

Efficient numerical Fresnel diffraction with Gabor frames

DAVID BLINDER,^{1,2,3,*}  TOBIAS BIRNBAUM,⁴  AND PETER SCHELKENS^{1,2} 

¹Department of Electronics and Informatics (ETRO), Vrije Universiteit Brussel (VUB), 1050 Brussel, Belgium

²IMEC, 3001 Leuven, Belgium

³Graduate School of Engineering, Chiba University, Chiba 263-8522, Japan

⁴Swave BV, 3001 Leuven, Belgium

*Corresponding author: david.blinder@vub.be

Received 21 May 2024; revised 29 October 2024; accepted 18 November 2024; posted 18 November 2024 (Doc. ID 530818); published 17 January 2025

Numerical Fresnel diffraction is broadly used in optics and holography in particular. So far, it has been implemented using convolutional approaches, spatial convolutions, or the fast Fourier transform. We propose a new way, to our knowledge, of computing Fresnel diffraction using Gabor frames and chirplets. Contrary to previous techniques, the algorithm has linear-time complexity, does not exhibit aliasing, does not need zero padding, has no constraints on changing shift/resolution/pixel pitch between source and destination planes, and works at any propagation distance. We provide theoretical and numerical analyses, detail the algorithm, and report simulation results with an accelerated GPU implementation. This algorithm may serve as a basis for more flexible, faster, and memory-efficient computer-generated holography methods. © 2025 Chinese Laser Press

<https://doi.org/10.1364/PRJ.530818>

1. INTRODUCTION

The numerical implementation of Fresnel diffraction is widely used in optics [1], especially in holography [2,3]. It can efficiently calculate how the wavefield propagates over space, forming the backbone of many applications in computer-generated holography [4] and holographic microscopy [5].

Currently, numerical Fresnel diffraction can be said to come in two main forms. The spatial domain form computes Fresnel diffraction directly as a spatial convolution. The first method is highly flexible and efficient for small distances, used for techniques such as lookup table methods [6,7], layer-based methods [8], and methods related to the wavefront recording plane method [9,10]. However, it does not scale well for larger propagation distances and may exhibit some artifacts as a result of the finite-sized zone plate.

The other main form uses the frequency domain. This computes diffraction as a convolution via Fourier space, which is more efficient than the spatial domain approach for larger distances. However, it is inherently less flexible, often only working for certain propagation distance ranges, needs zero padding, can result in aliasing for some distances, may need multiple fast Fourier transforms (FFTs), and requires extensions whenever there are differences in resolution, pixel pitch, or lateral position between the source and destination planes. That is why they come in many variants, such as shifted Fresnel diffraction [11] or aliasing-reduced Fresnel diffraction [12]. Recent approaches also support more flexible angular spectrum methods, including efficient

off-axis diffraction algorithms [13,14], scalable angular spectrum models [15], and the homeomorphic Fourier transform [16], but do not resolve all of the aforementioned limitations.

We propose a novel approach for computing Fresnel diffraction using Gabor frames. Because the point-spread function (PSF) is sparse in Gabor space, we can achieve linear computational complexity independently of propagation distance for any signal [17]. The transfer function exhibits both properties from spatial and frequency domain methods. Because of the analytical derivations, the method has inherently no limitations w.r.t. the parameters of the source and destination planes; any combination of pixel pitches, resolutions, and propagation distances is supported. The method requires no zero-padding. Because of its phase-space-based modeling, it is more flexible for integration with computer-generated holography (CGH) algorithms relying on, e.g., ray tracing [18] or ray-wavefront conversion [19].

This summarizes our specific contributions:

- we derive an analytical model for Fresnel diffraction with chirplets and Gabor frames;
- we utilize this to devise an algorithm for calculating accurate discrete Fresnel diffraction with freely chosen parameters;
- we propose an efficient approximation for high-speed algorithm evaluation, leading to a linear complexity algorithm;
- the method is implemented on GPU with C++/CUDA and compared to conventional algorithms;

• multiple experiments are performed, analyzing the PSF shape and measuring the computation speed and numerical accuracy.

The remainder of this paper is organized as follows. In Section 2, we introduce the theory for numerical diffraction, phase space, and frames and derive an analytical model using chirplets for Fresnel diffraction with Gabor frames. Section 3 details the proposed algorithm and details an efficient approximation for fast implementations. The algorithm is evaluated in Section 4, covering the PSF analysis, comparing the accuracy with different Fresnel diffraction algorithms, and for different parameters. Finally, we conclude and discuss future work in Section 5.

2. THEORY

A. Fresnel Diffraction

The Fresnel diffraction operator expresses the propagation of a wavefield from a source plane $z = z_1$ to a parallel destination plane $z = z_2$. The continuous Fresnel integral is given by [1]

$$U(x, y; z_2) = \frac{e^{ikd}}{ikd} \iint_{-\infty}^{+\infty} U(\xi, \eta; z_1) \cdot \exp\left(\frac{ik}{2d}((x - \xi)^2 + (y - \eta)^2)\right) d\xi d\eta, \quad (1)$$

where i is the imaginary unit, $d = z_2 - z_1$, $k = \frac{2\pi}{\lambda}$ is the wave-number with wavelength λ , and U is the complex-valued wavefield. Note that the Fresnel diffraction operator is separable along x and y . This means we can only consider the operation along x , giving us

$$U(x; z_2) = \int_{-\infty}^{+\infty} U(\xi; z_1) \exp\left(\frac{ik}{2d}(x - \xi)^2\right) d\xi = \int_{-\infty}^{+\infty} U(\xi; z_1) \mathcal{K}(x - \xi) d\xi, \quad (2)$$

where $\mathcal{K}(\cdot)$ is the Fresnel convolution kernel. We ignore the leading constant for simplicity, which does not affect our calculations.

1. Fresnel Diffraction in the Spatial Domain

This integral can be computed numerically in the spatial domain by means of a spatial convolution. However, one should not discretize the integral expression in a straightforward way, since this would lead to aliasing. $\mathcal{K}(\cdot)$ has a well-defined unique instantaneous frequency ν in every point (provided the assumption of stationary instantaneous phase [20]), given by

$$\nu(x) = \frac{1}{2\pi} \frac{\partial}{\partial x} \angle(\mathcal{K}(x)) = \frac{x}{2\lambda d}, \quad (3)$$

where $\angle(\cdot)$ is the phase function (or angle, “complex” argument). This instantaneous frequency may exceed the Nyquist rate of the sampled signal, especially for short propagation distances d . This can be addressed by only evaluating $\mathcal{K}(\cdot)$ in the region where $|\nu(x)| \leq \frac{1}{2p}$, i.e., by staying below the Nyquist limit, and setting it to zero otherwise. Note that p is the pixel pitch of the hologram. See Ref. [21] for a more general discussion on the shape of the masking region. Formally, we can use

$$U(x; z_2) = \int_{-\infty}^{+\infty} U(\xi; z_1) \Pi\left(\frac{2p}{\lambda d}(x - \xi)\right) \mathcal{K}(x - \xi) d\xi = U(x; z_1) * \Pi\left(\frac{2px}{\lambda d}\right) \mathcal{K}(x), \quad (4)$$

where $\Pi(\cdot)$ is the rectangular function and $*$ is the convolution operator. The spatial convolution algorithm has an $\mathcal{O}(n^2)$ complexity and is thus generally unsuitable for efficiently modeling diffraction. Although the complexity is generally high, the PSF support will be limited for short distances d , leading to fast run times that can surpass frequency domain algorithms. This has been leveraged, e.g., in Ref. [9]. Conversely, a very small d will result in a PSF with a tiny number of samples, leading to larger errors. In addition, the sharp transition to zero will cause ringing artifacts.

