Image reconstruction from noisy digital holograms

C.P. Mariadassou, BE, MTech,
Indexing terms: Holograms, Image processing

Abstract: The method of projection onto convex sets (POCS) is used in signal reconstruction to find a function that satisfies a collection of constraints, provided each of these constraints defines a convex set. If the constraints are inconsistent then a modified method of POCS may be used to find the fixed point of the POCS operator. This problem arises in the area of signal reconstruction from noisy multiple-frequency digital holograms, where it is required to compute a signal that satisfies known data subject to the constraint that it has a finite region of support. For noisy data there may exist no such signal and hence the constraint is inconsistent with the known data. Thus a modified POCS method needs to be applied. The paper reviews the method of image reconstruction from digital holograms. Two ways of applying the modified POCS method (called the PICTICS method) are presented. Studies reported here show the effectiveness of the method for image reconstruction from noisy multiple-frequency holograms.

1 Introduction

In this paper we consider the problem of image reconstruction from digital holograms [1, 2]. In the noise-free case the data from multiple-frequency holograms can be combined in an iterative manner using the projection onto convex sets (POCS) [3-5] method for computing a solution that has finite support and that satisfies the known data at each of the multiple frequencies. However, if there is noise in the data then no solution may exist that satisfies the given finite support constraint and the known data. Hence, what we mean by a solution is itself an issue to be addressed. If the data is available at only one frequency then our aim would be to find a solution that has a finite region of support and has the least error with respect to the known data. A similar description of the nature of the solution for the multiple-frequency case is not possible. The problem of image reconstruction from noisy phase of a digital hologram was considered in Reference 6. It was found that reconstruction from quantised phase was better than that from full phase. No attempt was made to make use of the magnitude data. In this paper we propose a method that makes use of the magnitude and full phase of a noisy digital hologram. In what follows we shall give a brief description of the problem we are considering and discuss the suitability of the POCS method for solving it.

Consider a linear shift invariant system. Let

\[ f = h \ast g + n \]  

where \( f \) is the transformed (received) signal, \( g \) is the (transmitted) signal to be recovered, \( h \) is the impulse response of the system, \( n \) is the noise function and \( \ast \) denotes convolution.

If \( g \) is a finite duration signal and \( T^{-1} \) is an inverse operator on \( f \), then

\[ T^{-1} \ast f = g + T^{-1} \ast n \]  

If \( n = 0 \) and \( T^{-1} \) is well behaved, then the solution obtained by applying the inverse operator may be satisfactory. However, owing to the presence of noise it is clear that even if \( T^{-1} \) is well behaved, \( T^{-1} \ast f \) may not be a finite-duration signal. Of course one can truncate the signal according to the previously known duration of the transmitted signal. This solution may not be the best estimate as there may exist some other signal of the same duration that fits the received data better. Our aim is to find a \( g_0 \) that minimises \( d(f, h \ast g) \) for all \( g \in C_f \), where \( C_f \) denotes the set of all signals with known finite duration. Here \( p \) refers to the metric on the space of solutions. Such a solution is referred to as a quasisolution [7]. It will be shown later that if \( d \) is known completely, then the quasisolution will be unique. We shall now discuss some of the earlier approaches for signal reconstruction in the presence of noise.

