國立台灣大學電資學院光電工程學研究所

碩士論文

Graduate Institute of Photonics and Optoelectronics

College of Electrical Engineering and Computer Science

National Taiwan University

Master Thesis

以改良式基因演算法的交配機制設計繞射光學元件之研究

The Study of an Improved Crossover Mechanism in the Genetic Algorithm for Design of Diffractive Optical Elements

林家鴻

Lin, Chia-Hung

指導老師：林晃巖 博士，徐巍峰 博士

Advisor: Hoang Yan Lin, Ph.D. and Wei-Feng Hsu, Ph.D.

中華民國97年7月

July, 2008

誌 謝

從三年前決定辭去工作準備研究所入學考， 到今天口試結束，論文完成， 對於當初已離開校園多年的筆者而言， 若沒有一路走來適時出現貴人們的相助， 這絕對不是一件簡單的事。

首先，在準備考試的一年中， 兩位補習班的老師，孫超群老師及劉明昌老師， 對於筆者近乎永無止盡的發問， 總是能清楚地爲筆者解惑， 這對於當初連 e^x 該如何微分都已經忘的一乾二淨的筆者而言， 眞的是一個好的開始。 而在準備過程中所認識的幾位同伴， 啓超、耀淵以及交大電子所的欣哲、彥忠， 能不嫌棄和筆者一起討論學習，更是筆者能夠順利通過入學考試的貴人。

進入研究所後，首先要感謝林晃巖老師願意給基礎不佳的筆者一個機會，讓筆者加入本實驗室。 並且包容筆者的任性，給筆者在學習的過程中充分的自由，包括研究題目的選擇。 而在學習的過程中，更是慶幸能和俊翔成爲同學， 有了他的指教，筆者才能更深入地思考所習得的知識。 感謝顧輝在實驗室中、日常生活上所給予的許多幫助， 也因爲顧輝，筆者才有機會接受到台北科技大學光電所徐巍峰老師無私的指導。 謝謝于堂不厭其煩地介紹與討論，使筆者有機會一窺另一領域的奧妙。 國忠對於人生規劃上的分析和建議，也讓筆者受益良多。 而，年輕的夥伴，婷玉、宇廷和東儒，也爲實驗室帶來歡笑和朝氣。

在論文的完成上，要感謝三位口試委員的指教， 林晃巖老師、徐巍峰老師以及南台科技大學光電系林正峰老師。 林晃巖老師總是能夠指出關鍵的問題，使筆者進一步地思考， 並告訴筆者何處該加強立論的邏輯，以增加論文的說服力。 徐巍峰老師知無不言言無不盡的指導，讓筆者完成了論文最重要的部分， 而，徐老師鉅細靡遺地閱讀修改筆者的論文，也讓論文的文字更加地精確。 另外，感謝林正峰老師在口試過程中提出的建議，使得論文能夠有更高的完成度。

感謝貞宏、顧輝、昭德、俊翔以及冠宇的在口試前具體的建議及鼓勵， 而包括阿

達、揚智在內的徐老師的研究生也撥冗陪筆者練習，並提供了很多準備時思考的方向，在此一併致謝。 謝謝顧輝、于堂、俊翔和婷玉特地來參加筆者的口試，爲筆者打氣，當筆者的聽衆， 也謝謝于堂在過程中爲筆者記錄。 此外，在口試場地和飲食的準備上，更要感謝顧輝和宇廷的大力相助。

最後，筆者要感謝多年來一直陪伴在筆者身邊，從女友成爲筆者妻子的紘子。 沒有她的陪伴，筆者無法從疾病的陰影走出來。 有了她的支持，筆者才能夠下定決心重回校園。 而也因爲她的體貼，筆者才能心無旁騖地專心於學業，順利完成論文。 感謝紘子，有了她，這一切才算完整；有了她，未來才算完整。

林家鴻

謹誌於台北

西元 2008 年 7 月

中 文 摘 要

在本篇論文中，我們以基因演法算來設計繞射光學元件。 在所使用的演算法中，爲了深入瞭解交配機制，我們排除突變機制的影響， 並且提出了一種新式的交配運算子。在設計此新式的交配運算子時，我們利用了繞射光學元件具有明確目標光場圖形的性質， 因此，我們亦將其命名爲「 目標圖形導向交配運算子 」。 藉由利用目標光場圖形，此運算子能夠刺激整個族群，提供演化時的動力。

在本論文中，我們探討了該運算子的特性。 藉由使用不同的初始族群，我們驗證了該演算法對初始條件的強韌性。 我們觀察到繞射效率及均方根誤差的快速收斂性質； 同時，我們也發現在訊雜比的演化需要較多的世代數。 另外，在演化結束後，我們發現在整個族群的個體表現上， 繞射效率與訊雜比無法達到同時優化，而必須在其中一者有所犧牲。

我們將所提出的方法應用於二階、多階及連續相位調變繞射光學元件的設計上。 我們也針對了以本方法所設計出來的結果，與其他演算法的結果進行比較。 例如，與疊代傅立葉轉換演算法與簡單型基因演算法的結果比較。 在與使用不同突變率之簡單型基因演算法的比較上， 我們提出的演算法可以在 5,000 世代時達到相似的結果， 這相當於只需要簡單型基因演算法 20 分之 1 的世代數。 而在與疊代傅立葉轉換演算法的比較上，本方法所得結果的繞射效率與均方根誤差與其相似， 並且訊雜比可以達到 2 倍以上。

Abstract

In this thesis, a method is proposed to design the diffractive optical elements (DOEs), in which the genetic algorithm (GA) is used with a novel crossover operator. The mutation mechanism is excluded from the method. The novel crossover operator is derived by taking advantage of the existence of the well-defined target in the DOE design and it is therefore entitled the TArget-oriented CrossOver (TACO) operator. By utilizing the target property in the DOE design, the TACO operator stimulates the evolution of solution toward the global maximum in the solution space.

