2}. \quad (16)$$

4.2.1. BPMC: spatial mask quality evaluation

For BPMC, we first study on the example of the “split dices” hologram sequence the mask quality quantitatively as a function of the thresholding parameter $q \in (0, 1)$, which is used for binarization of the hologram in step 1 of the mask calculation. The results are summarized in Fig. 12(a). Thereby q times the maximal amplitude of O is used as the threshold. In general, q should be chosen as small as possible such that masks do cover the most information possible within O for compensation, while not overlapping. However, if chosen too small no segmentation will be possible anymore. We implemented the search for optimal values of q as a binary search, which did yield $q = 0.7, 0.6, 0.9\%$ and NMSE values of 1.7, 1.7, 2.0% for frames 1 – 3, respectively.

Visually, we can verify the good mask quality for an optimally chosen q in the phase-space by comparing the phase-space footprint in the hologram plane of the entire scene in frame 1, *i.e.* of $H(1)$ in Fig. 12(b), with the successfully extracted object 1, $BP^{-1}(S_1(1))$ in Fig. 12(c).

4.2.2. GMMC: Gabor mask quality evaluation

GMMC utilizes Gabor masks whose quality we study first quantitatively as a function of the size of the space-frequency mask dilations (given in Gabor atom indices) and as a function of redundancy r of the Gabor system. The NMSE results are reported in Table 2. The values evaluated for a mask covering the entire scene, are an approximation of the per-object mask qualities, which would be depend on the specific scene geometries (amount of occlusion etc.).

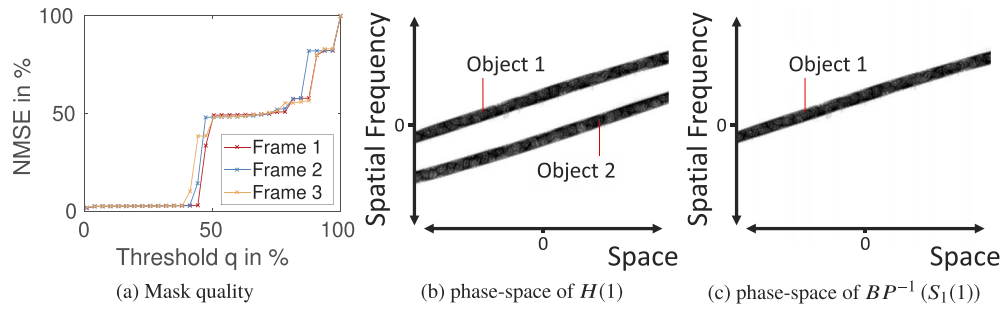


Fig. 12. Quality of joint object masks created via BPMC measured with NMSE and evaluated as a function of the binarization threshold q , shows in (a) that $q \leq 35\%$ provides good mask qualities. (b) and (c) show the phase-space of the entire scene of frame 1 of the “split dices” and of the extracted object 1 after re-propagation to the hologram plane, respectively.

Table 2. NMSE in % after Gabor mask extraction of all objects for “spyhole”, “split dices” (bold).

	Dilation 0 px		Dilation 1 px		Dilation 2 px		Dilation 3 px		Dilation 4 px	
Redundancy 2	24.13	58.19	2.33	16.70	0.07	0.21	<0.01	0.02	<0.01	0.02
Redundancy 4	26.19	39.85	4.96	4.89	0.22	0.37	0.01	0.09	<0.01	0.01
Redundancy 8	17.03	43.43	3.42	9.63	0.33	0.86	0.16	0.27	0.01	0.01

The evaluation was performed on the apodized frames 2 of the “spyhole” and “split dices” hologram sequences to avoid artifacts from periodic boundary conditions in the (1)DGT.

The NMSE for $r = 2$ without any dilation stems from space-frequency discretization errors of the binary masks in combination with the finite spatial and frequency resolution of the Gabor grid, as described in section 3.2.4.1. This error can be mitigated by dilating the masks, resulting in a rapid decline even with only minimal dilation. The additional dilation should be kept as small as possible (wrt. N, M), in the case of multiple objects in the scene, to avoid bleeding of the individual object masks into each other. We find a value of $2\,px$ results in near-lossless masking.

Alternatively, the NMSE can be lowered slightly by increasing the redundancy r of the Gabor system by multiples of 4 at the cost of increased computational complexity when the dilation is small. Increasing r by 4 allows doubling the number of translations N and modulations M henceforth halving the respective resolutions. If only either N or M is doubled, depending on the hologram type, its resolution, and the dilation, the errors might increase due to the discrete nature of the Gabor grid. For dilations $\geq 2\,px$, an increase in redundancy leads even to marginally worse masks because the number of Gabor atoms increases as r^2 for 2D signals and therefore the error caused by any signal-mask mismatch is blown up by the same factor.