2. Fresnel Diffraction in the Frequency Domain

The second main class of techniques computes the Fresnel transform in the frequency domain with the FFT. For this, the Fresnel kernel is expressed analytically in Fourier space as

$$\hat{\mathcal{K}}(\xi) = \mathcal{F}\{\mathcal{K}(x)\} = \sqrt{\frac{i\pi d}{k}} \exp\left(\frac{id\xi^2}{2k}\right) \quad (5)$$

using the Fourier transform $\mathcal{F}\{\cdot\}$ so that we can compute the convolution as

$$U(x; z_2) = \mathcal{F}^{-1}\{\mathcal{F}\{U(x; z_1)\} \hat{\mathcal{K}}(\xi)\}. \quad (6)$$

The main advantage of this approach is the $\mathcal{O}(n \log n)$ computational complexity, independently of the parameters of $\mathcal{K}(\cdot)$. This form is mainly used for long-distance propagation with pixel pitch preservation, i.e., the pixel pitches of the source and destination planes match. It is one of the most widely used approaches for computing Fresnel diffraction. Other variants exist using one or three FFTs, where the pixel pitch may change as a function of d [22].

However, several considerations must be made when computing the discrete version of the Fresnel diffraction operator in the frequency domain. Aliasing can occur for large distances d and zero padding is required. The frequency domain approach is less flexible than the spatial domain convolution whenever there are differences in the sampling pitch p and the number of samples between the source and destination planes or whenever there is a lateral offset. To this end, several extensions have been proposed to (partially) address these limitations, as mentioned in the introduction [11,12,23]. Multiple shifted propagations can be used whenever the destination plane resolution is a whole multiple of the source plane, but this cannot efficiently calculate arbitrary resolution changes.

We propose a third way for encoding Fresnel diffraction, relying on a joint space-frequency (or time-frequency) representation, also known as the Wigner transform or phase space.

B. Phase Space Modeling of Fresnel Diffraction

One important way of finding a suitable signal representation is the notion of sparsity; this is a concept in signal processing, where a small number of coefficients can accurately represent some class of signals in a well-chosen transform space. This property is useful for many applications, such as data

compression, filtering, compressed sensing, and accelerating calculations. Classical signal analysis using the Fourier transform for analyzing and representing many kinds of generally non-stationary holograms tends to be suboptimal [24]. This fact is apparent when looking at the PSF in Fresnel diffraction, cf. Fig. 1(a): although the PSF can be dense simultaneously in both spatial x and frequency ν domains, it is sparse in phase space.

We can generally model many paraxial diffraction operations as linear canonical transforms [25]. Many of these operations can be expressed as linear phase space transforms, including the Fourier transform (as a 90° rotation), fractional Fourier transform (arbitrary rotation), magnification (area-preserving coordinate scaling), and, in particular, Fresnel diffraction (shearing). For a mathematically rigorous description of, among others, reversible phase space transforms, we refer to Refs. [26–29].

The shearing in phase space [cf. Fig. 1(b)] shows how Fresnel diffraction is a very structured convolution operation amenable to efficient modeling and processing beyond a pure Fourier representation. One common reversible phase space representation is the Wigner transform. As it is quadratic with respect to the signal, it introduces cross-terms for any signal that does not form a single straight line in phase space [27,28,30]—such as holograms, which typically consist of a superposition of millions of PSFs, encoded in phase space as lines or sigmoids [20]. Furthermore, it is often highly overcomplete (typically requiring N^2 coefficients for an N -sample input signal). It is thus not a suitable candidate to leverage the phase space sparsity of holograms.

Therefore, we aim to select a better representation suitable for efficiently and accurately modeling Fresnel diffraction for any combination of propagation distances, lateral offsets, sampling pitches, and wavelengths. This requires a representation satisfying several requirements:

- locality or phase space compactness, to have most of the signal energy concentrated over a small area in phase space, promoting sparsity for arbitrary diffraction conditions;

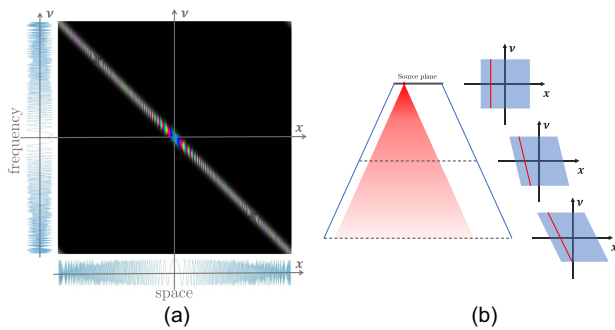


Fig. 1. Phase space models of Fresnel diffraction. (a) Although a PSF can be dense in space (x) and the frequency (ν) domain, it is sparse in phase space; it is color-coded using brightness for magnitude and hue for phase. (b) Side-by-side comparison of how propagating light will progressively shear over phase space. An example point is color-coded red in phase space. (a) Phase space of a PSF. (b) Fresnel diffraction shears phase space.

- efficiently “tiling” phase space, representing any complicated signal effectively with as few elements as possible;
- efficient computation of the forward and inverse transforms;
- being analytically and numerically compatible with Fresnel diffraction, i.e., preserving the aforementioned properties after Fresnel diffraction over any distance;
- linearity with respect to the signal, to avoid cross-terms.

We argue that chirped Gabor frames are good candidates for modeling Fresnel diffraction, which we will introduce in the next section. This is because, locally, any hologram in the Fresnel diffraction regime can be well approximated using chirplets.

C. Chirplets

We define a chirplet as a function composing the exponential function with a quadratic function with complex-valued coefficients. Formally, we define the set of chirplets \mathcal{G} as

$$\mathcal{G} = \{\exp(\alpha t^2 + \beta t + \gamma) \mid \alpha, \beta, \gamma \in \mathbb{C} \wedge \Re(\alpha) < 0\}, \quad (7)$$

where we consider elements where the real part $\Re(\alpha) < 0$ so that they are square-integrable and tending to zero for $t \rightarrow \pm\infty$. They can be viewed as an extension of Gaussians, where the imaginary part of the coefficients $\Im(\alpha)$, $\Im(\beta)$, and $\Im(\gamma)$ are additional degrees of freedom, cf. Table 1.