The properties of the TACO operator are studied. The robustness against different intial populations is verified. The fast convergence of efficiency and root-mean-square error (RMSERR) is observed. On the other hand, the growth of signal-to-noise ratio (SNR) is relatively slow. The trade-off between the efficiency and the SNR of the final population is also observed.

The proposed method is applied to the design of the binary, multi-level, and continuous phase DOEs. The performance is compared with that of other optimization algorithms, such as the iterative Fourier transform algorithm (IFTA) and the simple genetic algorithm (SGA). Compared with the SGA using different mutation rates, the proposed method yields comparable results within a total of 5,000 generations, only 1/20 of that used in the SGA. Also, the efficiency and the RMSERR of the result yielded by the proposed algorithm are comparable to the IFTA, while the SNR is twice as high as that of the IFTA.

Contents

List of Figures

List of Tables

Chapter 1

Introdcution

1.1 Diffrctive Optical Elements (DOEs)

Literally, a DOE is an optical device which utilizes the diffraction nature of light wave to produce the desired light field. By “diffraction”, “any deviation of light rays from rectilinear paths which cannnot be interpreted as reflection or refraction” is meant, as defined by Sommerfeld [1].

To accomplish the task of distributing the light field in a desired way, the size, the material and the surface profile of DOEs need to be carefully designed and chosen. For example, the feature size must be compatible with the wavelength for the diffraction phenomenon to have significant effect. (Here, “feature size” means “the size of the minimum aperture of the DOE”, as shwon in Fig. 1.1) The refractive index of a material should be properly selected to provide the desired optical path difference (OPD). Also, the surface profile has to be carefully calculated and arranged periodically, for instance, in order to produce the desired diffraction order of the light.

There are many ways to catagorize the DOEs, and one of them is by the way of

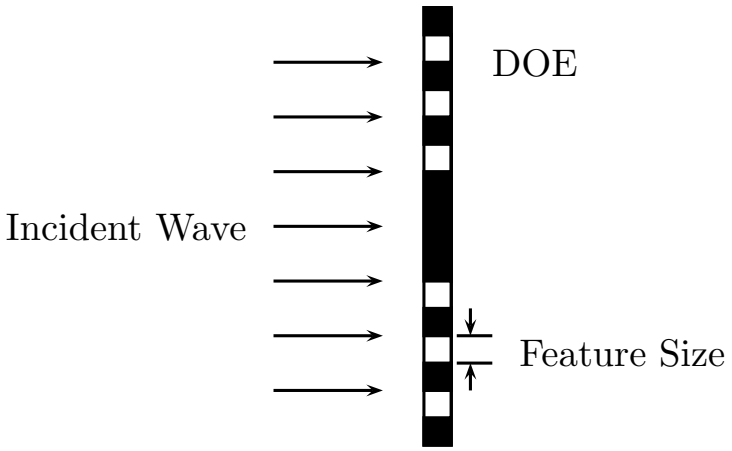

Figure 1.1 Feature size of a DOE

modulation. According to the modulation methods, DOEs can be catagorized into the phase modulation DOEs and the amplitude modulation DOEs. The difference between the two kinds of DOEs is that for amplitude modulation DOEs, the amplitude of the incident wave is absorbed or reflected by the DOEs. The phase modulation DOEs, on ther other hand, will change the phase and allow the amplitude after the DOEs to be the same as the incident light, theoritically. Therefore, the phase modulation DOEs have higher diffraction efficiency than the amplitude modulation DOEs.

The history of design of DOEs may be dated back to that the Fresnel zone plate was designed by Rayleigh. As the advance of the modern computation technology, the complex phase or amplitude modulation information can be calculated and encoded on a transparency, such as computer generated holograms (CGHs).

Further, thanks to the advanced semiconductor process technology, multi-level qunatized DOEs can be realized to approach the continuous profiles which are difficult to manufacture under the size of the wavelength. The semiconductor process technology also makes the manufacture of the pixel-wise encoded kinoforms easiler. Figure 1.2 shows the different kinds of DOEs.