Notable is also the influence of resolution on the Gabor mask quality. A doubling in hologram resolution (“spyhole”: $8192\,px$, “split dices”: $4096\,px$) results in general in halving the NMSE.

Next, we verify the mask qualities for ($r = 2$ and $2\,px$ dilation) visually by investigating the segmented phase-space footprints of frame 2 of the “spyhole” sequence. The footprint of the entire scene is shown in Fig. 4(c). The hologram segmentation achieved with occlusion handling and Gabor masks is, for reference, compared to the poor segmentation achieved with spatial masks in Fig. 13. A segmentation of the “spyhole” holograms is not possible with spatial masks as there exists no joint focal distance and occlusions render purely spatial masking insufficient. Instead, in Fig. 13(b), we see parts of object 1, the spyhole, are still present as they are merely clipped during extraction of object 2. A back-propagation to central a focal distance, as used for the spatial segmentation, looks visually similar to Fig. 4(c) and the detected mask for object

2 coincided with the darker, central region. In contrast, we see in Fig. 13(d), which shows the segmented dice sub-hologram, that the Gabor masks are accurate enough to model the fact that only the rims of the spyhole occlude the dice — visible by the two bright lines crossing the phase-space footprint of the dice. In section 4.3.2, we will present the reconstructions from the individual GMMC and BPMC segmented sub-holograms.

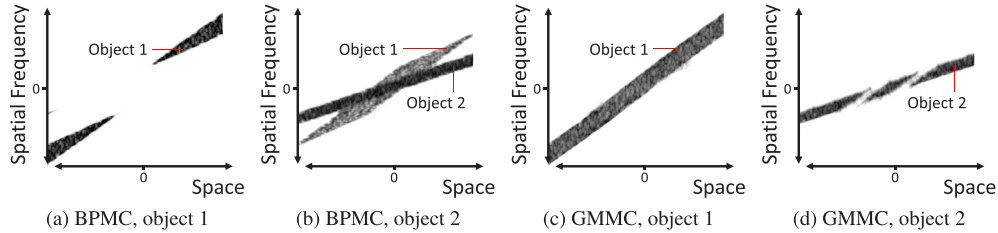


Fig. 13. (a) and (b) shows the phase-space footprint after a poor extraction of objects 1, 2 of frame 2 of the “spyhole” hologram sequence via BPMC. (c) and (d) show the phase-space footprints of the same objects extracted with GMMC. Note, in (d) only the rims of the spyhole occlude the dice in this part of phase-space.

4.3. Multi-object motion compensation

4.3.1. “split dices” hologram sequence

The example of the “split dices” hologram sequence demonstrates that the correct prediction of frames 2 and 3 using frames 1 and 2, respectively, is possible with GMMC as well as BPMC. The ground truth reconstructions and the spatial BPMC masks are depicted in Fig. 3. The compensated frames 2 and 3 as well as the errors in the hologram plane, relative in magnitude to the ground truth, are depicted in Fig. 14. The optimal threshold parameters q for BPMC were