Many of the earlier approaches are based upon regularisation techniques. These methods assume some knowledge of the statistics of the signal or noise. For example, it may be assumed that the noise is stationary and that the distribution of noise or its power spectrum is known a priori. Linear filters, such as Wiener filtering, or nonlinear filters, such maximum entropy estimation, were developed on this basis. Recently, another approach based on the method of POCS has been proposed [8, 9], in which it is required to compute a feasible solution satisfying the available data to within a certain degree. For example, it may be assumed that the magnitude of the noise is bounded. In other words \( |n| = |f - h \ast g| \leq d \), where \( d \) is a known positive real value. This method works well if \( d \) is known accurately. If the estimated value of \( d \) is more than the actual value, then the reconstructed signal will be a conservative estimate in that the variance of the estimate may be large. If the estimated value of \( d \) is lower than the actual value, then the method of POCS may not converge at all. Another possible approach is to assume that \( |n| < r \), that is the norm of the noise, which is usually the squared sum of the noise at individual points, is bounded [9]. In both the above mentioned cases the
The problem may have no solution at all if the data $f_2$ is corrupted with noise. Let us consider first the noise free case. It is trivial to show that
\[ C = \{ g : f_1 = h \ast g \text{ for given } f_1, h, \text{ and } \lambda = \lambda_0 \} \]
is a closed convex set [11]. Let $P_i$ denote the projection onto the set $C_i$ and $P_{fi}$ denote the projection onto the set $C_{fi}$. By the method of POCS we can iteratively compute $g \in C_{fi} \bigcap \cdots \bigcap C_1$ as follows:
\[ g_{n+1} = P_{fi} P_{fi+1} \cdots P_{f2} P_1 g_n \]
where the initial estimate is any $g_0 \in H$. Notice that if the data is noisy, that is, if the known data is
\[ f_1 = h \ast g + n \]
where $n$ is a noise function, then $C$ may be empty, and hence the above mentioned iterative procedure may not converge. However, if we choose an initial estimate $g_0$ in $C_{fi}$, then strong convergence can be shown if the result of successive application of the operator
\[ 1 + s(P_{fi}, P_{fi+1}, \ldots, P_{f2}, P_1) \quad 0 < s < 1 \]
is bounded. Lemma 1 assures strong convergence if any subsequence is convergent strongly. Note, however, that the exact nature of the solution is not well understood, although an attempt has been made in the literature to give an intuitive feel for the nature of the solution [13]. A precise characterisation of the nature of the solution is possible if only two convex sets are involved. For example, it can be shown [3] that the distance between the dual fixed points of the operators $P_{fi}, P_1$ is the minimum distance between the two convex sets $C_{fi}$ and $C_1$. In other words, let $x$ be a fixed point of $P_{fi}, P_1$. Then
\[ \| x - P_{fi} x \| = \min \| y - z \| \]
where $y, z \in C_{fi}, C_1$. We can also show that the operator $1 + s(P_{fi}, P_1 - 1)$, $0 < s < 2$, is a nonexpansive operator from $C_{fi} \to C_{fi}$ and that the sequence generated by the following iterative procedure converges strongly:
\[ g_{n+1} = g_n + s(P_{fi}, P_1) g_n - g_n \]
Recall that $T_1 = 1 + s_1(P_1 - 1), 0 < s_1 < 2$, is also a nonexpansive operator. Hence the application of the operator $1 + s(T_1, T_i - 1)$ assures strong convergence. Using a value of $s_1$ that is not necessarily unity helps in speeding up the convergence. $P_{fi}$ cannot be replaced by a corresponding $T_{fi}$ since then the result of every iteration may not belong to $C_{fi}$ and hence strong convergence cannot be assured. Note also that $P_{fi}$ is a linear operator. We intend to show that the operator $1 + s(P_{fi}, P_1 - 1)$ is equivalent to $1 + s(T_{fi}, T_i - 1)$. In other words, for every choice of $s$ and $s_1$,
\[ 1 + s(T_{fi}, T_i - 1) = 1 + s(P_{fi}, P_1 - 1) \]
where
\[ s' = ss_1 \]
This may be shown as follows: for any $g \in C_{fi}$,
\[ g + s(P_{fi}, T_i) g = g + s(P_1 g, s_1(P_2 g - g) - g) \]
Since $P_{fi}$ is a linear operator and for $g \in C_{fi}, P_1 g = g$, the above equation reduces to
\[ g + s(P_{fi}, T_i) g = g + s(P_1 g, P_2 T_i g - g - g) \]

\[ = g + ss(P_1 g, P_2 g - g) \]

\[ \text{IEEE PROCEEDINGS, Vol. 137, Pt. F, No. 5, OCTOBER 1990} \]
From the above equation the desired result trivially follows. Moreover, the fixed points of \( I + s(P_{f_{1}} P_{1}) \) are just the fixed points of \( P_{f_{1}} P_{1} \), which in turn are just the fixed points of \( P_{f_{1}} P_{1} \).

