Holographic Predictive Search: Extending the Scope

Peter J. Christopher[a,*], Ralf Mouthaan[a], George S. D. Gordon[b], Timothy D. Wilkinson[a]

[a]*Centre of Molecular Materials, Photonics and Electronics, University of Cambridge*
[b]*Department of Electrical and Electronic Engineering, University of Nottingham*

Abstract

Holographic Predictive Search (HPS) is a novel approach to search-based hologram generation that uses a mathematical understanding of the optical transforms to make informed optimisation decisions. Existing search techniques such as Direct Search (DS) and Simulated Annealing (SA) rely on trialling modifications to a test hologram and observing the results. A formula is used to decide whether the change should be accepted. HPS operates presciently, using knowledge of the underlying mathematical relationship to make exact changes to the test hologram that guarantee the 'best' outcome for that change.

In this work, we extend the scope of the original research to cover both phase and amplitude modulating Spatial Light Modulators (SLMs), both phase sensitive and phase insensitive systems and both Fresnel and Fraunhofer diffraction. In the cases discussed, improvements of up to 10x are observed in final error and the approach also offers significant performance benefits in generation time. This comes at the expense of increased complexity and loss of generality.

Keywords: Computer Generated Holography, Holographic Predictive Search, Direct Search, Simulated Annealing, Holographic Search Algorithms

1. Introduction

The expansion of Computer Generated Holography (CGH) in recent years has seen application in areas including super resolution microscopy, optical tweezing, quantum mechanics and optical communication [1, 2, 3, 4, 5, 6]. For applications where quality is the primary consideration,

---

*Corresponding author
Email address: pjc209@cam.ac.uk (Peter J. Christopher)
URL: www.peterjchristopher.me.uk (Peter J. Christopher)

Holographic Search Algorithms (HSAs) are a common approach with algorithms like such as Direct Search (DS) and Simulated Annealing (SA) being used [7, 8].

In our recent paper we introduced Holographic Predictive Search (HPS), an algorithm that offers the potential to improve upon existing HSAs [9]. That work exclusively considered the case of phase modulated holograms where the replay field phase is of interest. Here we expand on this to treat both phase and amplitude modulated holograms; both phase sensitive and phase insensitive replay fields and both the Fresnel and Fraunhofer diffraction regimes.

HPS uses a prescient model of the Fourier and Fresnel Transforms used in far- and mid- field holography to provide a predictive alternative to traditional *blind* search approaches. While HPS offers considerable performance improvements over rival HSAs of up to $10\times$ lower convergence times it comes at the expense of reduced flexibility. Here we expand on the single case presented initially to provide algorithmic variants for an array of system combinations. Each is presented with an analysis of performance and of the relative advantages of HPS over other HSAs. Reviews of CGH are available [10, 11] so we start with only the bare minimum of background required.

2. Background

The Discrete Fourier Transform (DFT) forms the core of the holographic process,

$$F_{u,v} = \frac{1}{\sqrt{N_x N_y}} \sum_{x=0}^{N_x-1} \sum_{y=0}^{N_y-1} f_{xy} e^{-2\pi i \left(\frac{ux}{N_x} + \frac{vy}{N_y} \right)} \tag{1}$$

$$f_{x,y} = \frac{1}{\sqrt{N_x N_y}} \sum_{u=0}^{N_x-1} \sum_{v=0}^{N_y-1} F_{uv} e^{2\pi i \left(\frac{ux}{N_x} + \frac{vy}{N_y} \right)} \tag{2}$$

where u and v represent the spatial frequencies and x and y represent the source coordinates. Fast Fourier Transforms (FFTs) are typically used to calculate the DFT with calculation times of $O(N_x N_y \log N_x N_y)$ where N_x and N_y are the respective x and y resolutions [12, 13].

The far-field pattern produced by passing coherent light through a Spatial Light Modulator (SLM) is equivalent to taking the DFT of the SLM aperture function multiplied by the static pixel shape parameter and coherent illumination [14]. For an ideal pixellated SLM acting on uniform unit intensity planar wavefronts with 100% fill factor pixels, the produced hologram is given by the DFT of the SLM aperture function. More generally, the projected hologram is often referred to as the *Replay Field* or *Replay Plane* and the SLM as the *Diffraction Field* or *Diffraction Plane*.

Real-world SLMs modulate light only in a limited fashion, typically phase or amplitude only [15, 16]. When addressed digitally, this is restricted further to discrete levels. Finding an SLM aperture function for a given far-field hologram $F(u, v)$ is identical to the problem of finding $f(x, y)$ where $F(u, v) = \mathcal{F}\{f(x, y)\}$ subject to these constraints. $\mathcal{F}$ here refers to the Fourier transform.

3. Holographic Search Algorithms

In our recent paper we introduced Holographic Predictive Search (HPS) and compared it with DS and SA. The procedure for these two algorithms is shown in Figures 1 and 2. Key to these algorithms is the fact that we can avoid performing a full DFT at each iteration, instead using the following $O(N_x N_y)$ update step

$$\Delta R_{u,v} = \frac{1}{\sqrt{N_x N_y}} \Delta H_{x,y} e^{\left[-2\pi i \left(\frac{ux}{N_x} + \frac{vy}{N_y}\right)\right]} \tag{3}$$

where change $\Delta H_{x,y}$ in aperture function causes a change $\Delta R_{u,v}$ in the replay field.
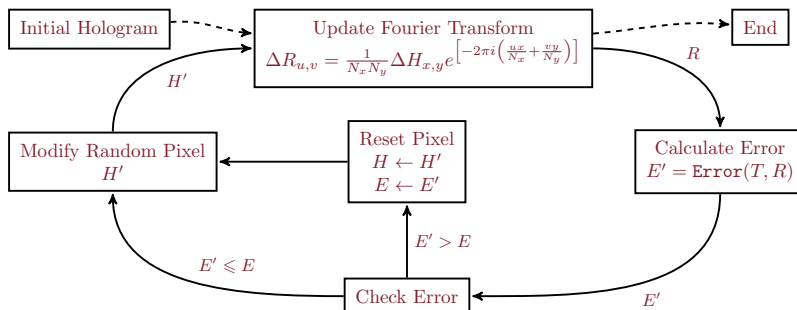

Figure 1: Fast Direct Search

One fundamental feature of these two algorithms is that they are blind, randomly selecting a new pixel value to trial. The change in error or quality metric is then observed and the change is either discarded or accepted. In the case of DS, the decision is binary with change always being accepted if the error is reduced. SA adds a probabalistic element that can sometimes select worse solutions. This slows the algorithm but reduces the chance of being captured in local minima.

4. Holographic Predictive Search

Our earlier work introduced Holographic Predictive Search (HPS) for hologram generation [9]. HPS operates in a similar manner to DS and SA but instead of blindly choosing a new pixel value
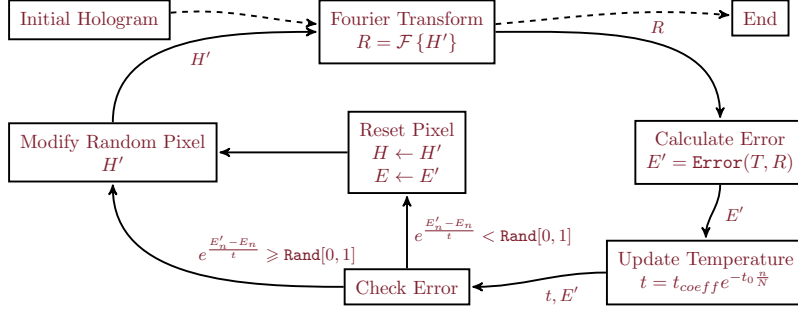
Figure 2: Simulated Annealing

and observing the change in error, HPS operates by generating a linear relationship for the *best* change in a pixel value. This led to significant performance improvements in convergence time and error reduction. The downside was an increase in computational complexity per iteration and a loss of generality in mathematical form.

As originally presented, this algorithm was only applicable in the case of a phase-modulated SLM, with a Fraunhofer hologram where the target image was phase sensitive. Here we extend the scope of this approach to cover both phase and amplitude modulating SLMs, both phase sensitive and phase insensitive systems and both Fresnel and Fraunhofer diffraction.
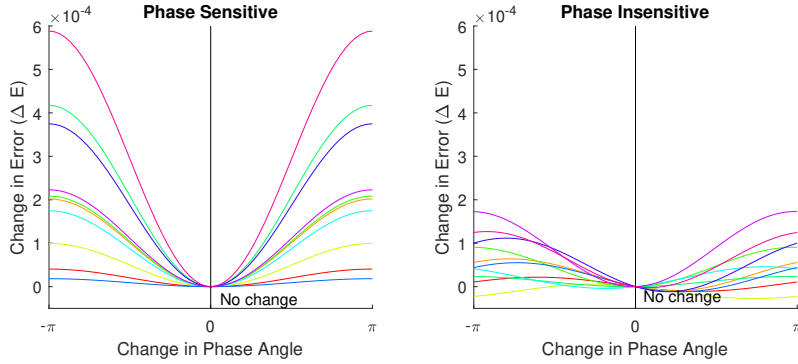

Figure 3: Selection of final errors depending on pixel changes for phase modulating SLMs for phase sensitive replay fields (left) and phase insensitive replay fields (right). The SLM is assumed to be 256×256 pixels with a flat unit intensity illumination.