Chirplets can be viewed as idealized chirped pulses with a Gaussian envelope, cf. Fig. 2.

Because chirplets are an extension of Gaussians, they inherit the optimality of Gaussian beam modes in that they are shape-preserving, have a well-defined Fourier representation, and spread minimally as they propagate. They can be viewed as a wave-like extension of a ray bundle, similar to field tracing [31], which are efficient generalizations for ray tracing in geometric optics for numerical diffraction.

The set of chirplets \mathcal{G} is closed under multiplications and convolutions. Consider two chirplets $u = \exp(at^2 + bt + c)$, $\hat{u} = \exp(\hat{a}t^2 + \hat{b}t + \hat{c})$; we have that $\forall u, \hat{u} \in \mathcal{G}$,

$$u \cdot \hat{u} = \exp((a + \hat{a})t^2 + (b + \hat{b})t + (c + \hat{c})) \in \mathcal{G}, \quad (8)$$

and

Table 1. Relation between Chirplet Parameters and Attributes

	α	β	γ
$\Re(\cdot)$	Width	Translation	Scaling
$\Im(\cdot)$	Chirp	Modulation	Phase shift

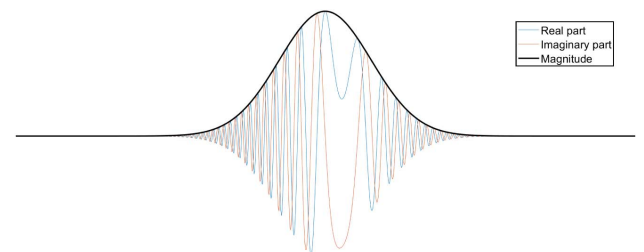


Fig. 2. Chirplet example with a quadratically changing phase.

$$\begin{aligned}
 u * \hat{u} &= \int_{-\infty}^{+\infty} u(\tau) \hat{u}(t - \tau) d\tau \\
 &= \sqrt{\frac{-\pi}{a + \hat{a}}} \exp\left(\frac{a\hat{a}}{a + \hat{a}} t^2 + \frac{a\hat{b} + \hat{a}b}{a + \hat{a}} t - \frac{(b - \hat{b})^2}{4(a + \hat{a})} + c + \hat{c}\right)
 \end{aligned} \tag{9}$$

are also part of \mathcal{G} . In addition, chirplets can be integrated over \mathbb{R} analytically as follows:

$$\forall u \in \mathcal{G}: \int_{-\infty}^{+\infty} u(t) dt = \sqrt{\frac{\pi}{-a}} \exp\left(c - \frac{b^2}{4a}\right). \tag{10}$$

This means we can freely multiply, convolve, or compute inner products with chirplets, making them suitable basis functions.

The Fresnel convolution kernel \mathcal{K} can be viewed as a “degenerate” chirplet since it almost satisfies the condition of belonging to \mathcal{G} , except for having a purely imaginary α , i.e., $\Re(\alpha) = 0$ rather than $\Re(\alpha) < 0$. This poses no problem in practice because multiplying or convolving any proper chirplet with \mathcal{K} will still result in a chirplet. Formally,

$$\forall u \in \mathcal{G}: u \cdot \mathcal{K} \in \mathcal{G} \wedge u * \mathcal{K} \in \mathcal{G}, \tag{11}$$

following immediately from Eqs. (8) and (9), seamlessly integrating Fresnel diffraction with chirplets. For notational simplicity, we will define $\mathcal{K} = \exp(iKx^2)$, where $K = \frac{\pi p^2}{\lambda d}$ is a dimensionless constant closely related to the Fresnel number.

Suppose we would like to express the relationship between the signals in the parallel input and output planes, spaced by a distance d , respectively represented with chirplets. For any $u = \exp(at^2 + bt + c)$ in the input plane and $\hat{u} = \exp(\hat{a}t^2 + \hat{b}t + \hat{c})$ the output plane; we can now directly express the relationship of their coefficients with the inner product

$$\begin{aligned}
 &\langle \hat{u}, u * \mathcal{K} \rangle \\
 &= \int_{-\infty}^{+\infty} \hat{u}(t) \cdot \overline{(u * \mathcal{K})(t)} dt \\
 &= \frac{\pi}{\sqrt{a\hat{a} + K(a + \hat{a})}} \exp\left(-\frac{1}{4} \frac{a\hat{b}^2 + \hat{a}b^2 + (b - \hat{b})^2}{a\hat{a} + K(a + \hat{a})} + c + \hat{c}\right),
 \end{aligned} \tag{12}$$

where the overbar is the complex conjugate operation. In addition to this mapping, one should be able to represent arbitrary signals with a collection of chirplets. This can be achieved with Gabor frames, which provide a mathematical framework for a reversible, linear, and local transform.

D. Gabor Frames

Frames generalize vector space bases to sets that may be linearly dependent, providing a redundant and stable representation of a signal [29,32]. Although an “overcomplete” basis means that the number of transform coefficients will be larger than the number of input samples to ensure reversibility, this allows for more flexibility in designing the transform properties. Formally, a family of vectors $\{\varphi_j\}_{j=1}^M$ in a separable Hilbert space \mathcal{H} is called a frame for \mathcal{H} , if there exist constants $0 < A \leq B < \infty$ such that [32]

$$\forall b \in \mathcal{H}, \quad A \|b\|^2 \leq \sum_{j=1}^M |\langle b, \varphi_j \rangle|^2 \leq B \|b\|^2. \tag{13}$$

For every frame $\{\varphi_j\}_{j=1}^M$, there exists a canonical dual frame $\{\hat{\varphi}_j\}_{j=1}^M$, satisfying

$$\forall b \in \mathcal{H}, \quad b = \sum_{j=1}^M \langle b, \varphi_j \rangle \hat{\varphi}_j = \sum_{j=1}^M \langle b, \hat{\varphi}_j \rangle \varphi_j, \tag{14}$$

namely, serving as an inverse operator.