Consider now the operator \( 1 + s(P_{f_{1}} P_{f_{1}} T_{2}, \ldots, T_{n}) \). Here \( T_{1}, T_{2}, \ldots, T_{n} \) are applied in sequence. It would be better if the operators could be applied in parallel and combined in a suitable manner. This helps to speed up convergence in a multiprocessor environment. More importantly it lays equal emphasis on every set of available data. If the operators are applied in a sequence, unequal and unknown relative emphasis is likely to be placed on each of the available data. The following lemma suggests a method of applying the operators in parallel.

**Lemma 2:** Let \( R_{1}, R_{2}, \ldots, R_{n} \) be a collection of non-expansive operators. Let

\[
R = a_{1} R_{1} + a_{2} R_{2} + \cdots + a_{n} R_{n}
\]

where \( a_{1} + a_{2} + \cdots + a_{n} = 1 \), and \( a_{1}, a_{2}, \ldots, a_{n} \) are all greater than zero. \( R \) is a nonexpansive operator. Furthermore, if \( R_{i} \) for some \( i \), \( 1 \leq i \leq n \), is a contraction mapping, then \( R \) is a contraction mapping.

Let us consider the nature of the solutions provided by the two operators \( P_{f_{1}} T_{1} T_{2} \ldots, T_{n} \) and \( P_{f_{1}}(a_{1} T_{1} + a_{2} T_{2} + \cdots + a_{n} T_{n}) \). Here, as in the rest of the paper, \( a_{1}, a_{2}, \ldots, a_{n} \) denote a set of real numbers satisfying the constraints specified in Lemma 2. The two operators may not necessarily have the same fixed points. However, if the set \( C_{f_{1}} \cap C_{1} \) is nonempty, then we can show that the set of fixed points of the two operators is identical and is precisely the set \( C_{f_{1}} \cap C_{1} \). It has been shown earlier that the set of fixed points of the operator \( P_{f_{1}} T_{1} T_{2} \ldots, T_{n} \) is the set \( C_{f_{1}} \cap C_{1} \). Our aim is to show that the set of fixed points of \( P_{f_{1}}(a_{1} T_{1} + a_{2} T_{2} + \cdots + a_{n} T_{n}) \) is also \( C_{f_{1}} \cap C_{1} \). It is obvious that \( C_{f_{1}} \cap C_{1} \) is a subset of the fixed points of \( P_{f_{1}}(a_{1} T_{1} + a_{2} T_{2} + \cdots + a_{n} T_{n}) \), since every \( g \in \cap \cap C_{1} \) is also a fixed point of \( P_{f_{1}}(a_{1} T_{1} + a_{2} T_{2} + \cdots + a_{n} T_{n}) \). The containment in the other direction can be shown as follows. Let \( y \) be a fixed point of \( P_{f_{1}}(a_{1} T_{1} + a_{2} T_{2} + \cdots + a_{n} T_{n}) \) and \( x \) an element of \( C_{f_{1}} \cap C_{1} \). Consider \( ||x - y|| \). Since \( x \) is also fixed point of \( P_{f_{1}}(a_{1} T_{1} + a_{2} T_{2} + \cdots + a_{n} T_{n}) \),

\[
||x - y|| = ||P_{f_{1}}(a_{1} T_{1} + a_{2} T_{2} + \cdots + a_{n} T_{n}) x - P_{f_{1}}(a_{1} T_{1} + a_{2} T_{2} + \cdots + a_{n} T_{n}) y||
\]

(12)

Again, since \( P_{f_{1}} \) is a linear operator, the above equation can be written as

\[
||x - y|| = a_{1} ||P_{f_{1}} T_{1} x - P_{f_{1}} T_{1} y||
\]