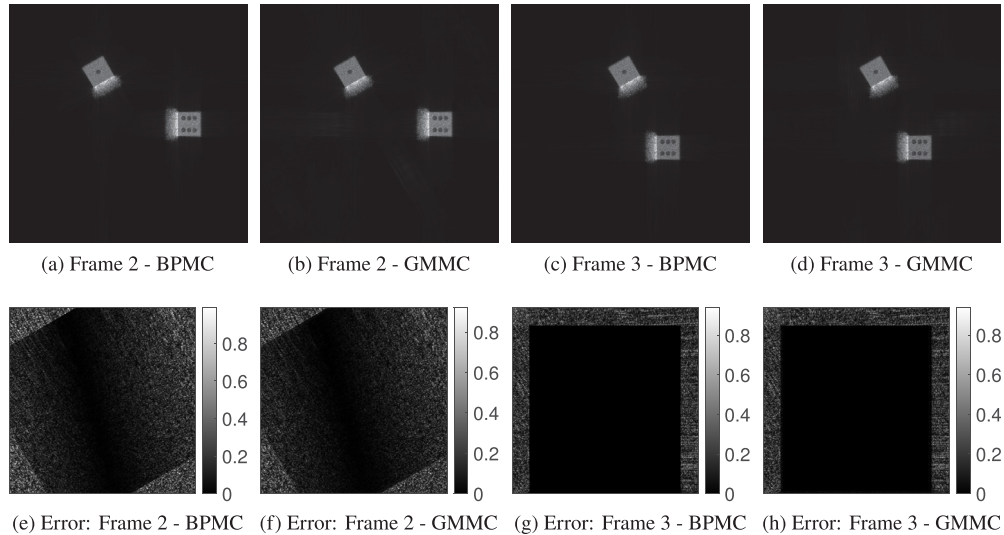


Fig. 14. Top: Reconstructions of approximate frames 2,3 are obtained by motion-compensation via BPMC and GMMC, given frames 1, 2, respectively. Bottom: The corresponding errors in the hologram plane are shown relative to the maximal magnitude of the ground truth. They are predominantly due to genuinely missing information.

chosen per frame. Despite that the spatial masks of BPMC are suitable in this case, the Gabor mask qualities of GMMC are much better for $\geq 2 \text{ px}$ in the apodized case. No visual artifacts can be observed with either method irrespective of the apodization.

Due to the perfectly possible spatial segmentation of the hologram after propagation, the dominant errors are caused by a genuine lack of information. This is visible in frame 2 by the missing corners of the predicted sub-hologram $\tilde{S}_1(t=1)$. In frame 3 the left outer edge is missing from $\tilde{S}_1(t=2)$ after compensation and the top and right edges are missing from $\tilde{S}_2(t=2)$.

4.3.2. “Spyhole” hologram sequence

Next, we show how only GMMC can be used to compensate three non-apodized frames of the “spyhole” scene, see Fig. 5. Figure 15(a)–15(b) show reconstructions of the Gabor mask segmented hologram frames 2 and 3, which have been obtained from frames 1 and 2,