To give an understanding of the differences between these two cases Figures 3 and 4 take an initial inverse Fourier transform of the *Mandrill* test image and plot the effect on mean squared

4

error (MSE) of changing the level on 10 randomly selected SLM pixels for two different categories of SLM - phase and amplitude modulating - and for two different categories of replay field - phase sensitive and insensitive. These were generated by selecting a random hologram pixel $H_{x,y}$ and plotting the change in final error for a range of pixel values. They show that the response to level changes of a single phase pixel has a near sinusoidal effect on the error of the replay field whereas changing an amplitude pixel has a more linear response.
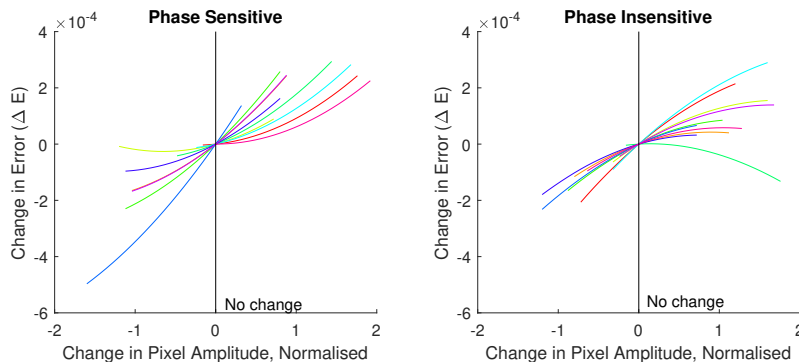

Figure 4: Selection of final errors depending on pixel changes for amplitude modulating SLMs for phase sensitive replay fields (left) and phase insensitive replay fields (right). The SLM is assumed to be 256×256 pixels with a flat unit intensity illumination.

The aim of HPS is to mathematically capture the nature of the curves shown and solve for the optimum location without having to apply Eq. 3 to every step. In practice the phase insensitive cases degenerate to quartic polynomials and it is only through the judicious use of simplifications that we can provide high speed solutions.

4.1. Extending the Scope

When originally presented, only the phase modulating SLM with phase sensitive replay field was considered as that case simultaneously provided the simplest analysis and an array of potential applications. Our aim in this paper is to develop the relationships for the other paradigms and to present analysis of their efficacy. There are three sets of variants that we must consider.

- **Transform Type** - Only far-field or Fraunhofer holograms can be modelled as an Fourier Transform. For mid-field holograms, an additional quadratic phase term must be added to form a Fresnel Transform.

- **SLM Modulation Behaviour** - SLMs typically modulate in either phase or amplitude. Each paradigm involves a different set of relationships between the SLM and the replay field.

- **Phase Sensitivity** - As originally presented, HPS aims to satisfy both the amplitude and phase constraints of the target. Many display applications, however, do not require the phase constraint due to the human eye's phase insensitivity. This additional freedom requires a separate formulation.

This paper sets out the necessary background to conform to every combinations of these constraints.

4.2. Error and Quality Metrics

The choice of whether phase sensitivity is considered also changes the error metric used where T and R are the target and actual replay fields

$$E_{MSE,ps}(T,R) = \frac{1}{N_x N_y} \sum_{u=0}^{N_x-1} \sum_{v=0}^{N_y-1} [|T_{u,v} - R_{u,v}|]^2$$

$$E_{MSE,pi}(T,R) = \frac{1}{N_x N_y} \sum_{u=0}^{N_x-1} \sum_{v=0}^{N_y-1} [|T_{u,v}| - |R_{u,v}|]^2 \tag{4}$$

Where $E_{MSE,ps}$ and $E_{MSE,pi}$ represent the *phase sensitive* and the *phase insensitive* mean squared errors (MSE) respectively.

The Structural Similarity Index (SSIM) is often used in preference to MSE when image quality rather than numerical error is the primary concern [17]. The algorithms presented here primarily target MSE and we return to this issue later.

5. Methods

It is challenging to fairly compare the performance of techniques across different system designs. In order to best serve our readers we have adopted the following conventions:

1. We have used the *Mandrill* test image shown in Figure 5a to provide the target intensities.

2. Amplitude holograms are generated with a rotationally symmetric version of Figure 5c to avoid error due to image symmetry.

3. Phase sensitive holograms are generated with the *Peppers* test image used used as the phase component as shown in Figure 5b with rotationally symmetric variant shown in Figure 5d.

4. For phase insensitive holograms, the entire target region is solved for. In the case of phase sensitive holograms, for reasons of degrees of freedom, we have scaled the target image to only fill the central quadrant of the initial replay field and set the surrounding regions to zero.

5. Planar unit intensity incident on the hologram is assumed with the target scaled to ensure conservation of energy.

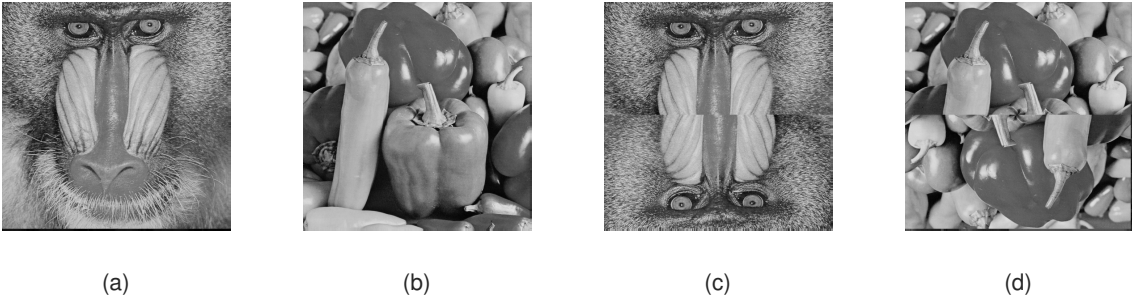|			
(a)	(b)	(c)	(d)

Figure 5: The two test images used showing the *Mandrill* and *Peppers* as well as their artificially symmetric counterparts.

These differences in method mean that the normalised error metrics used should be treated as distinct in each case and cannot be compared quantitatively between cases..

6. Fraunhofer Domain

6.1. Phase Modulated SLM, Phase Sensitive Replay Field

HPS was originally developed for the case of a Fraunhofer or far-field system displayed on a phase modulating SLM with a phase sensitive target [9]. That algorithm is shown in Figure 6 with problem geometry as shown in Figure 7.

The single pixel x, y on the SLM is set to zero and the resulting change in the replay field $R^\dagger$ and error $E^\dagger$ calculated using Eq. 3. An expression for the error after assigning a new phase to the selected pixel is then given by

$$\Delta E' = 1 - \frac{2}{\sqrt{N_x N_y}} \left[\cos \theta_{H'} \sum_{u=0}^{N_x-1} \sum_{v=0}^{N_y-1} \sqrt{E_{u,v}^\dagger} \cos C_{u,v} + \sin \theta_{H'} \sum_{u=0}^{N_x-1} \sum_{v=0}^{N_y-1} \sqrt{E_{u,v}^\dagger} \sin C_{u,v} \right]$$
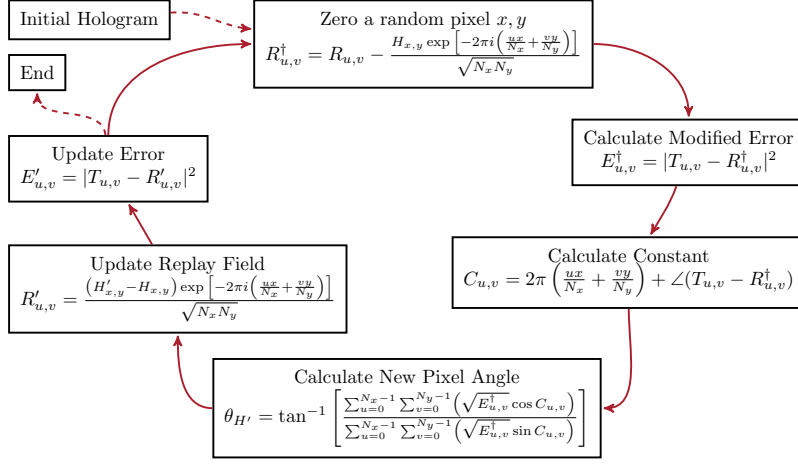
7