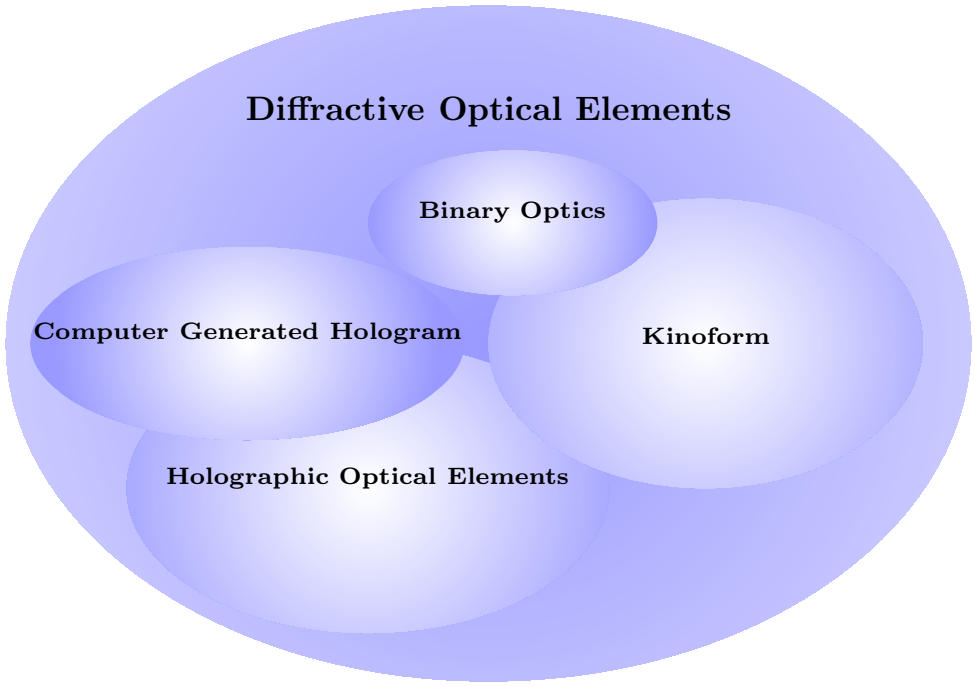

Figure 1.2 Catagory of DOEs

There are many applications of DOEs, from imaging systems to beam shaping of lasers [2]. For example, a laser beam with a uniform field distribution is sometimes needed in applications such as microlithography [3]. The illuminators may utilize DOEs to homogenize the input laser beam with the Gaussian profile into a flat-top output beam [4, 5]. Figure 1.3 shows the schematic diagram of the application of DOEs to beam shaping of lasers.

1.2 Design of DOEs

The fundamental theory of DOEs is the diffraction theory, which predicts the propagation of wave when light passes through the free-space system containing the DOEs. The formalism of the diffraction theory can be derived from the Maxwell's equations and

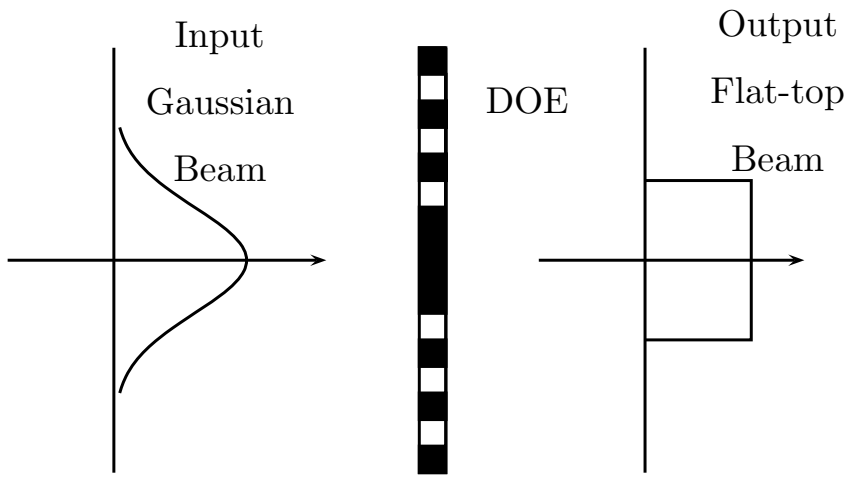

Figure 1.3 Schematic diagram of beam shaping of lasers

the may possess different complexity, depending on the feature size of the DOE and the wavelength of the incident light. In the paraxial region, the formulation can be simplified while giving results with sufficient accuracy. Usually the scalar diffraction theory will be adequate. However, in a real DOE design problem, except for the analytical-type DOEs, it is not possible to realize the calculation in an analytical way. Instead, modern computers together with numerical methods are used to carry out these tasks.

Besides the diffraction theory, an optimization method is needed in order to design the DOEs with desired performance. Some methods use algorithms which unidirectionaly conduct the transformation from incident field to target field, such as direct binary search (DBS) [6, 7] and simulated annealing (SA) [8, 9] algorithms. On the other hand, there are algorithms which transform the incident field and the target field bidirectionally, such as Gerchberg-Saxton (GS) algorithm [10], which is also known as iterative Fourier tansformation algorithm (IFTA).

There are other algorithms which sort to the random process in such a way that the whole process emulates the evolution of natural species, namely “survival of the fittest”

proposed by Darwin. Genetic algorithm (GA) [11] is the most well-known method of this kind.

Although the fundamental theory is the same for all kinds of DOEs, the optimization methods to design DOEs are chosen according to the applications and the desired performance.

1.3 Motivation

The theory and application of GA has been widely studied since its debut in the late 1960s. There are several merits of optimization using GA. For example, the implementation of GA is readily to be parallelized, which then reduces a great amount of search time. Also, the feature of parallelization enables GA to search in multiple points of the solution space for the global optimum. These merits make GA a superior optimization algorithm and apply GA widely to different areas.

The study on DOE design using GA is relatively less compared with other methods such as IFTA. The application of GA and the combination of GA and SA in DOE design have been reported [11, 12]. There are studies on how the crossover and mutation mechanisms affect GA in DOE design [13, 14]. In [13], it is concluded that, without mutations, the regular crossover patterns only had limited effect on the growth of performance. On the other hand, although the mutation mechanism enhances the searching of the global optimum, the expense of computation resource is inevitable [14].

To further increase the performance of GA in DOE design, in this thesis, the GA used in [13] will be improved for DOE design. In particular, a method is proposed in which a novel crossover operator is used in the GA instead of the regular exchange crossover pattern and the mutation mechanism. The proposed method results in the DOE profiles

with superior performance in a relatively short time.

1.4 Thesis Orginization

In this study, the output diffractions in the optical far field of the DOEs are calculated by the scalar diffraction theory which is briefly illustrated in Chapter 2. Based on the far-field diffraction formula, the numerical model of DOEs in the optimization processer is described. In particular, the encoding of DOEs and the relative parameters and merit functions for evaluating performance of DOEs are discribed. Chapter 3 introduces the basis and implementations of GA, including the fundamental operators and the several kinds of implementations. The results of the algorithm used in [13] and [14] are summarized. In Chapter 4, the proposed crossover operator, entitiled Target-Oriented CrossOver (TACO) operator, is derived. The relative parameters used in the algorithm with the TACO operator are determined. The simulation results of DOEs designed by the proposed method are presented in Chapter 5. These results are compared with the results of DOEs designed using IFTA. Finaly, the conclusion of this thesis and possible future study direction will be given in Chapter 6.