We will focus on a particular class of frames: stationary Gabor systems. Such a Gabor frame can be represented as a lattice $(m, n) \in \mathbb{Z}^2$ in phase space, expressing a collection of translations T and modulations M of a window g :

$$\{M_{ma} T_{nb} g\}_{m,n \in \mathbb{Z}}, \tag{15}$$

where $a, b > 0$ are fixed, and the operators T_p and M_q operating on some $f \in \mathcal{H}$ are defined as

$$T_p: (T_p f)(x) = f(x - p) \quad \text{and} \quad M_q: (M_q f)(x) = e^{2\pi i q x} f(x). \tag{16}$$

Many windows g are admissible, but they usually are chosen to be symmetric, with a maximum in the middle and tapering away from the center, e.g., rectangular windows, B-spline windows, or Hann windows, among many others. We chose a Gaussian window g to be compatible with chirplets. Only values for a and b where $ab \leq 1$ are admissible. Otherwise, the phase space lattice becomes too “sparse,” making inversion impossible. As the lattice gets denser for smaller values of ab , the dual window \hat{g} for the canonical dual frame gets increasingly compact and regular; see Fig. 3. This causes a trade-off between the number of basis functions and their efficient computability. In practice, oversampling factors of $\approx \frac{4}{3}$ (or $ab = \frac{3}{4} = 0.75$) provide sufficiently regular dual windows, which is superior to the $2\times$ oversampling factor required for the zero-padding in frequency-domain-based Fresnel diffraction. In other words, the memory requirements of the proposed Gabor diffraction algorithm are lower than that of conventional Fresnel diffraction, only needing an $\approx 33\%$ increase of coefficients in the intermediate representation as opposed to the memory doubling required for zero-padding for frequency domain Fresnel diffraction.

Since Gaussian Gabor frame elements are special cases of chirplets [with $\Im(\alpha) = 0$], we have all the ingredients to express Fresnel diffraction directly via the Gabor transform

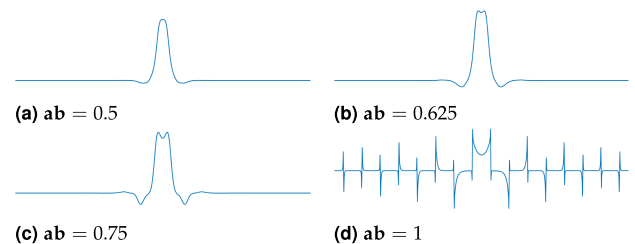


Fig. 3. Examples of canonical dual windows for different values of ab , computed for a discrete signal with $N = 1024$ samples and 64 modulations.

instead of the spatial or frequency domain. The practical methodology will be explained in the next section.

3. METHODOLOGY

A. Implementation

The methodology for Fresnel diffraction of Gabor frames is summarized in Fig. 4. The general approach is similar in spirit to the frequency domain Fresnel transform, but instead of going back and forth with FFTs and applying operations in Fourier space, we use the Gabor transform.

First, we can independently process the rows and columns of the input hologram since the Fresnel transform is separable along x and y . This can also be done in parallel to accelerate the calculations. Every row and column are transformed into an invertible linear combination of chirplets; this is achieved by using the Gabor transform with the dual frame. Then, these chirplets are transformed using the analytical mapping expressed in Eq. (12); how this can be achieved in practice is elaborated in Section 3.B. The resulting Gabor coefficients now describe the hologram in the destination plane. We obtain the output hologram after a final inverse Gabor transform using the dual frame again.

We use the dual frame twice in the forward and inverse Gabor transforms. This is because the analytical mapping is expressed using chirplets with a Gaussian window, requiring initial coefficients obtained with the dual frame rather than the standard frame. This means we can use the same (precomputed) dual window for both transform parts.

The Gabor transform can be computed with a freely selectable number of modulations, independent of the sample count n , making it a transform of linear complexity. What remains to be done is to find a linear-complexity implementation of the analytical chirplet mapping given by Eq. (12), resulting in $\mathcal{O}(n)$ complexity Fresnel diffraction.

B. Chirplet's Region of Influence in Phase Space

Although the mapping would indicate at first glance a relationship exists between all input and output coefficients, this mapping transformation is sparse: the large majority of mappings will be near-zero and thus negligible. Note that this is independent of the input signal, which does not have to be sparse. Suppose that we have $u = \exp(-s(x - \delta)^2 + imx)$ as the input chirplet and $\hat{u} = \exp(-\hat{s}(x - \hat{\delta})^2 + i\hat{m}x)$ as the output chirplet. Namely, the support is encoded by (s, \hat{s}) , the translation by

$(\delta, \hat{\delta})$, and the modulation by (m, \hat{m}) . From Eq. (12), we can see that the inner product is characterized by an exponential function of the chirplet parameters. The magnitude will thus only be affected by the real part of the expression within the exponential, that is,

$$\Re(\log(\hat{u}, u * \mathcal{K})) \propto \mathcal{E}(\hat{\delta}, \hat{m}) := (4K^2(s + \hat{s})^2 + 4s^2\hat{s}^2)^{-1} \cdot (K^2(s + \hat{s})(4s\hat{s}(\delta - \hat{\delta})^2 + (m - \hat{m})^2) + s\hat{s}(4K(m\hat{s} + \hat{m}s)(\delta - \hat{\delta}) + m^2\hat{s} + \hat{m}^2s)), \quad (17)$$

as we can ignore the leading constant since it will remain constant for varying $\hat{\delta}$ or \hat{m} . The function $\mathcal{E}(\hat{\delta}, \hat{m})$ is concave everywhere since we have

$$\frac{d^2}{d\hat{\delta}^2} \mathcal{E} = -\frac{2K^2s\hat{s}(s + \hat{s})}{K^2(s + \hat{s})^2 + s^2\hat{s}^2} < 0 \quad (18)$$

and

$$\frac{d^2}{d\hat{m}^2} \mathcal{E} = -\frac{K^2(s + \hat{s}) + s\hat{s}^2}{2K^2(s + \hat{s})^2 + 2s^2\hat{s}^2} < 0, \quad (19)$$

meaning that the mapping has a unique maximum in phase space and will quickly decay to zero everywhere else. In other words, every input chirplet will only significantly affect a small number of output chirplets, thereby enabling efficient numerical implementations.

The location of this phase space maximum is highly structured across chirplets. This can be deduced from the linear canonical transform model previously discussed in Section 2.B, expressing the Fresnel diffraction as a phase space shear. This principle is illustrated in Fig. 5.

This means one can accurately compute the value of an output Gabor coefficient by only considering a small sheared neighborhood in the appropriate location in phase space corresponding to a few input Gabor coefficients. The size of this $\chi \times \omega$ region-of-interest (ROI) neighborhood can be chosen based on the user's computational complexity versus speed requirements.