Figure 6: Holographic Predictive Search for phase modulated, far-field holography with phase sensitive target. Modified with permission from [9].

$$C_{u,v} = 2\pi \left(\frac{ux}{N_x} + \frac{vy}{N_y} \right) + \angle(T_{u,v} - R_{u,v}^\dagger) \tag{5}$$

This formulation only uses terms known at runtime and allows for solutions to find the most desirable value of new SLM pixel phase $\theta_{H'}$. Note that $\angle X$ here refers to the phase angle of X.

$$\theta_{H'} = \tan^{-1} \left[\frac{\sum_{u=0}^{N_x-1} \sum_{v=0}^{N_y-1} \left(\sqrt{E_{u,v}^\dagger} \sin C_{u,v} \right)}{\sum_{u=0}^{N_x-1} \sum_{v=0}^{N_y-1} \left(\sqrt{E_{u,v}^\dagger} \cos C_{u,v} \right)} \right] \tag{6}$$

This allowed for greatly improved speed of convergence at the expense of additional iteration complexity.

Below we discuss the seven other cases.

6.2. Phase Modulated SLM, Phase Insensitive Replay Field

The phase insensitivity of the eye means that display applications are often phase insensitive. This greatly increases the problem freedom but also changes the predictive geometry into a non-linear problem. The updated regime is shown on the Argand diagram in Figure 8.

Zeroing an individual SLM pixel x, y introduces error to location u, v in the replay field R given by Eq. 3 with $\Delta H_{x,y} = -H_{x,y}$. This replay field we term $R^\dagger$. The task is to find new pixel value $H'_{x,y}$ of unit magnitude so that the error in the new replay field R' is minimised. Expressing θ and θ', the old and new pixel phases, in terms of unknown $\angle H'_{x,y}$ and known x, y, u, v, N_x and N_y,
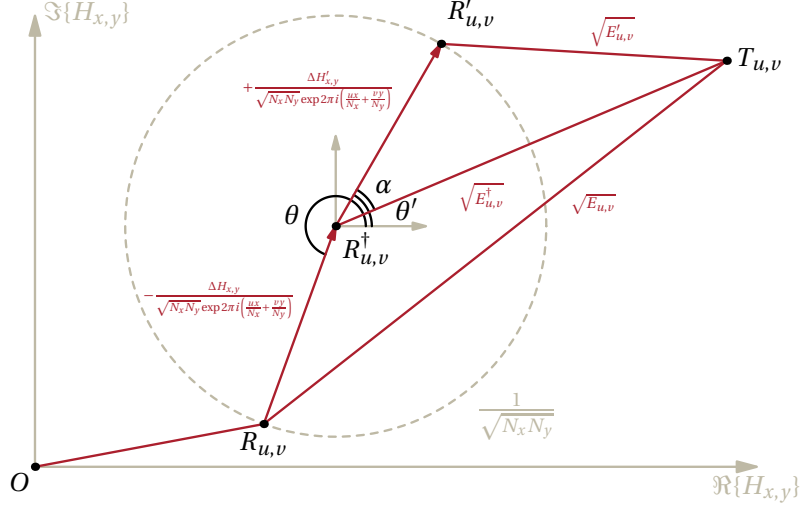
8

Figure 7: Problem geometry in phase modulated, phase insensitive case. Used with permission from [9].

$$\theta = \angle H_{x,y} - 2\pi \left(\frac{ux}{N_x} + \frac{vy}{N_y} \right),$$

$$\theta' = \angle H'_{x,y} - 2\pi \left(\frac{ux}{N_x} + \frac{vy}{N_y} \right) \tag{7}$$

The error after zeroing pixel x, y is given as $E^{\dagger}_{u,v} = \left(|T_{u,v}| - |R^{\dagger}_{u,v}| \right)^2$. The new error $E'_{u,v}$ is given by

$$E'_{u,v} = \left[|T_{u,v}| - |R'_{u,v}| \right]^2$$

$$= \left[|T_{u,v}| - \sqrt{|R^{\dagger}_{u,v}|^2 + \frac{1}{N_x N_y} + \frac{2|R^{\dagger}_{u,v}| \cos \alpha}{\sqrt{N_x N_y}}} \right]^2$$

$$\Rightarrow \Delta E'_{u,v} = E'_{u,v} - E^{\dagger}_{u,v}$$

$$= \frac{1}{N_x N_y} + \frac{2|R^{\dagger}_{u,v}| \cos \alpha}{\sqrt{N_x N_y}} + 2|T_{u,v}||R^{\dagger}_{u,v}|$$

$$- 2|T_{u,v}| \sqrt{|R^{\dagger}_{u,v}|^2 + \frac{1}{N_x N_y} + \frac{2|R^{\dagger}_{u,v}| \cos \alpha}{\sqrt{N_x N_y}}} \tag{8}$$

As

$$\alpha = \theta' - \angle R^{\dagger}_{u,v} = \angle H'_{x,y} - \left[2\pi \left(\frac{ux}{N_x} + \frac{vy}{N_y} \right) + \angle R^{\dagger}_{u,v} \right] \tag{9}$$
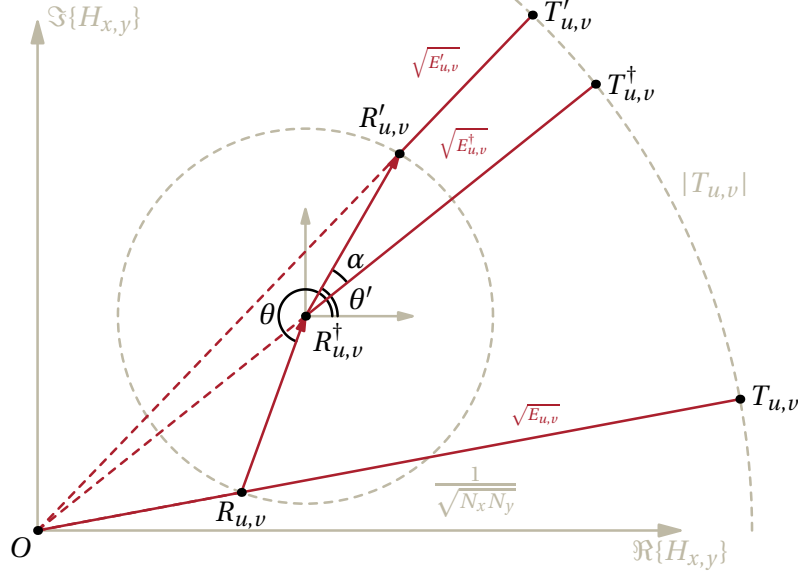
9

Figure 8: Problem geometry in phase modulated, phase insensitive case.

we can apply the Taylor expansion of $\sqrt{1+z}$ to give

$$\Delta E'_{u,v} = D_{u,v} + F_{u,v} \left[\cos \theta_{H'} \cos C_{u,v} + \sin \theta_{H'} \sin C_{u,v} \right]$$

where

$$D_{u,v} = \frac{1}{N_x N_y} + 2|T_{u,v}||R^{\dagger}_{u,v}| - 2|T_{u,v}|\sqrt{|R^{\dagger}_{u,v}|^2 + \frac{1}{N_x N_y}}$$

$$F_{u,v} = \frac{2|R^{\dagger}_{u,v}|}{\sqrt{N_x N_y}} - \frac{2|T_{u,v}||R^{\dagger}_{u,v}|}{\sqrt{N_x N_y}\sqrt{|R^{\dagger}_{u,v}|^2 + \frac{1}{N_x N_y}}}$$

$$C_{u,v} = 2\pi \left(\frac{ux}{N_x} + \frac{vy}{N_y} \right) + \angle R^{\dagger}_{u,v} \tag{10}$$

Summing over all pixels

$$\Delta E' = \sum_{u=0}^{N_x-1} \sum_{v=0}^{N_y-1} \Delta E'_{u,v}$$

$$= \sum_{u=0}^{N_x-1} \sum_{v=0}^{N_y-1} D_{u,v} + \cos \theta_{H'} \sum_{u=0}^{N_x-1} \sum_{v=0}^{N_y-1} F_{u,v} \cos C_{u,v} + \sin \theta_{H'} \sum_{u=0}^{N_x-1} \sum_{v=0}^{N_y-1} F_{u,v} \sin C_{u,v} \tag{11}$$