Chapter 2

Fundamental Theory of DOEs—Diffractive Optics

2.1 Diffraction Theory

The diffraction theory begins from Maxwell's equations. Depending on the complexity of the problem, the formulation may have different appearances, which are essentially divided into two mainstreams, namely the vector theory and the scalar theory. In the design of DOEs, the feature size of the DOEs and the distance between the observation plane and the DOEs play an important role in deciding whether the vector theory or the scalar theory is used.

When the feature size of the DOEs is smaller than the magnitude of a wavelength of the incident wave or the near field after the DOEs is of interest, the vectory theory which sorts to rigorously solving the Maxwell's equations is needed to calculate the diffractive field. On the other hand, when the diffracting aperture is much larger than the wavelength of the incident wave and the observed output diffrated field is far from the aperture, the

analysis can be further simplified and the scalar theory yields accurate results. [15]

In the thesis, the target plane of the DOEs illuminated by a monochromatic coherent light with uniform amplitude is in the optical far field, and therefore the scalar theory can give accurate predictions. The formulation of the scalar theory is described in the following sections.

2.1.1 Huygens-Fresnel Principle

The fundation of the scalar diffraction theory is the Huygens-Fresnel principle. As shown in Fig. 2.1, the diffracting aperture Σ lies in the ξ–η plane and the output diffractive field is to be observed on the x–y plane. According to the Rayleigh-Sommerfeld diffraction formula [15], the Huygens-Fresnel principle can be expressed as

$$U(P_0) = \frac{1}{j\lambda} \iint_{\Sigma} U(P_1) \frac{e^{jkr_{01}}}{r_{01}} \cos\theta ds, \tag{2.1}$$

where $U(P_0)$ and $U(P_1)$ represent the fields after the aperture and the diffractive field, respectively, and λ is the wavelength of the monochromatic wave. It should be noted that the expression of (2.1) is for monochromatic waves only, which is the basic assumption of this thesis. In (2.1), $\cos\theta$ can be expressed by the distance between source point P_0 and field point P_1 and the distance between ξ–η plane and x–y plane:

$$\cos\theta = \frac{z}{r_{01}}. \tag{2.2}$$

Therefore, (2.1) can be written into

$$U(x,y) = \frac{z}{j\lambda} \iint_{\Sigma} U(\xi,\eta) \frac{e^{jkr_{01}}}{r_{01}^2} d\xi d\eta, \tag{2.3}$$

with

$$r_{01} = \sqrt{z^2 + (x-\xi)^2 + (y-\eta)^2}. \tag{2.4}$$

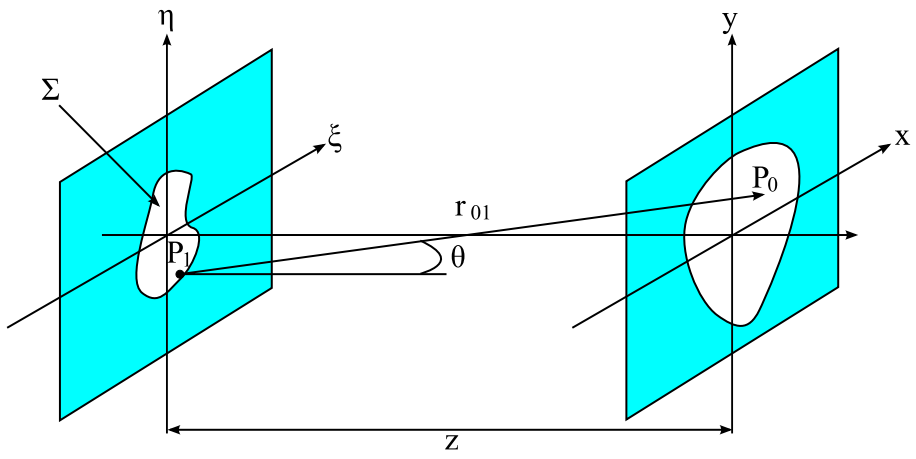

Figure 2.1 Geometry for explanation of the Huygens-Fresnel principle

2.1.2 Fresnel Diffraction

The Huygens-Fresnel priciple expressed in (2.3) can be simplied under certain conditions into the Fresnel diffraction. First, for $z \gg (x-\xi)$ and $z \gg (y-\eta)$, one can approximate r_{01} using binominal expansion with the first two terms such as

$$r_{01} \approx z\left[1+\frac{1}{2}\left(\frac{x-\xi}{z}\right)^2+\frac{1}{2}\left(\frac{y-\eta}{z}\right)^2\right]. \tag{2.5}$$

Upon substituting (2.5) into (2.3), it should be noted that r_{01} appears two times in (2.3), and it is not necessary that both of them be replaced with (2.5). For the r_{01}^2 term, considering its appearing in the denominator and the effect of square, it is obvious that the effect of the error is so small that the result will be the same if one just replaces r_{01} with z. For the r_{01} in the phase term, the error arising from substituting z for r_{01} is more sensitive and needs more care. This is caused by the multiplying of r_{01} by the wave number k. The value of wave number k is usually of the order of 10^7 1/m for the visible light and amplifies the error by the same order accordingly. Besides, the value of exponetial may change greatly even with a little change of the phase. Therefore, instead of merely replacing r_{01} in the phase term with z, (2.5) is used in order for sufficient accuracy.

Then, the substitution yields

$$U(x,y) = \frac{e^{jkz}}{j\lambda z} \iint_{-\infty}^{\infty} U(\xi,\eta) e^{j\frac{k}{2z}\left[(x-\xi)^2+(y-\eta)^2\right]} d\xi d\eta. \tag{2.6}$$

Further, by factoring the term $e^{\frac{jk}{2z}(x^2+y^2)}$ out of the integral, (2.6) is written into

$$U(x,y) = \frac{e^{jkz}}{j\lambda z} e^{j\frac{k}{2z}(x^2+y^2)} \iint_{-\infty}^{\infty} \left\{U(\xi,\eta) e^{j\frac{k}{2z}(\xi^2+\eta^2)}\right\} e^{-j\frac{2\pi}{\lambda z}(x\xi+y\eta)} d\xi d\eta, \tag{2.7}$$

which, as well as (2.6), is referred to as the Fresnel diffraction integral. The detail analysis of the accuracy of the Fresnel approximation can be found in [15].