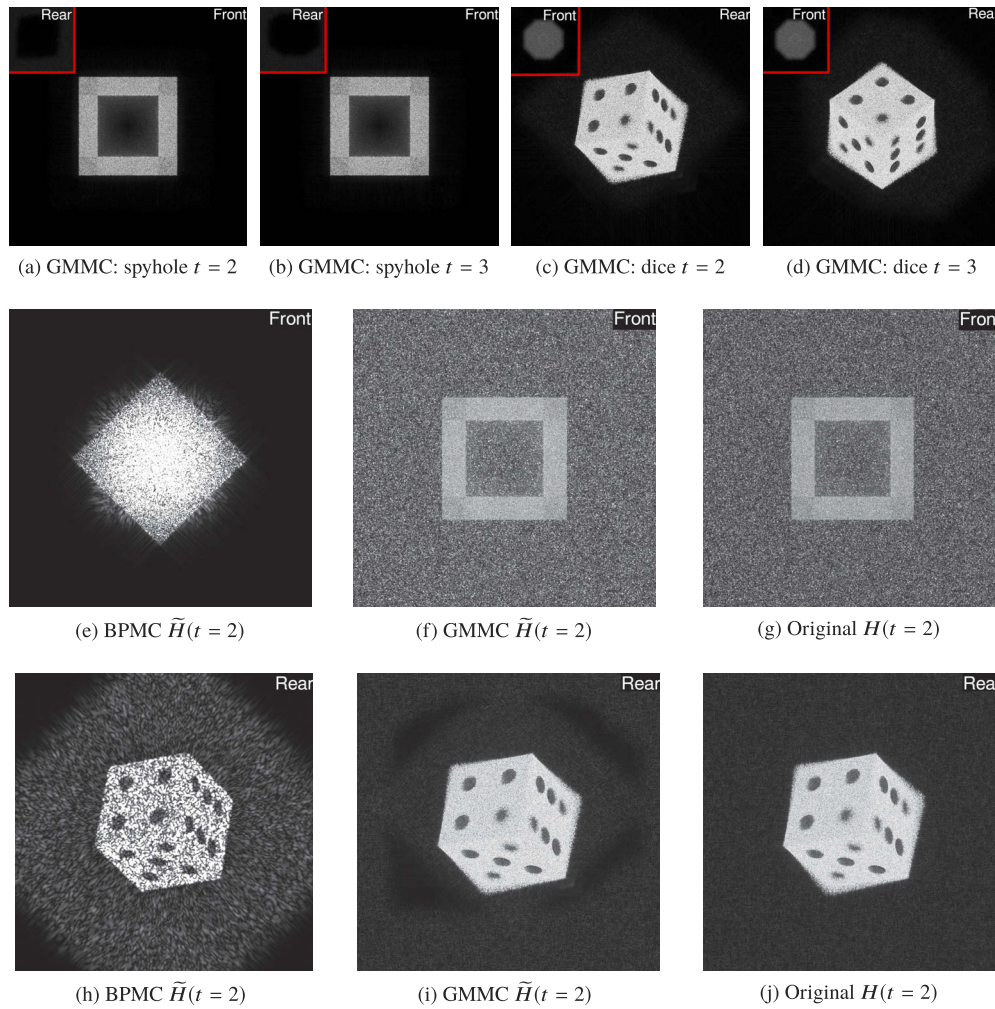


Fig. 15. (a)–(d): Reconstructions of the sub-holograms $\tilde{S}_{\{1,2\}}$ of the “spyhole” holograms — segmented and motion-compensated with GMMC. (e)–(j): Reconstructions of the merged final prediction $\tilde{H}$, using the BPMC, GMMC, and the ground truth H are shown side by side for $t=2$. As expected BPMC fails. The individual front and rear reconstructions are shown magnified.

respectively. Reconstructions labeled “Front” are focused at the front of the spyhole, whereas “Rear” corresponds to the central focal plane of the dice. The motion of the point cloud underlying the ground truth frame is shown in Fig. 5(a)–5(c). As seen only one object is visible per sub-hologram. The space-frequency segmentation is then leveraged to achieve a high-quality motion compensation of the moving and partially occluded dice object, see Fig. 15(f), 15(i) versus the original in Fig. 15(g), 15(j). We provide [Visualization 1](#) in the supplemental material for a clear side-by-side comparison of GMMC prediction and ground truth for multiple depths and viewing angles.

For BPMC (Fig. 15(e), 15(h)) three errors are noticeable. First, because the segmentation is incomplete both the spyhole and the dice are transformed. Second, the spatial mask acts as a limiting numerical aperture for the rear dice object which exhibits a lowered angular resolution — visible by the larger speckle grains in both reconstructions and the smaller spread of the dice in the front reconstruction (Fig. 15(e)) due to the higher frequencies being clipped by the aperture. Third, a bright fog surrounding the dice is present. It is caused by discontinuities introduced into the diffraction pattern of the spyhole upon the merger of the wrongfully motion-compensated parts of the spyhole with its stationary rest. Due to the last two artifacts of BPMC, even simple global motion compensation would be superior, whilst only GMMC produces correct predictions.

5. Conclusion and future work

We proposed a novel method called GMMC to compensate for the motions of multiple independently moving objects in holographic video sequences. The proposed method can handle an arbitrary number of independently moving and mutually occluding objects. GMMC leverages a newly introduced Gabor mask-based hologram segmentation scheme of objects in the space-frequency domain. We compared GMMC against BPMC, which is a simpler reference method for the same task. BPMC relies solely on spatial hologram segmentation and is thereby similar to segmentation schemes used in digital holographic microscopy. BPMC may only be used whenever there exists a focal plane which brings all objects in focus so that they become spatially separable through natural image segmentation schemes applied to the hologram amplitude.

We demonstrated both motion compensation methods for holographic videos containing multiple independently moving objects. Both techniques can be used either for more efficient CGH, as proposed in [3,4], or holographic video compression, *e.g.* [5,6]. With BPMC high mask qualities with $\leq 2\%$ NMSE, of the overall signal missing from the mask, were demonstrated for spatially separable scenes. Furthermore, GMMC successfully motion-compensated a scene with partial occlusions and a look-through object. High-quality Gabor masks with an NMSE of only 0.01% are achievable. Future work may express the motion compensation methods [5,6] in Gabor space instead of the spatial domain, thereby reducing the computational complexity of GMMC by eliminating the current need of one DGT and one IDGT per object. Also, GMMC may be adapted to enable compensation in scenes with non-uniform lighting or reflections through the use of additional scene information.

Funding

H2020 European Research Council (617779); Fonds Wetenschappelijk Onderzoek (12ZQ220N).

Disclosures

The authors declare no conflicts of interest.

References

1. J. Underhill, “In conversation with CyArk: digital heritage in the 21st century,” *Int. J. for Digit. Art Hist.* (2018).