Taking $\mathrm{d}\Delta E' / \mathrm{d}\theta_{H'} = 0$ to find the the value of $\theta_{H'}$ where $\Delta E'$ is minimum

$$\sin\theta_{H'} \sum_{u=0}^{N_x-1} \sum_{v=0}^{N_y-1} F_{u,v} \cos C_{u,v} - \cos\theta_{H'} \sum_{u=0}^{N_x-1} \sum_{v=0}^{N_y-1} F_{u,v} \sin C_{u,v} = 0 \tag{12}$$

which is trivially solvable.

$$\theta_{H'} = \tan^{-1}\left[\frac{\sum_{u=0}^{N_x-1}\sum_{v=0}^{N_y-1}(F_{u,v}\sin C_{u,v})}{\sum_{u=0}^{N_x-1}\sum_{v=0}^{N_y-1}(F_{u,v}\cos C_{u,v})}\right] \tag{13}$$

We can choose the correct solution by using $\mathrm{d}^2\Delta E'/\mathrm{d}\theta_{H'}^2 > 0$

$$\cos\theta_{H'} \sum_{u=0}^{N_x-1} \sum_{v=0}^{N_y-1} F_{u,v} \cos C_{u,v} + \sin\theta_{H'} \sum_{u=0}^{N_x-1} \sum_{v=0}^{N_y-1} F_{u,v} \sin C_{u,v} > 0 \tag{14}$$
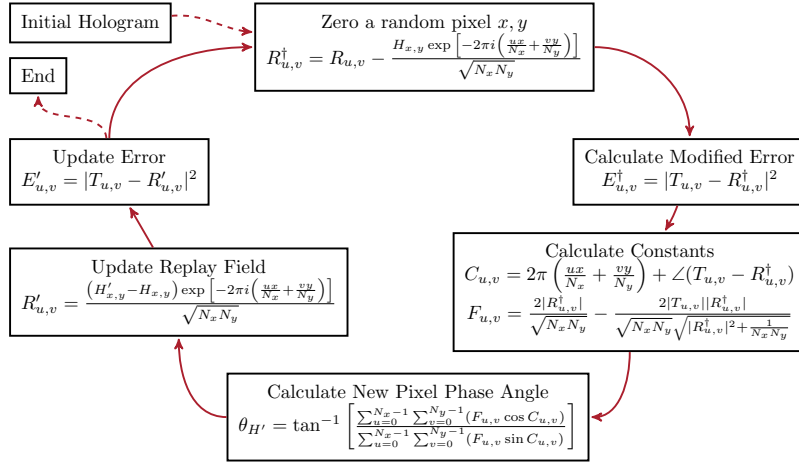
This translates into the algorithm shown in Figure. 9

Figure 9: Holographic Predictive Search for phase modulated, far-field holography with phase insensitive target

Tests of convergence for a 256×256 *Mandrill* test image on a 2^8 level phase SLM gives the performance graph as shown in Figure 10. The normalisation assumes a unit energy illumination to every SLM pixel and the target is scaled accordingly.

Comparison with the DS case presents a significant improvement in convergence speed but this is less marked than the phase sensitive case. This is not unexpected as an examination of Figure 3 will show that the degree of variation per pixel is much lower in the PI case. i.e. the hologram initial hologram is much closer to the theoretical *best* hologram than in the PS case. Nonetheless, the error at the end of $100,000$ iterations is $10\times$ lower for HPS than for DS.
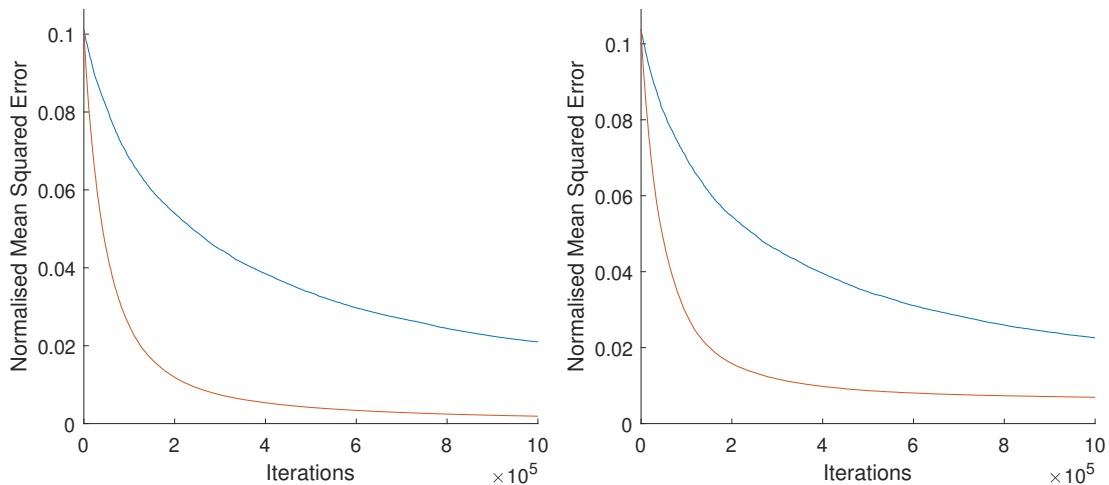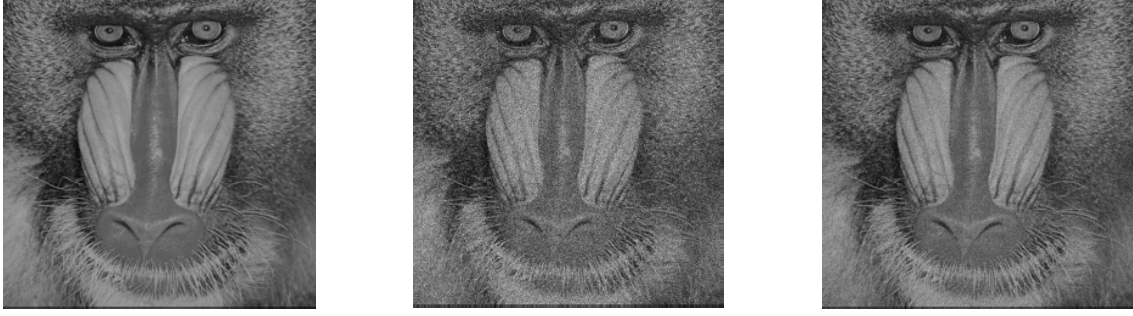
Figure 10: Comparison of Direct Search (blue) against phase modulated, phase insensitive HPS (orange) for the 256×256 pixel *Mandrill* test image being displayed on a 2^8 (left) and 2^4 (right) phase level spatial light modulator. Each trend line is taken as the mean of 5 independently seeded runs.

Figure 10 also provides a comparison of the 2^8 modulation level (left) vs 2^4 modulation level (right) cases. This shows a similar performance improvement in both cases. We return to the effect of number of modulation levels later, but for now it should be noted that the primary difference is in an increased convergent error with other features remaining indistinguishable.

The algorithm reconstruction is shown in Figure 11 showing target image (left), DS (centre) and HPS (right). The SSIM values given are calculated with a dynamic range of 1.0.

Note also the use of the Taylor series expansion. An exact solution quickly devolves into a quartic polynomial which, while solvable, proves expensive computationally. Note also that it would be entirely possible to use more terms of the Taylor series for increased accuracy at the cost of performance. Our initial tests suggested that the use of only two terms was sufficient with error in predicted angle $\theta_{H'}$ never going above 1%.

Finally, we note that while HPS compares favourable with DS in this case, it is unlikely to offer benefits over Gerchberg-Saxton for high numbers of modulation levels [18]. We return to this discussion later.

| Target amplitude | DS amplitude (SSIM=0.59) | HPS amplitude (SSIM=0.82) |

Figure 11: Comparison of Direct Search (centre) against phase modulated, phase insensitive HPS (right) for 256×256 pixel *Mandrill* test image (left) being displayed on a 2^8 phase level spatial light modulator. Both algorithms were run for $1,000,000$ iterations.

6.3. Amplitude Modulated SLM, Phase Sensitive Replay Field

Similar to its phase modulated counterpart, amplitude modulated phase sensitive HPS has a problem geometry as shown in Figure 12. Unlike in the phase modulated case, the new pixel phase angle $\theta_{H'}$ is equal to zero. This allows us to skip the step of zeroing a pixel and ignore $R^{\dagger}$ and $E^{\dagger}$. Instead we can work in terms of $\Delta r' = r' - r$ where r and r' are the old and new pixel magnitudes respectively.