Consider the source and destination plane represented by the discrete Gabor transform with M modulations and T translations. Suppose that we want to compute a particular destination Gabor coefficient with coordinates (t_d, m_d) , where $t_d \in \{0, \dots, T - 1\}$ and $m_d \in \{0, \dots, M - 1\}$. Given the

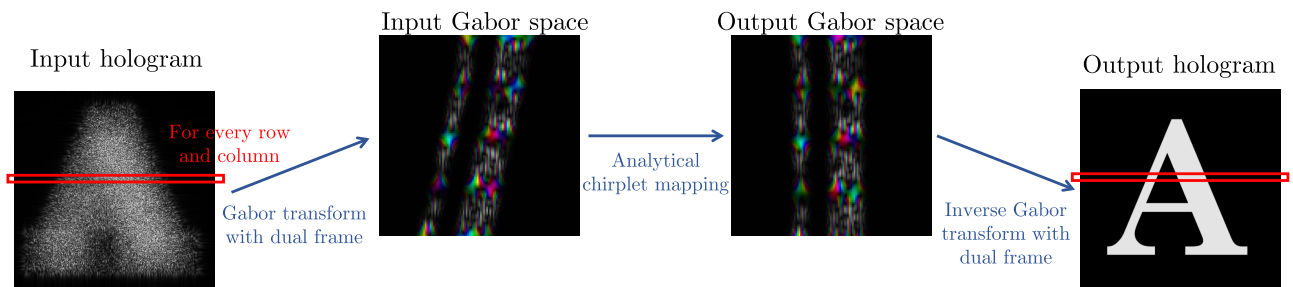


Fig. 4. Summary of the proposed Fresnel transform pipeline. Every row and column can be processed independently thanks to the separability of the Fresnel transform. Using the dual window, we obtain a structured collection of chirplets forming a Gabor frame. These are transformed to obtain the target Gabor coefficients, after which the resulting signal can be retrieved with an inverse Gabor transform. Reprinted from Fig. 20 in Ref. [33].

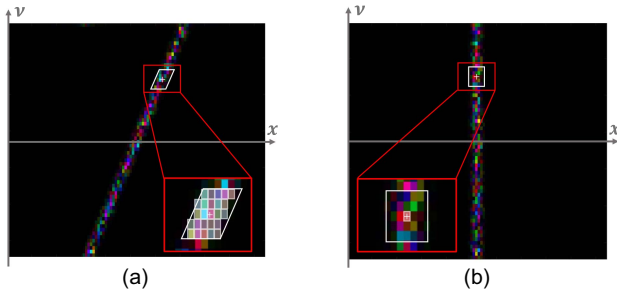


Fig. 5. Approximate Gabor coefficient mapping for fast Fresnel diffraction. The diagrams depict the discrete Gabor coefficients, showing translations along the x -axis and modulations along the ν -axis, color-coded using brightness for magnitude and hue for phase. The example hologram (a) is an idealized Gaussian-modulated random diffuser placed a few cm from the hologram plane, brought into focus in (b). The coefficient in (b) marked by a “+” sign is calculated using only the highlighted five rows of five coefficients in (a), chosen to best align with the parallelogram-shaped region. (a) Source Gabor space. (b) Destination Gabor space.

chosen size $\chi \times \omega$ of the ROI, which source Gabor coefficient array coordinates (t_s, m_s) will have the highest contributions?

We observe that the shear is purely horizontal since the Fresnel transform is an all-pass filter. This indicates that the most relevant source frequencies will be centered around m_d . We let m_s vary in the range of

$$m_s \in \left\{ \left\lfloor m_d - \frac{\omega}{2} \right\rfloor, \dots, \left\lfloor m_d + \frac{\omega}{2} \right\rfloor \right\}, \quad (20)$$

where $\lfloor \cdot \rfloor$ is the floor operator (rounding down). For each row m_s , we have a different spatial displacement matching the phase space shear. The shearing displacement for that row is given by

$$s_s = \frac{\pi(M/2 - m_s)}{KMT}, \quad (21)$$

where T is the number of translations so that we only consider source Gabor coefficient coordinates in that row satisfying

$$t_s \in \left\{ \left\lfloor t_d + s_s - \frac{\chi}{2} \right\rfloor, \dots, \left\lfloor t_d + s_s + \frac{\chi}{2} - 1 \right\rfloor \right\}, \quad (22)$$

where $\lfloor \cdot \rfloor$ is the rounding operator. In this case, we treat coefficients for coordinates outside the source Gabor coefficient array as zero. Experiments measuring the impact of this method are reported in Section 4.D.

4. EXPERIMENTS

We will cover multiple aspects of the newly proposed propagation algorithm for the experiments section. We will look at the properties of the PSF, test the propagation operator on a sample image, and measure the computation speed and accuracy.

A. Analysis of the Point-Spread Function

As discussed in Section 2.A and Ref. [1], the continuous Fresnel kernel $\mathcal{K}(x) = \exp(iKx^2)$ cannot be straightforwardly discretized. Different approaches can be taken depending on

assumptions made about sampling conditions, propagation distance, wavelength, and desired signal properties.

For this experiment, we examine the PSF resulting from a single point at a distance of $d = 25$ mm from a one-dimensional 1024-sample hologram with pixel pitch $p = 6 \mu\text{m}$ and wavelength of $\lambda = 633$ nm. Three different algorithms were tested, cf. Fig. 6: the spatial domain method using Eq. (4), the proposed Gabor domain method, and the frequency domain method using Eq. (6) with zero padding.

The spatial domain method exactly matches $\mathcal{K}(x)$ in the valid (i.e., nonaliased) region but goes sharply to zero, which will cause visual artifacts, especially apparent at edges (see also Section 4.B). The frequency domain method will smoothly drop toward zero in the spatial domain, being defined everywhere over the hologram. However, it will not match $\mathcal{K}(x)$ precisely in the spatial domain, showing highly oscillatory small perturbations, so-called ringing, especially in the low-frequency regions of the signal. The frequency domain method will also become less accurate at long distances due to the required band-pass filtering [12]. The proposed method combines good properties from the spatial and frequency methods. In the middle of the PSF corresponding to low frequencies, the signal exactly matches $\mathcal{K}(x)$ like the spatial method. As we go outwards, there is a transient region similar to the frequency method, but its extent is smaller due to the locality of the Gabor frame; it will drop faster and eventually become identically zero at larger distances, unlike the frequency domain method. Nonetheless, the overall properties of the proposed PSF are closer to the FFT version.

The desired properties of the discrete Fresnel diffraction operator are partially controllable depending on the properties of the Gabor frame, offering simultaneous accurate approximations to the PSF in the center for low frequencies and smooth transients at the edges. This aspect is controllable depending on the chosen space and frequency resolution of the Gabor transform by altering the shape of the Gabor lattice expressed by the number of translations T and modulations M .

B. Propagating Test Wavefronts

For this series of experiments, we compare the relative accuracy of a propagated image for the three Fresnel diffraction methods: the spatial domain, frequency domain, and proposed Gabor domain method. We use the 1024×1024 pixels hologram depicting a bird, whose properties were set to $p = 4 \mu\text{m}$ and $\lambda = 532$ nm. We chose three different propagation distances $d = 5, 10, 20$ mm and compared the PSNR of the reconstructions for the three methods in Table 2. Example reconstructions are shown in Fig. 7.

This confirms that the methods are visually highly similar but not identical due to the different PSF approximations. The frequency domain method and the Gabor method are more similar to each other, especially for shorter propagation distances, where the spatial method has a smaller PSF with fewer samples, leading to larger errors.