2.1.3 Fraunhofer Diffraction

To further simplify (2.7), we assume that the far field condition is satisfied such that

$$z \gg \frac{k(\xi^2+\eta^2)_{\max}}{2}. \tag{2.8}$$

Under this condition, the quadratic phase factor $e^{j\frac{k}{2z}(\xi^2+\eta^2)}$ in the integral of (2.7) is further canceled out. This yields the Fraunhofer approximation as

$$U(x,y) = \frac{e^{jkz}}{j\lambda z} e^{j\frac{k}{2z}(z^2+y^2)} \iint_{-\infty}^{\infty} U(\xi,\eta) e^{-j\frac{2\pi}{\lambda z}(x\xi+y\eta)} d\xi d\eta, \tag{2.9}$$

which is directly proportional to the Fourier transform of $U(\xi,\eta)$ by introducing the variables f_x and f_y by

$$f_x = \frac{x}{\lambda z} \text{ and} \tag{2.10}$$

$$f_y = \frac{y}{\lambda z}. \tag{2.11}$$

Then (2.9) can be written to

$$U(x,y) = \frac{e^{jkz}}{j\lambda z} e^{j\frac{k}{2z}(z^2+y^2)} \mathcal{F}\{U(\xi,\eta)\}\Big|_{f_x=\frac{x}{\lambda z}, f_y=\frac{y}{\lambda z}}, \tag{2.12}$$

where $\mathcal{F}\{U(\xi,\eta)\}$ denotes the two-dimensional Fourier transform of $U(\xi,\eta)$.

In this thesis, the diffractive field generated by the DOE is assumed to be in the far field, which can be easily realized by inserting a lens between the DOE and the observation plane such that the DOE and the observation plan are located at the focal plane in front and behind the lens, respectively [15, 16]. Therefore, in the algorithms of the following chapters, the diffractive field will be calculated using formulae based (2.12).

2.2 Modeling of DOEs in Numerical Methods

The design of DOEs needs the use of a powerful computer and a proper numerical method in the optimization of the DOE performance. Thus, the modeling of DOEs in the numerical methods and the corresponding description of the diffraction formula as introduced in the previous sections are essential. These modeling methods are discussed in this section.

2.2.1 Multi-level Phase Modulation DOEs

The phase or the amplitude modulation of the incident wave provided by a planar DOE can be represented by a transmittance function

$$t(\xi,\eta) = A(\xi,\eta)e^{j\phi(\xi,\eta)}, \tag{2.13}$$

where $A(\xi,\eta)$ and $\phi(\xi,\eta)$ denote the amplitude and the phase modulation, respectively. Since the amplitude modulation DOEs will absorb or reflect some of the incident wave, insofar as the efficiency is concerned, it is usually desirable to design and use the phase modulation DOEs, which provide a theoritically lossless transformation. That is, the transmittance function has an amplitude of unity and can be written as

$$t(\xi,\eta) = e^{j\phi(\xi,\eta)}. \tag{2.14}$$

Thus, the field $U_i'(\xi,\eta)$ immediately behind the DOE is the product of the transmittance function and the incident field $U_i(\xi,\eta)$ such as

$$U_i'(\xi,\eta) = U_i(\xi,\eta)t(\xi,\eta), \tag{2.15}$$

$$= U_i(\xi,\eta)e^{j\phi(\xi,\eta)}, \tag{2.16}$$

$$= |U_i(\xi,\eta)|\, e^{j(\phi_i(\xi,\eta)+\phi(\xi,\eta))}. \tag{2.17}$$

It should be noted that for the above (2.15) to (2.17) to be applicable, the DOE is assumed to be thin enough so that it satisfies the "thin-lens" approximation. The "thin-lens" approximation requires that, for a ray passing through the DOE, the entry point and exit point of the ray have almost the same trasverse coordinates (ξ,η), namely the translation of the ray inside the DOE can be neglected [15]. This assumption is easy to understand if one reflects on the boundary conditions imposed by the Maxwell's equations. For DOEs made of dielectric materials, which is usually the case, the Maxwell's equations require the tangential components of the field to be continuous across the boundary. This will in turn cause the distortion of the field inside the DOEs. However, when the DOEs are not too thick for the distortion to spread out from the boundary to the whole field, the field immediately behind the DOEs will has similar distribution with the incident field [16], except for the phase retardation provided by the DOEs.

The DOEs need to be thin enough, usually no thicker than one wavelength, to satisfy the "thin-lens" appoximation. To do so, it is essential to realize the fact that for a certain phase difference, it makes no difference in effect if a multiple of 2π is added to or subtracted from it. Hence, for any desired profile of phase difference, one can always make the thickness of the DOEs less than one wavelength by subtracting a multiple of wavelength from the original profile and still keep the resulted phase difference the same.

On the other hand, it is usually difficult to manufacture the continuous profiles under

the scale of the wavelength of visible light. However, modern semiconductor process technology enables us to approach the continuous profiles by discontinuous multiple level profiles. The discontinuous multiple level profiles are made by quantizing the original continuous profiles into levels of order of power of 2, each of which provides a phase difference of $\phi_{\max}/M$, where M is the number of levels and $\phi_{\max}$ is the max phase differece provided by the original profiles. Figure 2.2 shows the DOE equivalence, the binary and the four-level quantization processes of a continuous thick device.