The initial error before modifying pixel x, y is given as $E_{u,v} = |T_{u,v} - R_{u,v}|^2$ which is knowable at runtime. The new error $E'_{u,v}$ is given by

$$
\begin{aligned}
E'_{u,v} &= |T_{u,v} - R'_{u,v}|^2 \\
&= \left[|T_{u,v} - R_{u,v}| - \frac{\Delta r}{\sqrt{N_x N_y}} \cos \beta \right]^2 + \left[\frac{\Delta r}{\sqrt{N_x N_y}} \sin \beta \right]^2 \\
&= E_{u,v} + \frac{\Delta r^2}{N_x N_y} \cos^2 \beta - 2\sqrt{E_{u,v}} \frac{\Delta r}{\sqrt{N_x N_y}} \cos \beta + \frac{\Delta r^2}{N_x N_y} \sin^2 \beta
\end{aligned}
\tag{15}
$$

Remembering that $\cos^2 \beta + \sin^2 \beta = 1$, the change in error for given $\frac{\Delta r}{\sqrt{N_x N_y}}$ is

$$
\begin{aligned}
\Delta E'_{u,v} &= E'_{u,v} - E_{u,v} \\
&= \frac{\Delta r^2}{N_x N_y} - 2\sqrt{E_{u,v}} \frac{\Delta r}{\sqrt{N_x N_y}} \cos \beta
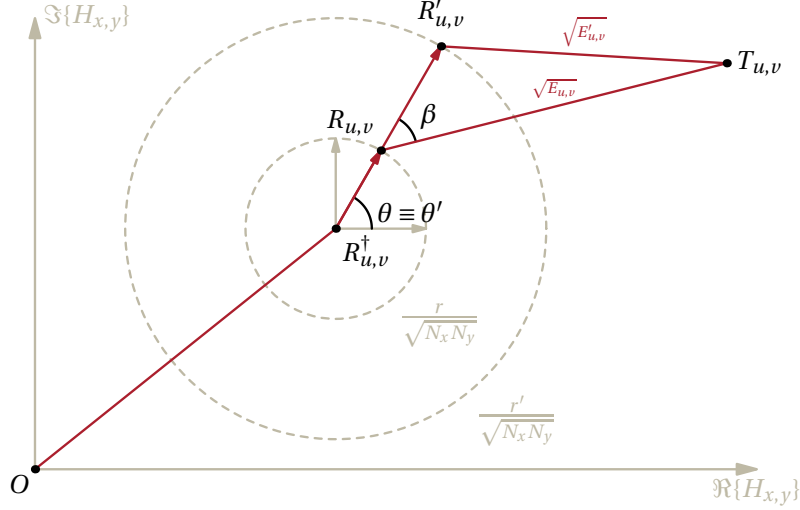\end{aligned}
\tag{16}
$$

13

Figure 12: Problem geometry in amplitude modulated, phase sensitive case.

$\beta_{u,v}$ is given from $\theta_{u,v}$

$$\beta_{u,v} = \theta_{u,v} - \angle(T_{u,v} - R_{u,v}) = -2\pi\left(\frac{ux}{N_x} + \frac{vy}{N_y}\right) - \angle(T_{u,v} - R_{u,v}) \tag{17}$$

where

$$\theta' = \theta = -2\pi\left(\frac{ux}{N_x} + \frac{vy}{N_y}\right) \tag{18}$$

Summing

$$\Delta E' = \sum_{u=0}^{N_x-1}\sum_{v=0}^{N_y-1} \Delta E'_{u,v}$$

$$= \Delta r^2 - \frac{2\Delta r}{\sqrt{N_x N_y}} \sum_{u=0}^{N_x-1}\sum_{v=0}^{N_y-1} \sqrt{E_{u,v}}\cos\beta \tag{19}$$

Taking $\mathrm{d}\Delta E'/\mathrm{d}\Delta r = 0$ to find the the value of Δr where $\Delta E'$ is minimum

$$\Delta r = \frac{1}{\sqrt{N_x N_y}} \sum_{u=0}^{N_x-1}\sum_{v=0}^{N_y-1} \sqrt{E_{u,v}}\cos\beta \tag{20}$$

The linear nature of this result means that we can cap Δr within the constraints of the SLM. This is shown algorithmically in Figure 13.

This is modelled for a 256×256 *Mandrill* test image with target phase given by the *Peppers* test image. On a simulated 2^8 level amplitude SLM this gives the performance graph as shown in
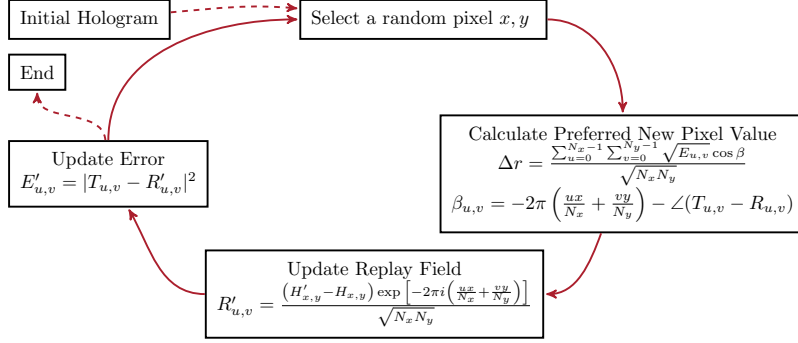
14

Figure 13: Holographic Predictive Search for amplitude modulated, far-field holography with phase sensitive target.

Figure 14. This results in an approximately $2\times$ improvement in convergence speed to reach a given target error.
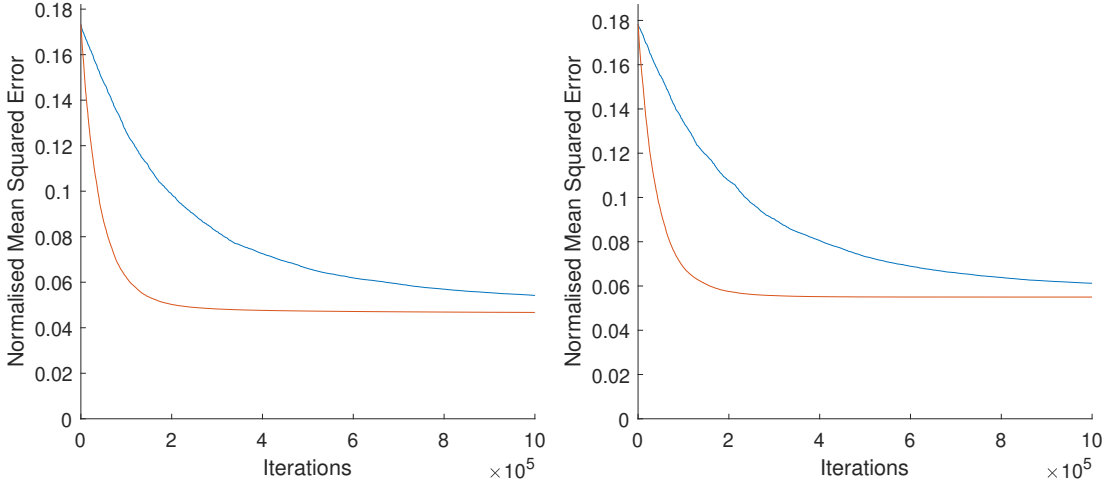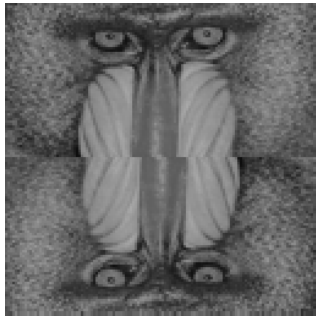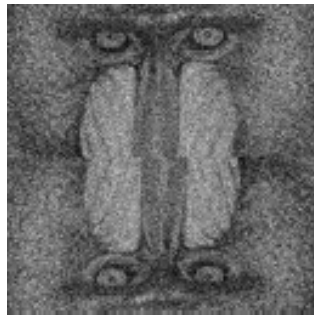

Figure 14: Comparison of Direct Search (blue) against amplitude modulated, phase sensitive HPS (orange) for the 256×256 pixel *Mandrill* test image being displayed on a 2^8 (left) and 2^4 (right) amplitude level spatial light modulator. Each trend line is taken as the mean of 5 independently seeded runs.