(13)

\[
+ a_{2} ||P_{f_{1}} T_{2} x - P_{f_{1}} T_{2} y||
\]

(14)

\[
+ \cdots + a_{n} ||P_{f_{1}} T_{n} x - P_{f_{1}} T_{n} y||
\]

(15)

\[
\leq a_{1} ||P_{f_{1}} T_{1} x - P_{f_{1}} T_{1} y||
\]

(16)

\[
+ a_{2} ||P_{f_{1}} T_{2} x - P_{f_{1}} T_{2} y||
\]

(17)

\[
+ \cdots + a_{n} ||P_{f_{1}} T_{n} x - P_{f_{1}} T_{n} y||
\]

(18)

\[
= ||x - y||
\]

(19)

(20)

(21)

(22)

(23)

\[
H(u, v) = \begin{cases} 
\exp (j k z^2 - (\alpha x)^2 - (\alpha y)^2)^{1/2} & \text{for } -1 < k < 1 \\
0 & \text{otherwise}
\end{cases}
\]
Since $h(x, y)$ is band-limited, $f(x, y)$ is also band-limited. However, due to the finite size of the aperture (object), $g(x, y)$ has a finite region of support, and hence $G(u, v)$ is not band-limited. Using the method of POCS, it is possible to recover $g(x, y)$ if $f(x, y)$ is known. The following iterative procedure converges to a unique solution:

$$g_{n+1}(x, y) = P_f P_s g_n(x, y)$$

(24)

where $P_f$ and $P_s$ can be defined as follows. Let $g'(x, y) = P_f g(x, y)$ and $G(u, v)$ and $G_g(u, v)$ represent the Fourier transforms of $g'(x, y)$ and $g(x, y)$, respectively. Therefore

$$G(u, v) = \begin{cases} f(u, v)/H(u, v) & \text{for } u^2 + v^2 < 1/2^2 \\ G_g(u, v) & \text{otherwise} \end{cases}$$

(25)

$g'(x, y)$ can be computed from $G(u, v)$.

That the solution is unique can be shown as follows. Let $g(x, y)$ be a fixed point of $P_f$. Since $g(x, y)$ has a finite region of support, $G(u, v)$ is analytic [14]. We have

$$G(u, v) = F(u, v)/H(u, v) \quad \text{for } u^2 + v^2 < 1/2^2$$

(26)

Hence, by analytic continuation $G(u, v)$ is unique.

In reality, $f(x, y)$ is known only at a finite set of sampling points and hence $F(u, v)$ cannot be known accurately. Also, the received data is corrupted by noise and hence eqn. 20 must be modified as

$$f(x, y) = h(x, y) + g(x, y) + n(x, y)$$

(27)

where $n(x, y)$ is some additive noise function. We resort to computing only the quasiosolution, as computing an actual solution may not be feasible. The following claim can be made.

**Lemma 3:** If the solution set is compact and $g \ast h = 0$ has only the zero solution, then a quasiosolution to the problem of computing $g$ from $f$ exists and is unique.

Let the solution set $C_f$ consist of all bounded functions with finite regions of support. Consider the operator $P = P_f, P_s$

(28)

where $P_f$ and $P_s$ are as defined earlier. Since $P$ is a nonexpansive operator on a closed, convex and bounded set, there exists a fixed point for the operator $P$. Hence, from the above lemma the solution is unique. From lemma 2 and the arguments that followed it, the iterative algorithm given below converges to a fixed point:

$$g_{n+1}(x, y) = g_n(x, y) + s' (Pf g_n(x, y) - g_n(x, y))$$

(29)

where $0 < s' < 2$.