Considering the efficiency issue and the real manufacturing process of DOEs, the following parts of this thesis discusses the design of the DOEs with quantized multi-level phase modulation profiles.

2.2.2 Encoding of the Multi-level Phase Modulation DOEs

To model a multi-level phase modulation DOE in numerical methods, it is convenient to consider it as a phase transmittance function represented by a matrix. By the matrix, the planar DOE is discretized into $m \times n$ zones or pixels, where m and n denote the size of matrix. In each element of the matrix, the phase modulation information of the corresponding pixel of the DOE is recorded, for example, 1, $e^{j\pi/4}$, $e^{j\pi/2}$ and $e^{j3\pi/4}$ for a four-level DOE. Similarly, the incident wave can also be discretized and represented by a matrix of the same size, each element of which records the amplitude and phase of the field at the corresponding coordinate.

Provided with the matrices of the incident wave and the DOE, the field immediately behind the DOE is obtained by element-to-element multiplication of the incident wave matrix and the DOE matrix. Figure 2.3 shows the encoding matrices of the incident field, the DOE, and the output diffractive field.

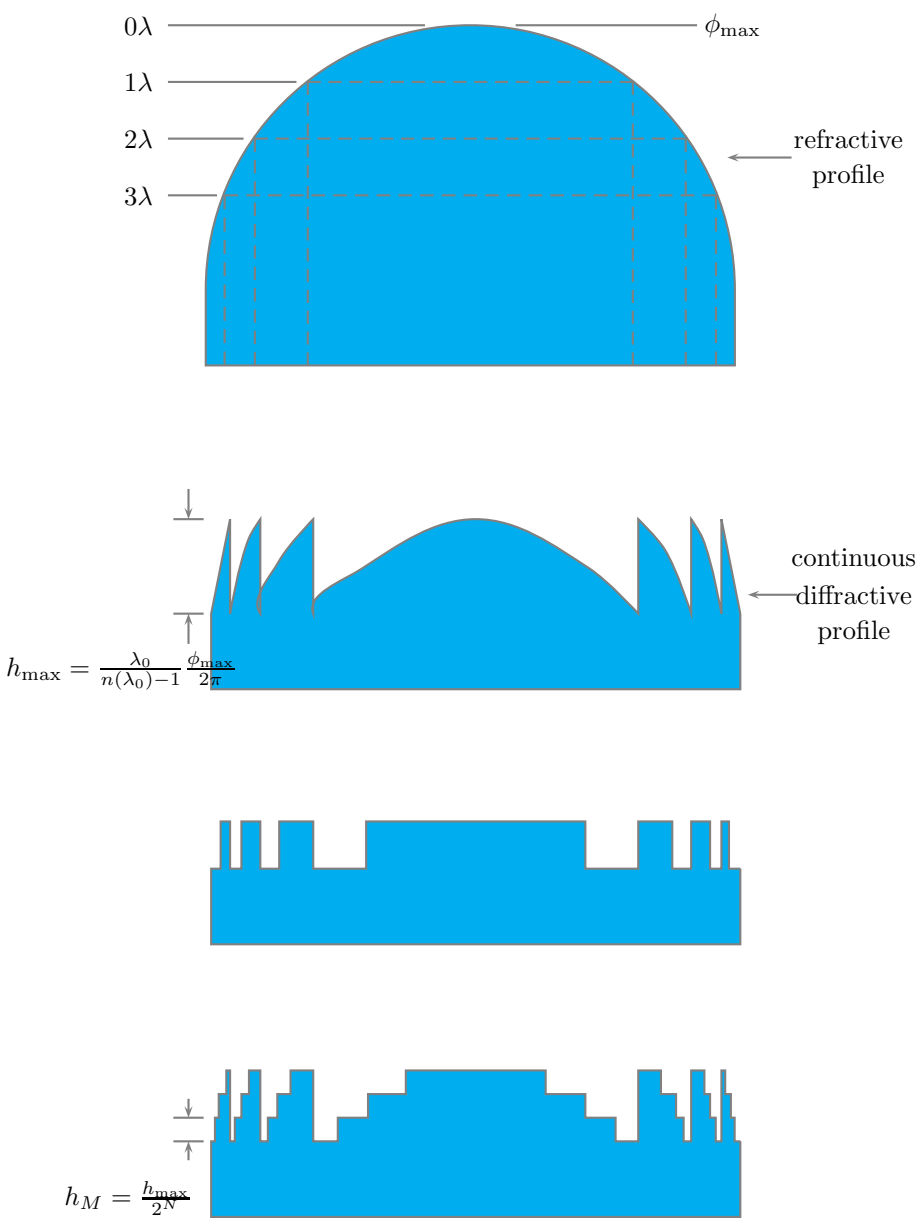

Figure 2.2 The DOE equivalence and quantization processes of a continuous thick profile