The algorithm reconstruction is shown in Figure 15 showing target image (left), DS (centre) and HPS (right). The SSIM values given are calculated with a dynamic range of 1.0.
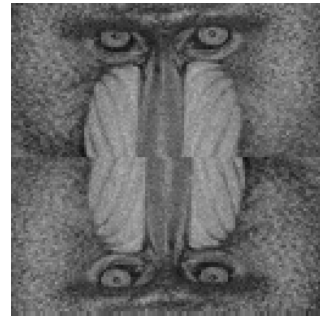
Very similar performance improvements were seen when used in place of a simulated annealing algorithm where HPS consistently outperformed SA approximately $2\times$ in terms of iterations required to reach a given target error. While this is interesting, it should be noted that phase
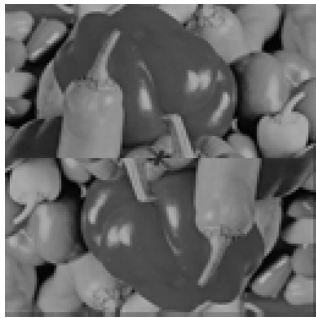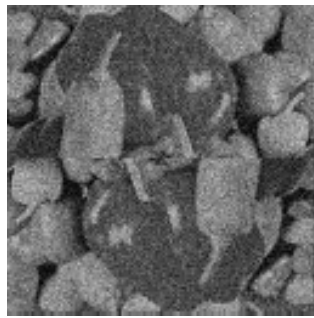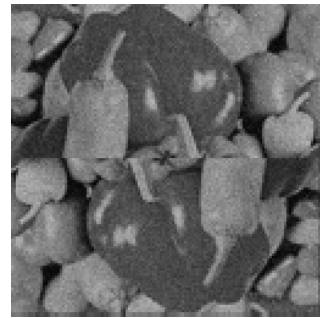
Figure 15: Comparison of Direct Search (centre) against amplitude modulated, phase sensitive HPS (right) for 128×128 pixel *Mandrill* test image amplitudes (top left) and *Peppers* test image phases (top left) being displayed on a 2^8 amplitude level 256×256 pixel spatial light modulator. Both algorithms were run for $1,000,000$ iterations.
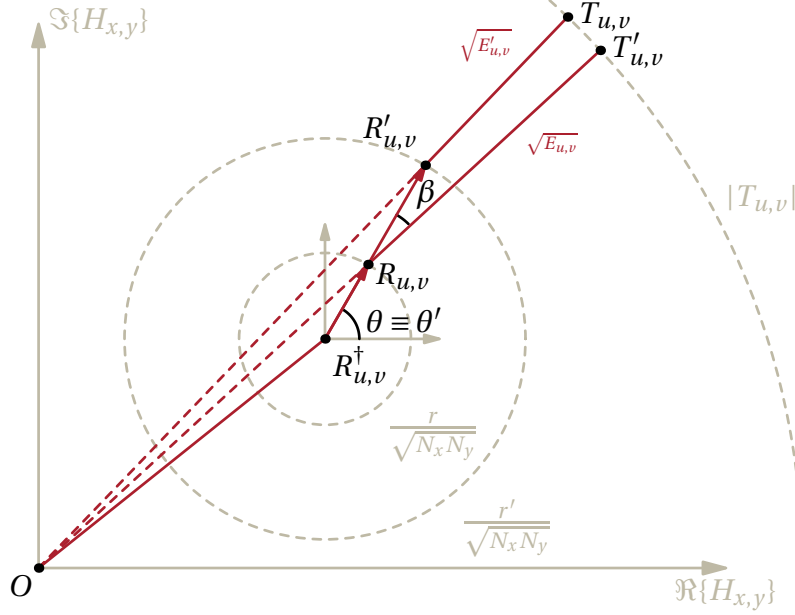
Figure 16: Problem geometry in amplitude modulated, phase insensitive case.

sensitive amplitude holography is unlikely to be a common paradigm and this result is presented primarily for completeness. We return to this later.

6.4. Amplitude Modulated SLM, Phase Insensitive Replay Field

The phase insensitive amplitude modulated behaviour is similar to the phase insensitive phase modulated case. The problem geometry is shown in Figure 16. Working again in terms of $\Delta r' = r' - r$ where r and r' are the old and new pixel magnitudes respectively.

$$E'_{u,v} = \left[|T_{u,v}| - |R'_{u,v}| \right]^2$$

$$= \left[|T_{u,v}| - \sqrt{ |R_{u,v}|^2 + \frac{\Delta r^2}{N_x N_y} + 2|R_{u,v}| \frac{\Delta r}{\sqrt{N_x N_y}} \cos \beta } \right]^2$$

$$\Rightarrow \Delta E'_{u,v} = E'_{u,v} - E_{u,v}$$

$$= \frac{\Delta r^2}{N_x N_y} + 2|R_{u,v}| \frac{\Delta r}{\sqrt{N_x N_y}} \cos \beta + 2|T_{u,v}||R_{u,v}|$$

$$- 2|T_{u,v}||R_{u,v}| \sqrt{ |R_{u,v}|^2 + \frac{\Delta r^2}{N_x N_y} + 2|R_{u,v}| \frac{\Delta r}{\sqrt{N_x N_y}} \cos \beta } \tag{21}$$

Unlike in Section 6.2, there is no easy Taylor substitution. Instead we assume that $\Delta r^2 + 2|R_{u,v}|\Delta r \cos\beta$ is smaller than $|R_{u,v}|^2$. This assumption can be seen to be valid for almost all non-zero target replay field values. For the *Mandrill* test image this results in $> 99.99\%$ of pixels being valid with this assumption. For different amplitude distributions this assumption becomes less valid. Fortunately for us, however, the system is insensitive to such pixels as they have near zero magnitude and this is further improved by the square relationship between intensity and amplitude. In the tests run here we found less that 0.02% of pixels for which this gave greater than 1% error in target value.

If we allow this assumption we can write

$$\Delta E'_{u,v} = E'_{u,v} - E_{u,v}$$
$$= (1 - |T_{u,v}|)(\frac{\Delta r^2}{N_x N_y} + 2|R_{u,v}|\frac{\Delta r}{\sqrt{N_x N_y}}\cos\beta) \tag{22}$$

Where $\beta_{u,v}$ is again given by

$$\beta_{u,v} = \theta_{u,v} - \angle(T_{u,v} - R_{u,v}) = -2\pi\left(\frac{ux}{N_x} + \frac{vy}{N_y}\right) - \angle(T_{u,v} - R_{u,v}) \tag{23}$$

Summing

$$\Delta E' = \sum_{u=0}^{N_x-1}\sum_{v=0}^{N_y-1}\Delta E'_{u,v}$$
$$= (1 - \sum_{u=0}^{N_x-1}\sum_{v=0}^{N_y-1}T_{u,v})\Delta r^2 + \frac{2\Delta r}{\sqrt{N_x N_y}}\sum_{u=0}^{N_x-1}\sum_{v=0}^{N_y-1}(1 - T_{u,v})|R_{u,v}|\cos\beta \tag{24}$$

Taking $\mathrm{d}\Delta E'/\mathrm{d}\Delta r = 0$ to find the the value of Δr where $\Delta E'$ is minimum

$$\Delta r = \frac{1}{\sqrt{N_x N_y}}\sum_{u=0}^{N_x-1}\sum_{v=0}^{N_y-1}(1 - T_{u,v})|R_{u,v}|\cos\beta \tag{25}$$

The linear nature of this result means that we can cap Δr within the constraints of the SLM. This translates into the algorithm shown in Figure. 17

Tests of convergence for a 256×256 *Mandrill* test image on a 2^8 level amplitude SLM gives the performance graph as shown in Figure 18. This again results in an approximately $2\times$ improvement in convergence iterations to reach a given target error.