Another experiment we performed was to use different settings for the source and destination holograms, demonstrating the flexibility of the proposed algorithm. It can compute holograms where the pixel pitch and hologram resolution are different using off-axis propagation without any modifications or

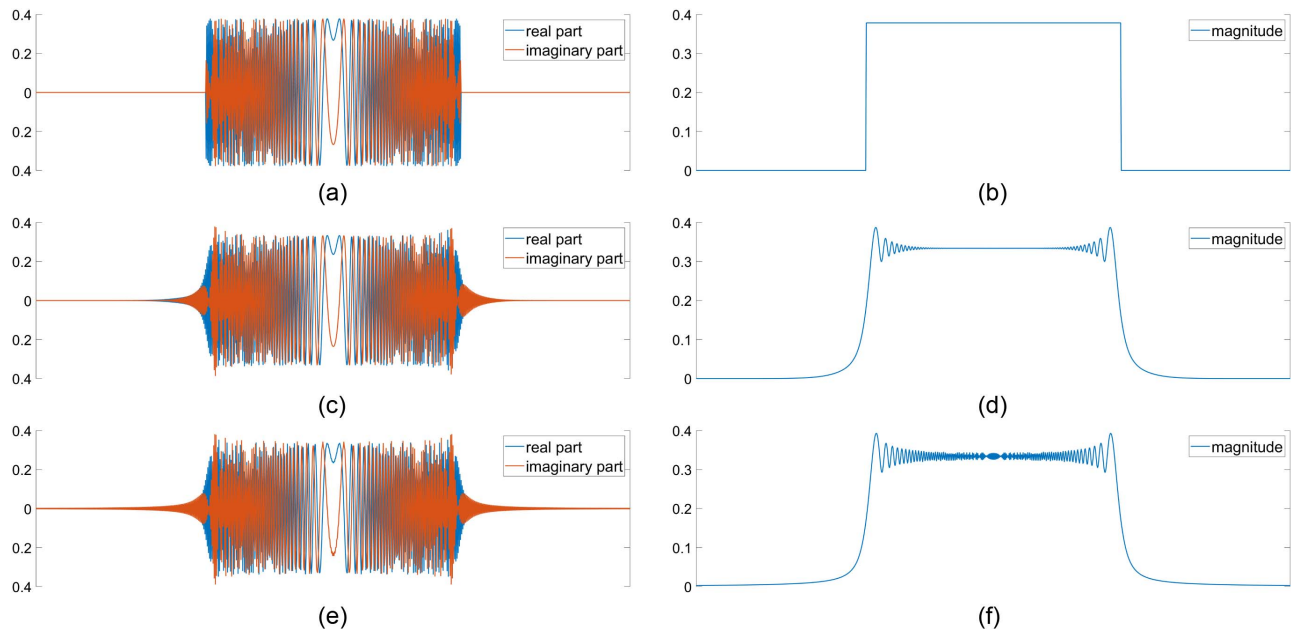


Fig. 6. Plots of the PSF of the discrete Fresnel diffraction, showing the real and imaginary parts and the magnitude. (a), (b) Spatial domain method, (c), (d) proposed Gabor domain method, and (e), (f) frequency domain method. (a) Spatial domain Fresnel transform (real/imaginary). (b) Spatial domain Fresnel transform (magnitude). (c) Gabor domain Fresnel transform (real/imaginary). (d) Gabor domain Fresnel transform (magnitude). (e) Frequency domain Fresnel transform (real/imaginary). (f) Frequency domain Fresnel transform (magnitude).

Table 2. Comparing the PSNR between Different Fresnel Diffraction Methods as a Function of Propagation Distance

d (mm)	PSNR (dB)		
	Spatial versus Gabor	Gabor versus Frequency	Spatial versus Frequency
5	33.6	68.3	33.6
10	40.4	67.3	40.4
20	44.6	63.8	44.6

zero-padding. We used a 1024×1024 pixels source hologram with a pixel pitch of $p_{\text{in}} = 4 \mu\text{m}$ and $\lambda = 532 \text{ nm}$, propagating over a distance of $d = 5 \text{ mm}$ to a destination hologram with a resolution of 1536×1536 pixels and a pixel pitch of

$p_{\text{out}} = 5 \mu\text{m}$, with an additional lateral displacement of $200 \mu\text{m}$ along x . This yielded a PSNR of 38.7 dB. The reconstructions are shown in Fig. 8.

The frequency domain algorithms do not support arbitrary changes in resolution. It is achievable with a spatial domain algorithm; however, it will no longer be a conventional discrete convolution, since the input samples have a different sampling density than the output samples.

This flexibility means that it is possible to compute small parts of the wavefield with any sampling density, allowing for the efficient computation of particular regions of interest, such as specific simulated viewpoints for holographic displays. Since only a subset of the Gabor coefficients of the input hologram is needed in that case, it may be possible to further accelerate the algorithm for view-dependent reconstructions by only computing the relevant coefficients of the input hologram.

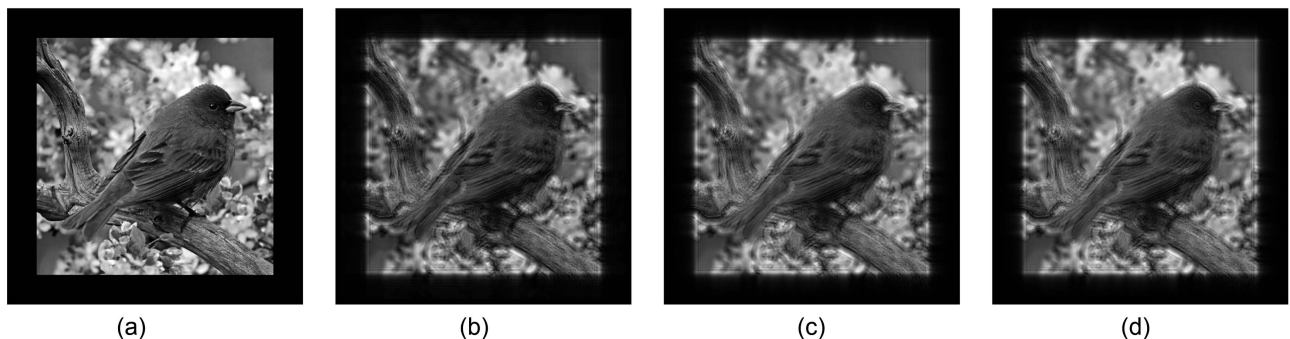


Fig. 7. Side-by-side comparison of an example hologram propagated with different algorithms, at $d = 5 \text{ mm}$. (a) The input hologram, followed by multiple reconstructions, using (b) the spatial domain method, (c) the proposed Gabor domain method, and (d) the frequency domain method. Reprinted from Fig. 21 in Ref. [33]. (a) Input hologram. (b) Output (spatial domain). (c) Output (Gabor domain). (d) Output (frequency domain).