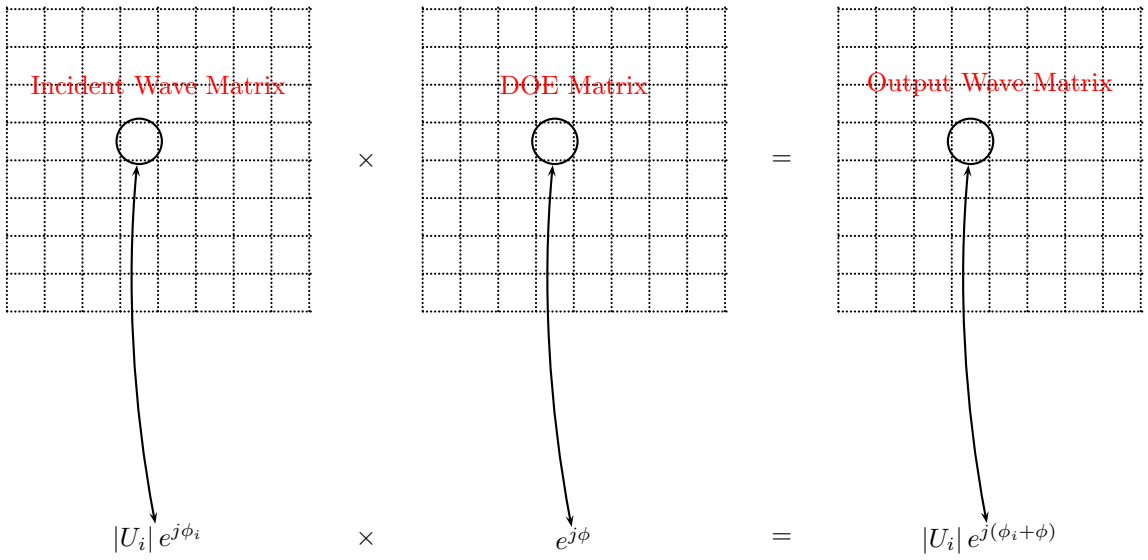

Figure 2.3 Encoding matrices of the incident field, the dOE and the output diffractive field

After the field immediately behind the DOE is calculated and encoded, the far field is calculated by the Fraunhofer approximation. The expression of (2.12) , however, necessitates a calculation of Fourier transform of the field immediately behind the DOE. Since the field has been discretized and encoded by a matrix, the discrete Fourier transform (DFT) is used [17] to calculate the matrix of the far field.

Also, because it is the field intensity that is detected by the detectors or human eyes, the intensity $|U(x,y)|^2$ is of concern, rather than $U(x,y)$. Therefore, the factor $\frac{e^{jkz}}{j}e^{j\frac{k}{2z}(x^2+y^2)}$ in (2.12) has no contribution to the intensity. In addition, the term $\frac{1}{\lambda z}$ is just a real constant since the incident wave is monochromatice coherent light and only the far field of the DOE is to be observed. Thus, insofar as the far field intensity is concerned, it is reasonalbe to use DFT to transform the field behind the DOE to the far field and neglect the factors in front of the integral in (2.12). Figure 2.4 shows the relation between the field after the DOE and the far field.

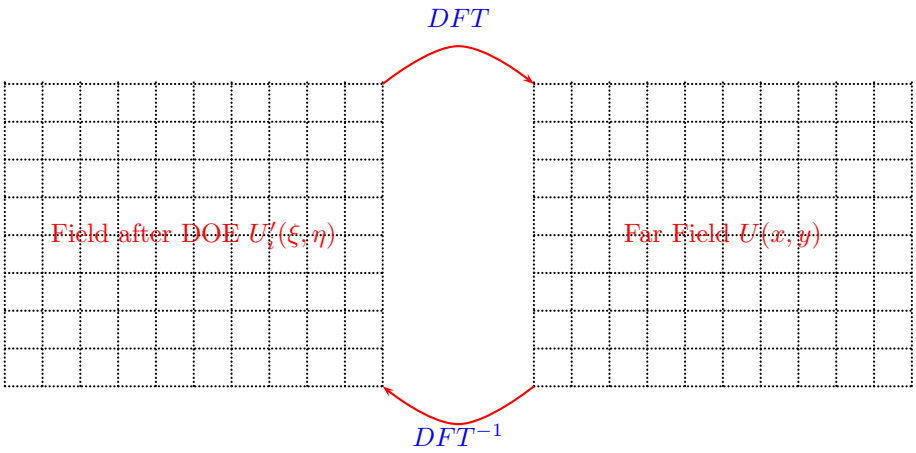

Figure 2.4 Relation between the field behind the DOE and the far field

2.2.3 Evaluation of Performance of DOEs

There are many indices or parameters to evaluate the quality of the field or image reconstructed by the DOEs. Here, the diffraction efficiency, the root-mean-square error (RMSERR), and the signal-to-noise ratio (SNR) are introduced and used to evalutate the performance of the DOEs designed in this thesis.

In the following definitions, the integral range S and N denote the signal region and the noise region of the field reconstructed by the DOEs, respectively. By signal region, the area defined by the target field is meant. The area outside this region is referred to as the noise region. $U(x,y)$ and $U_t(x,y)$ represent the complex amplitude of the reconsructed field and the target field, respectively. The definitions are as the following.

- Efficiency

$$\text{Efficiency} \equiv \frac{\iint_S |U(x,y)|^2 dxdy}{\iint_{S+N} |U(x,y)|^2 dxdy}. \tag{2.18}$$

The efficiency depicts how much energy is confined inside the signal region. The higher the efficiency, the more energy diffracted by the DOEs contributes to the

desired image.

- Root-mean-square Error

$$\text{RMSERR} \equiv \left[\frac{\iint_{S+N} \{|U_t(x,y)|^2 - |U(x,y)|^2\}^2 dxdy}{\iint_{S+N} |U_t(x,y)|^4 dxdy} \right]^{1/2}. \tag{2.19}$$

  The RMSERR evaluates the similarity between the reconstructed field and the desired image. The lower the RMSERR, the more the reconstructed field is close to the desired image.

- Signal-to-Noise Ratio

$$\text{SNR} \equiv \frac{\min_S |U(x,y)|^2}{\max_N |U(x,y)|^2}. \tag{2.20}$$

  The SNR shows the ratio of the smallest intensity in the signal region to the largest intensity in the noise region. Therefore, an SNR of larger than unity indicates that the intensity in each pixel forming the output diffractive field is larger than any of the noise.

In the design of DOEs, the resulted DOEs seldom possess good efficiency, RMSERR and SNR simultaneously. In most of the cases, chances are the optimization of one parameter causes the degradation of others. For example, if only the efficiency is optimized, the RMSERR or the SNR may degrade to an unacceptable level [18]. Therefore, the merit of optimization has to be selected carefully. Alternatively, instead of a single parameter, a combination of several parameters with different weightings can be used to be the merit of optimization. The resulted combination is usually called the fitness or merit function.