The algorithm reconstruction is shown in Figure 19 showing target image (left), DS (centre) and HPS (right). The SSIM values given are calculated with a dynamic range of 1.0.
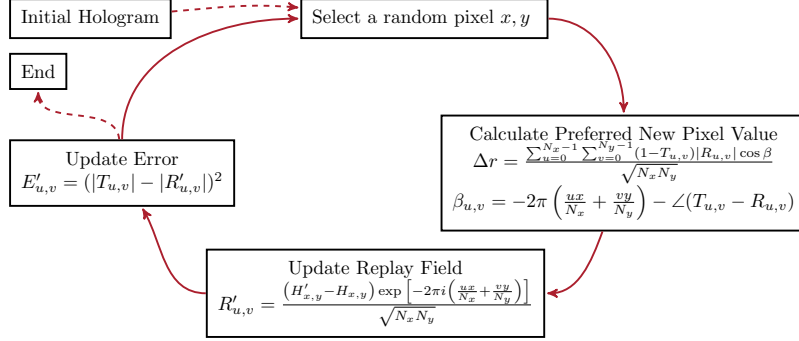
Figure 17: Holographic Predictive Search for amplitude modulated, far-field holography with phase insensitive target.
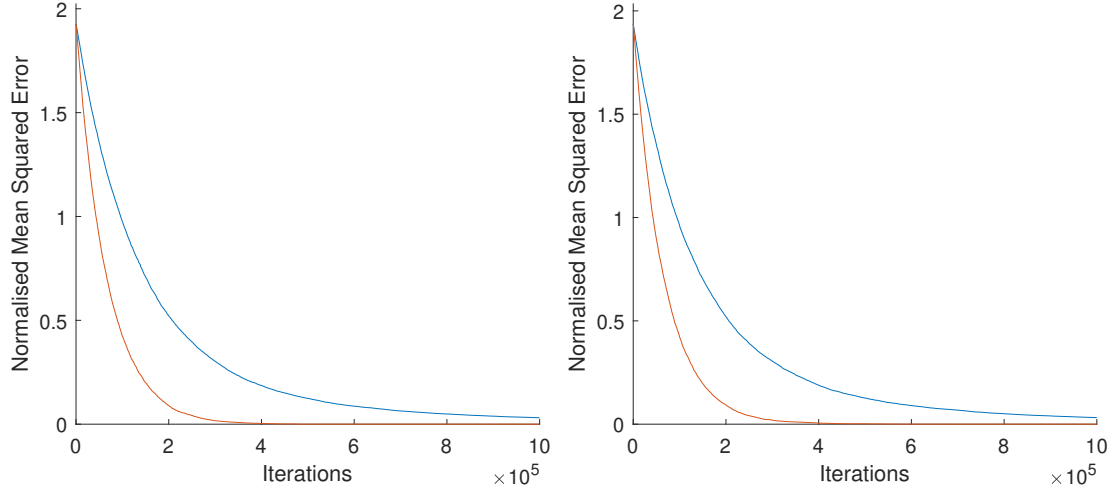

Figure 18: Comparison of Direct Search (blue) against amplitude modulated, phase insensitive HPS (orange) for the 256×256 pixel *Mandrill* test image being displayed on a 2^8 (left) and 2^4 (right) amplitude level spatial light modulator. Each trend line is taken as the mean of 5 independently seeded runs.
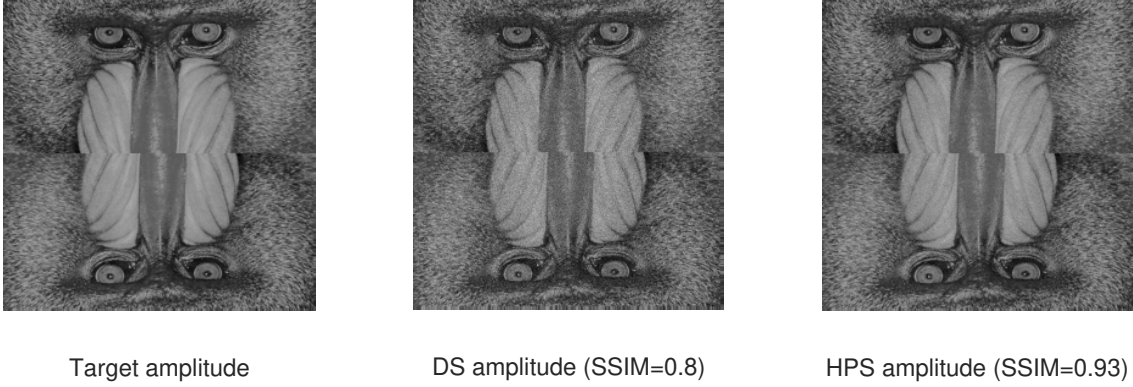
19

| Target amplitude | DS amplitude (SSIM=0.8) | HPS amplitude (SSIM=0.93) |

Figure 19: Comparison of Direct Search (centre) against amplitude modulated, phase insensitive HPS (right) for 256×256 pixel *Mandrill* test image (left) being displayed on a 2^8 amplitude level spatial light modulator. Both algorithms were run for $1,000,000$ iterations.

7. Fresnel Domain

Fortunately, the Fresnel transform variants of HPS turn out to have a very similar form to their Fraunhofer counterparts. The only distinction is the addition of the quadratic phase term to $C_{u,v}$ and $\beta_{u,v}$

$$
C_{u,v} = 2\pi \left(\frac{ux}{N_x} + \frac{vy}{N_y} \right) + \angle R_{u,v}^\dagger + e^{\frac{i\pi}{\lambda z}(x^2 + y^2)}
$$
$$
\beta_{u,v} = 2\pi \left(\frac{ux}{N_x} + \frac{vy}{N_y} \right) - \angle(T_{u,v} - R_{u,v}) + e^{\frac{i\pi}{\lambda z}(x^2 + y^2)} \tag{26}
$$

where z is the perpendicular separation between diffraction field and replay field and λ is the illumination wavelength.

8. Summary, Discussion and Recommendations

The required relationships for the different HPS variants are summarised in Table 1. While we have briefly presented the relationships behind HPS, there are a significant number of points that should be discussed.

Firstly, competitor algorithm families should be considered. HSAs such as DS can be used to generate some of the best quality holograms, albeit at the expense of slower generation times and HPS uniformly out-performed DS for every case discussed in this paper in both speed and convergent quality.

Table 1: Mathematical relationships for different HPS variants

Phase Modulated SLM, Phase Sensitive Replay Field, Fraunhofer Domain	Phase Modulated SLM, Phase Sensitive Replay Field, Fresnel Domain
$C_{u,v} = 2\pi\left(\dfrac{ux}{N_x} + \dfrac{vy}{N_y}\right) + \angle R_{u,v}^{\dagger}$ $\theta_{H'} = \tan^{-1}\left[\dfrac{\sum_{u=0}^{N_x-1}\sum_{v=0}^{N_y-1}\left(\sqrt{E_{u,v}^{\dagger}}\cos C_{u,v}\right)}{\sum_{u=0}^{N_x-1}\sum_{v=0}^{N_y-1}\left(\sqrt{E_{u,v}^{\dagger}}\sin C_{u,v}\right)}\right]$	$C_{u,v} = 2\pi\left(\dfrac{ux}{N_x} + \dfrac{vy}{N_y}\right) + \angle R_{u,v}^{\dagger} + e^{\frac{i\pi}{\lambda z}(x^2+y^2)}$ $\theta_{H'} = \tan^{-1}\left[\dfrac{\sum_{u=0}^{N_x-1}\sum_{v=0}^{N_y-1}\left(\sqrt{E_{u,v}^{\dagger}}\cos C_{u,v}\right)}{\sum_{u=0}^{N_x-1}\sum_{v=0}^{N_y-1}\left(\sqrt{E_{u,v}^{\dagger}}\sin C_{u,v}\right)}\right]$
Phase Modulated SLM, Phase Insensitive Replay Field, Fraunhofer Domain	**Phase Modulated SLM, Phase Insensitive Replay Field, Fresnel Domain**
$C_{u,v} = 2\pi\left(\dfrac{ux}{N_x} + \dfrac{vy}{N_y}\right) + \angle R_{u,v}^{\dagger}$ $F_{u,v} = \dfrac{2|R_{u,v}^{\dagger}|}{\sqrt{N_xN_y}} - \dfrac{2|T_{u,v}||R_{u,v}^{\dagger}|}{\sqrt{N_xN_y}\sqrt{|R_{u,v}^{\dagger}|^2 + \frac{1}{N_xN_y}}}$ $\theta_{H'} = \tan^{-1}\left[\dfrac{\sum_{u=0}^{N_x-1}\sum_{v=0}^{N_y-1}\left(F_{u,v}\cos C_{u,v}\right)}{\sum_{u=0}^{N_x-1}\sum_{v=0}^{N_y-1}\left(F_{u,v}\sin C_{u,v}\right)}\right]$	$C_{u,v} = 2\pi\left(\dfrac{ux}{N_x} + \dfrac{vy}{N_y}\right) + \angle R_{u,v}^{\dagger} + e^{\frac{i\pi}{\lambda z}(x^2+y^2)}$ $F_{u,v} = \dfrac{2|R_{u,v}^{\dagger}|}{\sqrt{N_xN_y}} - \dfrac{2|T_{u,v}||R_{u,v}^{\dagger}|}{\sqrt{N_xN_y}\sqrt{|R_{u,v}^{\dagger}|^2 + \frac{1}{N_xN_y}}}$ $\theta_{H'} = \tan^{-1}\left[\dfrac{\sum_{u=0}^{N_x-1}\sum_{v=0}^{N_y-1}\left(F_{u,v}\cos C_{u,v}\right)}{\sum_{u=0}^{N_x-1}\sum_{v=0}^{N_y-1}\left(F_{u,v}\sin C_{u,v}\right)}\right]$
Amplitude Modulated SLM, Phase Sensitive Replay Field, Fraunhofer Domain	**Amplitude Modulated SLM, Phase Sensitive Replay Field, Fresnel Domain**
$\beta_{u,v} = 2\pi\left(\dfrac{ux}{N_x} + \dfrac{vy}{N_y}\right) - \angle(T_{u,v} - R_{u,v})$ $\Delta r = \dfrac{1}{\sqrt{N_xN_y}}\sum_{u=0}^{N_x-1}\sum_{v=0}^{N_y-1}\sqrt{E_{u,v}}\cos\beta$	$\beta_{u,v} = 2\pi\left(\dfrac{ux}{N_x} + \dfrac{vy}{N_y}\right) - \angle(T_{u,v} - R_{u,v}) + e^{\frac{i\pi}{\lambda z}(x^2+y^2)}$ $\Delta r = \dfrac{1}{\sqrt{N_xN_y}}\sum_{u=0}^{N_x-1}\sum_{v=0}^{N_y-1}\sqrt{E_{u,v}}\cos\beta$
Amplitude Modulated SLM, Phase Insensitive Replay Field, Fraunhoferr Domain	**Amplitude Modulated SLM, Phase Insensitive Replay Field, Fresnel Domain**
$\beta_{u,v} = 2\pi\left(\dfrac{ux}{N_x} + \dfrac{vy}{N_y}\right) - \angle(T_{u,v} - R_{u,v})$ $\Delta r = \dfrac{1}{\sqrt{N_xN_y}}\sum_{u=0}^{N_x-1}\sum_{v=0}^{N_y-1}(1-T_{u,v})|R_{u,v}|\cos\beta$	$\beta_{u,v} = 2\pi\left(\dfrac{ux}{N_x} + \dfrac{vy}{N_y}\right) - \angle(T_{u,v} - R_{u,v}) + e^{\frac{i\pi}{\lambda z}(x^2+y^2)}$ $\Delta r = \dfrac{1}{\sqrt{N_xN_y}}\sum_{u=0}^{N_x-1}\sum_{v=0}^{N_y-1}(1-T_{u,v})|R_{u,v}|\cos\beta$