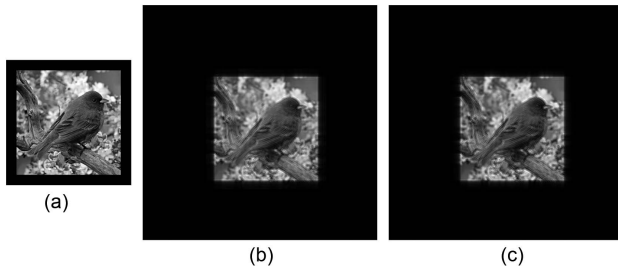


Fig. 8. Example of scaled numerical diffraction with an offset and resolution change. This is done with (b) the spatial domain approach and (c) the proposed Gabor domain method. The figures depict the holograms with correct relative sizes. (a) Source hologram. (b) Destination hologram (spatial). (c) Destination hologram (Gabor).

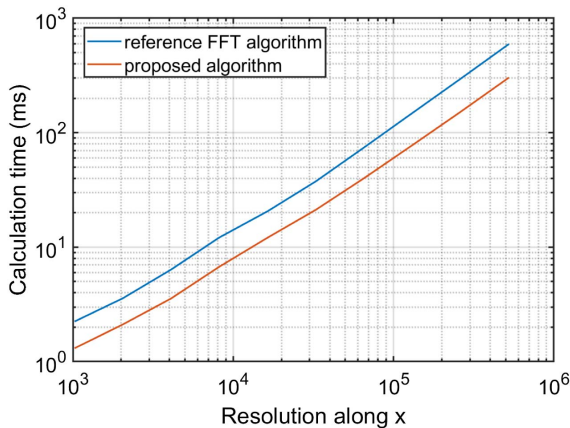


Fig. 9. Log-log plot of calculation time, comparing the reference FFT-based Fresnel diffraction algorithm with the proposed algorithm as the hologram resolution increases.

C. Measuring Calculation Speed

We implemented the proposed algorithm on GPU for the calculation time experiments, comparing it to the frequency domain method utilizing the FFT. We used a machine with an AMD Ryzen Threadripper PRO 5995WX CPU, 256 GB RAM, NVIDIA RTX A6000 GPU, and a Windows 11 OS. The code was implemented in C++20 with CUDA 12.1, enabling CUDA compute capability 8.6 and using 32-bit floating-point precision.

We calculated the Fresnel diffraction operator of an $N \times 512$ pixels hologram, where N was a variable parameter. The properties of the hologram were set to $p = 8 \mu\text{m}$, $\lambda = 633 \text{ nm}$, and $d = 1 \text{ cm}$. We set the ROI window size for the proposed Gabor propagation algorithm to 5×5 coefficients. The results are graphed in Fig. 9.

The parameter N was varied from $2^{10} = 1024$ up to $2^{19} = 524,288$, and the experiment was run 20 times for each value of N , reporting average calculation time. The proposed algorithm is consistently faster, reporting speed gain ratios rising from approximately $1.7\times$ up to $2.0\times$ for increasing N . Although the FFT algorithm has $\mathcal{O}(n \log n)$ complexity, it is

highly optimized and streamlined for GPUs, so exceeding it is a challenge.

The computational advantage of our proposed $\mathcal{O}(n)$ complexity algorithm should be more pronounced in highly memory-bound environments (such as with FPGA or ASIC) thanks to the phase space locality of the algorithm. It can be more efficient when the output resolution is not an integer multiple of the input resolution. Efficient approximations of the complex exponential and more sophisticated GPU implementations could also further extend the gap with FFT-based algorithms.

Depending on the application's precision and speed requirements, one should be careful what specific programming instructions are used for computing the different coefficients [34,35]; this is especially important for lower-precision fixed-point integer representations, which are beneficial for significant further acceleration.

D. Quantifying the Impact of the Phase Space ROI

In this section, we evaluate the accuracy of the version of the approximated method by varying the size of the ROI neighborhood $\chi \times \omega$ detailed in Section 3.B. We do this by computing the PSNR of a random complex-valued signal with $N = 512$, with hologram properties of $p = 6 \mu\text{m}$, $\lambda = 633 \text{ nm}$, and $d = 1 \text{ cm}$. The results were compared with the reference, non-approximated version of the algorithm. The results are reported in Table 3.

These show that it is not useful to increase only χ or ω , as the accuracy will eventually saturate. Rather, it is better to increase both parameters since the magnitude appears to decay radially like a Gaussian in phase space, confirming the appropriateness of the model derived in Eqs. (18) and (19). Thus, the PSNR will increase roughly proportionally to the ROI area for a given ratio $\chi \times \omega$.

5. CONCLUSION

We combine two of Dennis Gabor's inventions, holography and the Gabor transform, creating a novel method for numerical Fresnel diffraction. Different notions from signal processing were brought together, such as sparsity, phase space, chirplets, frames, and numerical Fresnel diffraction, resulting in a novel mathematical model for the propagation of coherent wavefronts of light. The algorithm is shown to be faster than FFT-based diffraction. Multiple experiments were performed, analyzing the PSF shape and the algorithm's speed and accuracy.

In future work, we aim to extend the algorithm to handle propagation between arbitrarily oriented planes or surfaces. These can be achieved with affine transformations of phase space; as this requires non-uniform resampling of Fourier space [24], extensions to the current chirplet algorithm model will be needed. Furthermore, this could be generalized to represent arbitrary 3D surfaces by a non-planar collection of chirplets that can be propagated efficiently to the hologram plane. This could contribute to various applications utilizing CGH, such as diffractive simulations for waveguide-based displays and non-line-of-sight imaging.

As an alternative Fresnel diffraction operator with a flexible phase space mapping operator, the proposed method could be

Table 3. Quantitative Assessment Using the PSNR^a

χ	ω									
	1	2	3	4	5	6	7	8	9	10
1	12.0	12.7	13.3	13.3	13.3	13.3	13.3	13.3	13.3	13.3
2	15.4	17.6	19.7	19.7	19.6	19.6	19.6	19.6	19.6	19.6
3	18.1	22.7	29.4	29.7	29.6	29.6	29.6	29.6	29.6	29.6
4	18.7	24.5	36.8	39.1	39.9	39.9	39.9	39.9	39.9	39.9
5	18.6	24.5	40.1	46.9	52.3	52.5	52.5	52.5	52.5	52.5
6	18.6	24.4	40.6	49.0	64.0	68.1	68.4	68.4	68.4	68.4
7	18.6	24.4	40.6	49.0	66.2	78.1	84.2	84.8	84.8	84.8
8	18.6	24.4	40.6	49.0	66.3	79.2	92.8	103.5	103.5	103.5
9	18.6	24.4	40.6	49.0	66.3	79.2	93.2	120.8	127.6	128.8
10	18.6	24.4	40.6	49.0	66.3	79.2	93.2	121.7	133.6	158.1

^aThese show the PSNR of diffracting a random complex-valued 512-sample signal for various combinations of χ and ω w.r.t. the non-approximated reference method (expressed in dB).

integrated with iterative CGH algorithms, using Gerchberg-Saxton or stochastic gradient descent, enabling the potential for more selective iterative phase space updates.