Now let us suppose $f(x, y)$ is known only on a finite set of points $I_f$. In this case no unique solution exists to the problem, even in the absence of noise. Let $C_{f_1}$ denote the set of all object field distributions $g(x, y)$ that give rise to the hologram data known on the set $I_f$. It is easily shown that $C_{f_1}$ is a closed convex set. Let $P_{f_1}$ denote the projection onto $C_{f_1}$ and define $P = P_f, P_{f_1}$. Then $P$ is a nonexpansive operator on the set of functions with known region of support and the iterative procedure of eqn. 29 converges to a fixed point of $P$. Notice also that any fixed point of $P$ is a quasiosolution to the problem. In other words, if $g(x, y)$ is a fixed point of $P$, then

$$\rho(g(x, y), C_f) = \rho(g(x, y), C_{f_1})$$

(30)

If the hologram data is known for various wavelengths $\lambda_1, \lambda_2, \ldots, \lambda_n$ on a set $I_f$, then each such hologram defines a convex set of functions that give rise to a hologram known on a finite set $I_f$. Let us denote as $C_{f_1}, C_{f_2}, \ldots, C_{f_n}$ the set of all functions $g(x, y)$ that gives rise to holograms $f_1(x, y), f_2(x, y), \ldots, f_n(x, y)$ respectively, and let $P_{f_1}, P_{f_2}, \ldots, P_{f_n}$ denote the corresponding projection operators. Define

$$P' = P_f s_1 T_{f_1} + s_2 T_{f_2} + \ldots + s_n T_{f_n}$$

(31)

where

$$s_1 + s_2 + \ldots + s_n = 1, \quad s_i > 0 \quad \text{for } 0 < i < n$$

and

$$T_{f_1} = 1 + s_1 (P_{f_1} - 1) \text{ with } 0 < s_1 < 2$$

It has been shown [4] that $T_{f_1}$ is a nonexpansive operator. From lemma 1 and lemma 2 it is seen that the following iterative procedure converges:

$$g_{n+1}(x, y) = sg_n(x, y) + (1 - s) P' g_n(x, y)$$

(32)

where $g_0(x, y)$ is any bounded signal with known finite support and $0 < s < 1$. Similarly, define

$$P' = P_f s_1 T_{f_1} + s_2 T_{f_2} \ldots, T_{f_n}$$

The following iterative procedure converges to a fixed point:

$$g_{n+1}(x, y) = sg_n(x, y) + (1 - s) P' g_n(x, y)$$

(33)

where $g_0(x, y)$ and $s$ are chosen as mentioned before.

Simulation studies were conducted using both eqns. 32 and 33. The results are given in the next Section.

4 Simulation studies

The purpose of the simulation studies is threefold, namely:

(i) to show the convergence of the PONICS method

(ii) to compare the methods of POCS and PONICS for different cases

(iii) to compare the parallel and sequential implementations of the PONICS method.

There are two types of noise: Gaussian noise (corresponding to the case of bounded noise energy) and uniformly distributed noise (corresponding to the case of bounded noise magnitude). In the present study we consider image reconstruction in the presence of uniformly distributed noise where the magnitude of the random noise is bounded.

In general, the POCs method operating on noisy data diverges within a few iterations, therefore the iterative reconstruction has to be terminated before reaching a given quality. The number of iterations that can be performed with the POCs method will be fewer for higher noise levels. Use of data at a greater number of frequencies does not solve this problem. On the other hand, our studies show that the PONICS method converges to a solution most of the time. Moreover, we find that the number of iterations required to obtain a given quality using the PONICS method is in general smaller than the number of iterations required using the POCs method.

In the following we consider an object plane of size $128 \times 128$ pixels enclosing an object of size $64 \times 64$ pixels. The receiver array consists of $128 \times 128$ points or sensor elements. The spacing between object and receiver planes is 2000 units and the spacing between adjacent sensors in the receiver plane is 0.54, where the wavelength $\lambda$ of the incident wave is 0.25 units. Uniformly distributed random noise is used to generate complex noise samples.
to be added to the received sensor signal samples. The signal-to-noise ratio in decibels is computed as the ratio of signal energy (sum of the squares of the signal amplitudes from all the sensors) and the noise energy (sum of the squares of amplitudes of the noise samples). Sparse data is generated by considering the output from selected sensors of the 128 × 128 elements on the receiver plane. For example, a 32 × 32 point sparse data set is obtained by considering every fourth point on the receiver plane.