Iterative algorithms - for example Gerchberg-Saxton [18] - are available for many relatively smooth systems and should be expected to be significantly faster than HSAs including HPS. This comes at the expense of final image quality where HPS is expected to still give the best performance. For more-discontinuous systems with lower numbers of modulation levels, iterative algorithms can fail to converge and HSAs become a suitable alternative. In this case HPS may be expected to offer better performance in speed as well as quality.

Secondly, the computational performance of HPS should be considered. In our previous work we found that HPS required approximately $70 - 80\%$ more time per iteration though this dropped to as little as 10% in the case of larger images where computation was memory bound. This increased iteration time should be taken into account when selecting an appropriate algorithm.

Thirdly, this study has been purely mathematical in nature. Account has not been taken for real-world imperfections such as lens aberration, non-flatness or speckle. The authors suggest that these effects are likely to affect all HSAs similarly but recommend further study of the sensitivity of different algorithms to real-world errors.

The fourth point is that the image quality metric used here is MSE. SSIM has seen increased use

in recent years as it more closely corresponds to the human eye behaviour. While recent authors have argued for a more close relationship between MSE and SSIM than thought previously [19] the authors acknowledge that this is a weakness in the HPS method and propose further investigation.

The final point regards the complexity of the HPS method. While in certain situations HPS offers significant performance improvements over rival techniques, the increased complexity and reduced generality will require a greater level of expertise than alternative techniques.

9. Conclusion

This work has presented seven new variants on the Holographic Predictive Search (HPS) algorithm which are summarised in Table 1. By using this, prescient search techniques can be used for a wide range of optical systems in both the far- and mid-field. Different modulation schemes and replay field constraints have all been discussed.

When compared to direct search and simulated annealing algorithms, HPS has been shown to be over $10\times$ faster than its competitors in specific cases at the expense of increased complexity and reduced flexibility. HPS also offers the best convergent error quality. Variants on the HPS algorithm have been presented for a range of optical configurations and the relative advantages and disadvantages presented.

Funding

Disclosures

The authors declare no conflicts of interest.

References

References

[1] G. Milione, T. A. Nguyen, J. Leach, D. A. Nolan, R. R. Alfano, Using the nonseparability of vector beams to encode information for optical communication, Optics letters 40 (21) (2015) 4887–4890.

[2] M. P. Backlund, A. Arbabi, P. N. Petrov, E. Arbabi, S. Saurabh, A. Faraon, W. Moerner, Removing orientation-induced localization biases in single-molecule microscopy using a broadband metasurface mask, Nature photonics 10 (7) (2016) 459.

[3] H. Cheng, P. Li, S. Liu, P. Chen, L. Han, Y. Zhang, W. Hu, J. Zhao, Vortex-controlled morphology conversion of microstructures on silicon induced by femtosecond vector vortex beams, Applied Physics Letters 111 (14) (2017) 141901.

[4] J. Wang, Advances in communications using optical vortices, Photonics Research 4 (5) (2016) B14–B28.

[5] J. Wang, X. Yang, Y. Li, Y. Chen, M. Cao, D. Wei, H. Gao, F. Li, Optically spatial information selection with hybridly polarized beam in atomic vapor, Photonics Research 6 (5) (2018) 451–456.

[6] X. Xie, Y. Chen, K. Yang, J. Zhou, Harnessing the point-spread function for high-resolution far-field optical microscopy, Physical review letters 113 (26) (2014) 263901.

[7] D. W. S. Brian K. Jennison, Jan P. Allebach, Iterative approaches to computer-generated holography, Optical Engineering 28 (6) (1989) 629 – 637 – 9. `doi:10.1117/12.7977012`. URL `https://doi.org/10.1117/12.7977012`

[8] S. Kirkpatrick, C. D. Gelatt, M. P. Vecchi, et al., Optimization by simulated annealing, science 220 (4598) (1983) 671–680.

[9] P. J. Christopher, Y. Wang, T. D. Wilkinson, Predictive search algorithm for phase holography, J. Opt. Soc. Am. A 36 (12) (2019) 2068–2075. `doi:10.1364/JOSAA.36.002068`. URL `http://josaa.osa.org/abstract.cfm?URI=josaa-36-12-2068`

[10] J.-H. Park, Recent progress in computer-generated holography for three-dimensional scenes, Journal of Information Display 18 (1) (2017) 1–12. `doi:10.1080/15980316.2016.1255672`.

[11] P. W. M. Tsang, T.-C. Poon, Y. M. Wu, Review of fast methods for point-based computer-generated holography, Photon. Res. 6 (9) (2018) 837–846.

[12] J. Carpenter, Graphics processing unit–accelerated holography by simulated annealing, Optical Engineering 49 (9) (2010) 095801. `doi:10.1117/1.3484950`.

URL http://opticalengineering.spiedigitallibrary.org/article.aspx?doi=10.1117/1.3484950

[13] M. Frigo, S. G. Johnson, The design and implementation of fftw3, Proceedings of the IEEE 93 (2) (2005) 216–231.

[14] J. W. Goodman, Introduction to Fourier Optics, Third Edition, Roberts and Company Publishers, 2004. doi:10.1117/1.601121.

[15] Y. Huang, E. Liao, R. Chen, S.-T. Wu, Liquid-Crystal-on-Silicon for Augmented Reality Displays, Applied Sciences 8 (12) (2018) 1–17. doi:10.3390/app8122366.

[16] J. L. de Bougrenet de la Tocnaye, L. Dupont, Complex amplitude modulation by use of liquid-crystal spatial light modulators, Appl. Opt. 36 (8) (1997) 1730–1741. doi:10.1364/AO.36.001730.
URL http://ao.osa.org/abstract.cfm?URI=ao-36-8-1730

[17] Z. Wang, a. C. Bovik, H. R. Sheikh, E. P. Simmoncelli, Image quality assessment: form error visibility to structural similarity, Image Processing, IEEE Transactions on 13 (4) (2004) 600–612. doi:10.1109/TIP.2003.819861.

[18] R. W. Gerchberg, W. O. Saxton, A practical algorithm for the determination of phase from image and diffraction plane pictures, Optik 35 (2) (1972) 237–246. doi:10.1070/QE2009v039n06ABEH013642.
URL http://ci.nii.ac.jp/naid/10010556614/

[19] R. Dosselmann, X. D. Yang, A comprehensive assessment of the structural similarity index, Signal, Image and Video Processing 5 (1) (2011) 81–91. doi:10.1007/s11760-009-0144-1.
URL https://doi.org/10.1007/s11760-009-0144-1