Future work may also include optimizing the Gabor frame parameters or utilizing more general frames and Gabor lattices. We plan to investigate how to integrate this model with different CGH algorithms, assess to what extent further optimizations are possible on GPU, and how suitable the algorithm is for hardware implementations on FPGA or ASIC.

Funding. Fonds Wetenschappelijk Onderzoek (12ZQ223N, G089424N, G0A3O24N); Japan Society for the Promotion of Science (International research fellow P22752).

Disclosures. The authors declare no conflicts of interest.

Data Availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

REFERENCES

- J. W. Goodman, *Introduction to Fourier Optics* (W. H. Freeman and Company, 2017).
- T.-C. Poon, *Digital Holography and Three-Dimensional Display: Principles and Applications* (Springer, 2006).
- P. Picart and J.-C. Li, *Digital Holography* (Wiley, 2013).
- K. Matsushima, *Introduction to Computer Holography: Creating Computer-Generated Holograms as the Ultimate 3D Image* (Springer Nature, 2020).
- M. K. Kim, *Digital Holographic Microscopy* (Springer, 2011).
- S.-C. Kim and E.-S. Kim, "Effective generation of digital holograms of three-dimensional objects using a novel look-up table method," *Appl. Opt.* **47**, D55–D62 (2008).
- Y. Pan, X. Xu, S. Solanki, *et al.*, "Fast CGH computation using S-LUT on GPU," *Opt. Express* **17**, 18543–18555 (2009).
- A. Symeonidou, D. Blinder, A. Munteanu, *et al.*, "Computer-generated holograms by multiple wavefront recording plane method with occlusion culling," *Opt. Express* **23**, 22149–22161 (2015).
- T. Shimobaba, N. Masuda, and T. Ito, "Simple and fast calculation algorithm for computer-generated hologram with wavefront recording plane," *Opt. Lett.* **34**, 3133–3135 (2009).
- A. Maimone, A. Georgiou, and J. S. Kollin, "Holographic near-eye displays for virtual and augmented reality," *ACM Trans. Graph.* **36**, 85 (2017).
- R. P. Muffoletto, J. M. Tyler, and J. E. Tohline, "Shifted Fresnel diffraction for computational holography," *Opt. Express* **15**, 5631–5640 (2007).
- T. Shimobaba, T. Kakue, N. Okada, *et al.*, "Aliasing-reduced Fresnel diffraction with scale and shift operations," *J. Opt.* **15**, 075405 (2013).
- W. Zhang, H. Zhang, K. Matsushima, *et al.*, "Shifted band-extended angular spectrum method for off-axis diffraction calculation," *Opt. Express* **29**, 10089–10103 (2021).
- H. Wei, X. Liu, X. Hao, *et al.*, "Modeling off-axis diffraction with the least-sampling angular spectrum method," *Optica* **10**, 959–962 (2023).
- R. Heintzmann, L. Loetgering, and F. Wechsler, "Scalable angular spectrum propagation," *Optica* **10**, 1407–1416 (2023).
- Z. Wang, O. Baladron-Zorita, C. Hellmann, *et al.*, "Theory and algorithm of the homeomorphic Fourier transform for optical simulations," *Opt. Express* **28**, 10552–10571 (2020).
- D. Blinder, T. Birnbaum, and P. Schelkens, "Fresnel diffraction with linear time complexity using Gabor frames," in *Optica Imaging Congress* (Optica Publishing Group, 2023), paper HW4C.3.
- D. Blinder, M. Chlipala, T. Kozacki, *et al.*, "Photorealistic computer generated holography with global illumination and path tracing," *Opt. Lett.* **46**, 2188–2191 (2021).
- K. Wakunami, H. Yamashita, and M. Yamaguchi, "Occlusion culling for computer generated hologram based on ray-wavefront conversion," *Opt. Express* **21**, 21811–21822 (2013).
- T. Birnbaum, T. Kozacki, and P. Schelkens, "Providing a visual understanding of holography through phase space representations," *Appl. Sci.* **10**, 4766 (2020).
- D. Blinder, T. Birnbaum, and P. Schelkens, "Pincushion point-spread function for computer-generated holography," *Opt. Lett.* **47**, 2077–2080 (2022).
- T. Shimobaba, J. T. Weng, T. Sakurai, *et al.*, "Computational wave optics library for C++: CWO++ library," *Comput. Phys. Commun.* **183**, 1124–1138 (2012).
- J. F. Restrepo and J. Garcia-Sucerquia, "Magnified reconstruction of digitally recorded holograms by Fresnel–Bluestein transform," *Appl. Opt.* **49**, 6430–6435 (2010).
- D. Blinder, C. Schretter, H. Ottevaere, *et al.*, "Unitary transforms using time-frequency warping for digital holograms of deep scenes," *IEEE Trans. Comput. Imaging* **4**, 206–218 (2018).
- J. J. Healy, M. A. Kutay, H. M. Ozaktas, *et al.*, *Linear Canonical Transforms: Theory and Applications* (Springer, 2015), Vol. **198**.
- M. A. De Gosson, *Symplectic Methods in Harmonic Analysis and in Mathematical Physics* (Springer, 2011), Vol. **7**.
- F. A. Franz Hlawatsch, *Time-Frequency Analysis: Concepts and Methods* (Wiley, 2008).
- B. Boashash, *Time-Frequency Signal Analysis and Processing: A Comprehensive Reference*, 2nd ed., Eurasip and Academic Press Series in Signal and Image Processing (Academic, 2016).
- K. Gröchenig, *Foundations of Time-Frequency Analysis* (Birkhäuser, 2001).

30. J. J. Healy, W. T. Rhodes, and J. T. Sheridan, "Cross terms of the Wigner distribution function and aliasing in numerical simulations of paraxial optical systems," *Opt. Lett.* **35**, 1142–1144 (2010).
31. F. Wyrowski and M. Kuhn, "Introduction to field tracing," *J. Mod. Opt.* **58**, 449–466 (2011).
32. O. Christensen, *An Introduction to Frames and Riesz Bases*, 2nd ed. (Springer, 2016).
33. J. Rosen, S. Alford, B. Allan, *et al.*, "Roadmap on computational methods in optical imaging and holography," *Appl. Phys. B* **130**, 166 (2024).
34. N. Chen, C. Wang, and W. Heidrich, "HTRSD: hybrid Taylor Rayleigh-Sommerfeld diffraction," *Opt. Express* **30**, 37727–37735 (2022).
35. S. Fujimori, T. Ito, and T. Shimobaba, "Rationalized diffraction calculations for high accuracy and high speed with few bits," *J. Opt. Soc. Am. A* **41**, 303–310 (2024).