Fig. 1 shows the object used in our simulation studies.

The effects of noise and multiple-frequency data are illustrated in Fig. 2 which shows reconstructed images after 50 iterations from PONICS and POCS using only 64 × 64 points of the received data. The intermediate points are set to zero. In Figs. 2–5, the iterations for the POCS method were terminated just before divergence, whereas the PONICS method was applied for 50 iterations. We find that the quality of the reconstructed image does not improve significantly beyond about 20 iterations. Fig. 2 shows that the quality of the reconstructed images using the PONICS method are as good as or better than those obtained using the POCS method.

**Fig. 2** Comparison of POCS and PONICS methods for reconstruction of images from noisy sensor array data with the data collected from a 64 × 64 array at different frequencies. Comparison is made for two different SNRs. Uniformly distributed noise is used.

**Fig. 3** Comparison of POCS and PONICS methods for reconstruction of images from noisy sensor array data with the data collected from a 32 × 32 array at different frequencies. Comparison is made for two different SNRs. Uniformly distributed noise is used.

**Fig. 4** Comparison of POCS and PONICS methods for reconstruction of images from noisy sensor array data with the data collected from a 16 × 16 array at different frequencies. Comparison is made for two different SNRs. Uniformly distributed noise is used.
The main advantage of the PONICS method is that the algorithm does not seem to diverge. This trend is evident for different noise levels and also for different numbers of frequencies. As expected, the degradation in the reconstructed image is greater for lower SNR. As the number of frequencies at which data is collected is increased, then the quality of the reconstructed image is improved significantly.

Fig. 5 Comparison of sequential and parallel PONICS methods for image reconstruction from noisy multiple frequency hologram data using a 32 x 32 array for two SNR values (0 dB and -10 dB)

Similar conclusions can be drawn even when the data is collected at fewer points on the receiver. Figs. 3 and 4 show the images reconstructed with 32 x 32 points and 16 x 16 points on the receiver plane. The Figures illustrate the sparsity of receiver data can be compensated by collecting data at a greater number of frequencies. The performance of sequential and parallel implementations of the PONICS method is illustrated in Fig. 5. The comparison is made for data collected at multiple frequencies using an array of 32 x 32 elements. The results are given for SNR = 0 dB and SNR = -10 dB. The results show that there is no appreciable difference in the two implementations.

5 Conclusions

In this paper we have applied a modified projection-onto-convex-sets algorithm method for image reconstruction from noisy multiple-frequency holograms. We also developed a method of applying the operators in parallel and showed that, if the constraints are consistent and the convex sets have a nonempty intersection, then both the sequential and parallel methods have the same set of fixed points. This result will not hold good if the constraints are inconsistent. Simulation studies conducted using both the operators showed no appreciable difference in the results obtained using either of the operators. We have also pointed out that a direct implementation of the POCS method may not work since the results of successive iterations may begin to diverge.

The significance of these studies is that it is possible to reconstruct an image from sparse noisy sensor array data. A small number of sensors reduces the receiver complexity and the sparsity of data is compensated for by collecting data at multiple frequencies. Noisy data may produce empty intersections of sets and hence the POCS method may diverge. The modified method enlarges the sets so that the intersection is no longer empty. Hence the proposed PONICS method can be used even for noisy data, although the projection operators are a little more complicated.

The simulation studies reported here are applicable for a simplified sensor array imaging set-up in which plane waves are assumed to be incident on a planar object. In practice, the spherical nature of the wavefront, diffraction effects due to object shapes and medium disturbances, all make the task of image reconstruction much more difficult than has been presented in this paper.

6 References

7 TIKHONOV, A.N. and ARSENIN, V.Y.: 'Solutions to ill posed problems' (John Wiley & Sons, New York, 1977)
15 YOSHIDA, K.: 'Functional analysis' (Springer Verlag, New York, 1975)