2. T. Laude, Y. G. Adhisantoso, J. Voges, M. Munderloh, and J. Ostermann, "A comparison of JEM and AV1 with HEVC: Coding tools, coding efficiency and complexity," in *2018 Picture Coding Symposium (PCS)*, (2018), pp. 36–40.
3. A. Symeonidou, R. M. Kizhakkumkara, T. Birnbaum, and P. Schelkens, "Efficient holographic video generation based on rotational transformation of wavefields," *Opt. Express* **27**(26), 37383–37399 (2019).
4. H.-K. Cao and E.-S. Kim, "Faster generation of holographic videos of objects moving in space using a spherical hologram-based 3-D rotational motion compensation scheme," *Opt. Express* **27**(20), 29139–29157 (2019).
5. D. Blinder, C. Schretter, and P. Schelkens, "Global motion compensation for compressing holographic videos," *Opt. Express* **26**(20), 25524–25533 (2018).
6. R. K. Muhamad, D. Blinder, A. Symeonidou, T. Birnbaum, O. Watanabe, C. Schretter, and P. Schelkens, "Exact global motion compensation for holographic video compression," *Appl. Opt.* **58**(34), G204–217 (2019).
7. K. Matsushima, "Formulation of the rotational transformation of wave fields and their application to digital holography," *Appl. Opt.* **47**(19), D110–D116 (2008).
8. H.-K. Cao, S.-F. Lin, and E.-S. Kim, "Accelerated generation of holographic videos of 3-D objects in rotational motion using a curved hologram-based rotational-motion compensation method," *Opt. Express* **26**(16), 21279–21300 (2018).
9. M. D. Panah and B. Javidi, "Segmentation of 3D holographic images using bivariate jointly distributed region snake," *Opt. Express* **14**(12), 5143–5153 (2006).
10. A. E. Mallahi, C. Minetti, and F. Dubois, "Automated three-dimensional detection and classification of living organisms using digital holographic microscopy with partial spatial coherent source: application to the monitoring of drinking water resources," *Appl. Opt.* **52**(1), A68–A80 (2013).
11. A. Gilles, P. Gioia, R. Cozot, and L. Morin, "Hybrid approach for fast occlusion processing in computer-generated hologram calculation," *Appl. Opt.* **55**(20), 5459–5470 (2016).
12. J. W. Goodman, *Introduction to Fourier Optics* (Roberts & Company, 2004).
13. T. Birnbaum, A. Ahar, D. Blinder, C. Schretter, T. Kozacki, and P. Schelkens, "Wave atoms for digital hologram compression," *Appl. Opt.* **58**(22), 6193–6203 (2019).
14. P. L. Søndergaard, "Efficient algorithms for the discrete Gabor transform with a long FIR window," *J. Fourier Anal. Appl.* **18**(3), 456–470 (2012).
15. K. Gröchenig, *Foundations of Time-frequency Analysis* (Birkhäuser, 2001).
16. A. E. Rhammad, P. Gioia, A. Gilles, M. Cagnazzo, and B. Pesquet-Popescu, "View-dependent compression of digital hologram based on matching pursuit," *Proc. SPIE* **10679**, 106790L (2018).
17. T. T. Chinen and T. R. Reed, "A performance analysis of fast Gabor transform methods," *Graph. Model. Image Process.* **59**(3), 117–127 (1997).
18. S. Qiu, F. Zhou, and P. E. Crandall, "Discrete gabor transforms with complexity $O(N \log(N))$," *Signal Process.* **77**(2), 159–170 (1999).
19. P. L. Søndergaard, "An efficient algorithm for the discrete Gabor transform using full length windows," in *SAMPTA 09*, L. Fesquet and B. Torrèsani, eds. (2009).
20. Z. Průša, P. L. Søndergaard, N. Holighaus, C. Wiesmeyr, and P. Balazs, "The Large Time-Frequency Analysis Toolbox 2.0," in *Sound, Music, and Motion*, (Springer International Publishing, 2014), LNCS, pp. 419–442.
21. S. Katz, A. Tal, and R. Basri, "Direct visibility of point sets," *ACM Trans. Graph.* **26**(3), 24 (2007).
22. D. Blinder, A. Ahar, S. Bettens, T. Birnbaum, A. Symeonidou, H. Ottevaere, C. Schretter, and P. Schelkens, "Signal processing challenges for digital holographic video display systems," *Sig. Process. Image Com.* **70**, 114–130 (2019).
23. R. M. Haralick and L. G. Shapiro, *Computer and Robot Vision*, vol. 1 (Addison-Wesley Reading, 1992).
24. C. B. Barber, D. P. Dobkin, D. P. Dobkin, and H. Huhdanpaa, "The quickhull algorithm for convex hulls," *ACM Trans. Math. Softw.* **22**(4), 469–483 (1996).