There are other parameters, such as the Strehl ratio and the space-bandwidth product (SBWP). For more details, readers are referred to [2] and [18].

Chapter 3

Optimization Methods for DOE Design—Genetic Algorithms (GAs)

Since the debut in the 1960, GAs have been evolved dramatically, theoretically or practically, from the original genetic plan by Holland in 1959 [19] to the first application by Goldberg in 1983 [20]. In this chapter, the fundamental operators of the GAs and their implementations are described first. Then the application of GA to the DOE design is introduced.

3.1 Genetic Algorithms

The basic concept of GAs lies in the natural phenomenon called "the survival of the fittest." Hence, the whole algorithm can be viewed as a miniature version of the evolution process of natural species.

The algorithm starts with an initial population. Each individual in the population shows its adaptation or fitness to the environment. The more adaptable to the environment an individual is, the more likely the individual survives. Reproduction is the natural

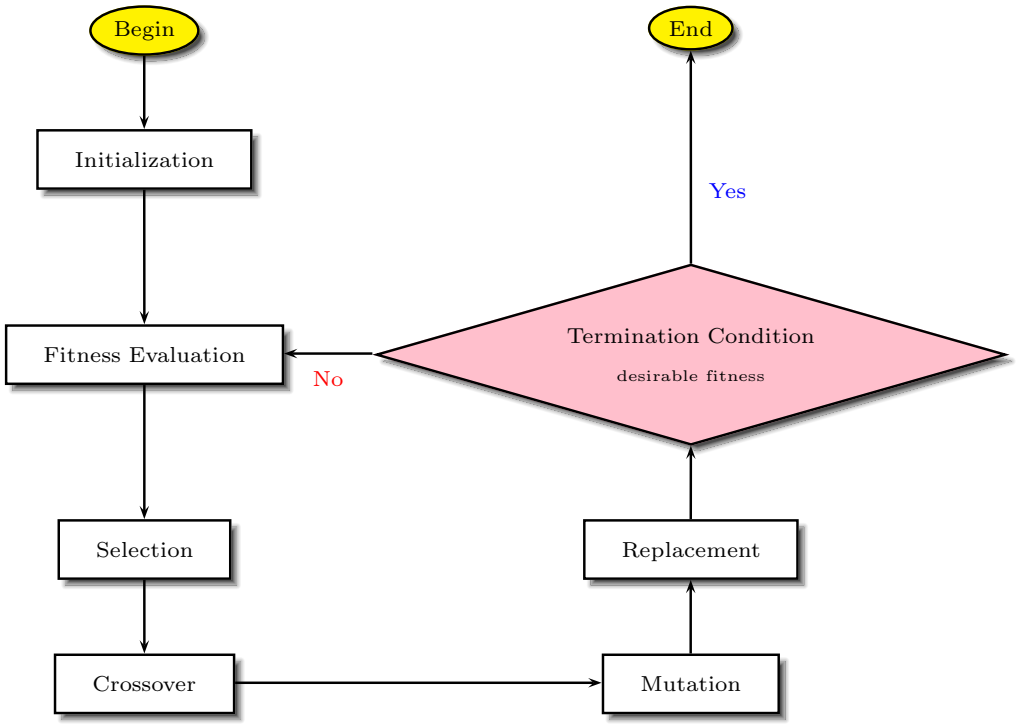

Figure 3.1 A simple genetic algorithm

instinct of species for leaving their genes. The reproduction includes the selection, the crossover and the mutation mechanisms. The offspring cycle these processes and continue the evolution of species.

GAs emulate this evolution and consist of steps or operators corresponding to the above processes. Figure 3.1 shows a simple genetic algorithm (SGA) [21]. The meaning and the implementation of each step of the algorithm shown in Fig. 3.1 are described in the following.

3.1.1 Gene Encoding and Population Initialization

At the beginning of the algorithm, an encoding mechanism is adopted to encode a solution into what is usually called a “chromosome” or an “individual”. Here, the “solution” means a random guess of the possible solutions since GAs will evolve it into the

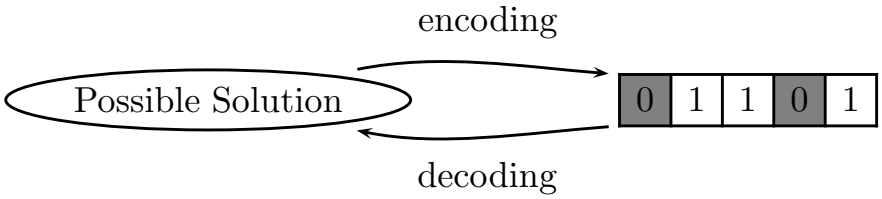

Figure 3.2 Binary encoding

final optimal one. There are many encoding methods to represent a chromosome. The most straightforward way is to encode the solution into a binary string consisting of 0's and 1's, as proposed first by Holland [21]. Figure 3.2 shows an example of the binary encoding. Other encoding schemes, such as real-valued encoding, integer-valued encoding etc. [21, 22], may be used depending on the applications or the problems to be solved.

Since the evolution begins from an initial population, as stated above, a group of chromosomes representing different possible solutions are randomly generated. The number of chromosomes in the initial population is usually decided according to the complexity of the problem and the availability of the computation resource.

3.1.2 Fitness Function

After the initialization of the population, the performance or the fitness of each chromosome is evaluated to determine its competence. The fitness is the main objective to be optimized since the it is a combination of the performance parameters. The configuration of the fitness varies from problems to problems. For instance, in the DOE design, the fitness can be one of the performance parameters such as the efficiency, the inverse of the RMSERR and the SNR. As an alternative, these parameters can be linearly combined to form a complex fitness function.

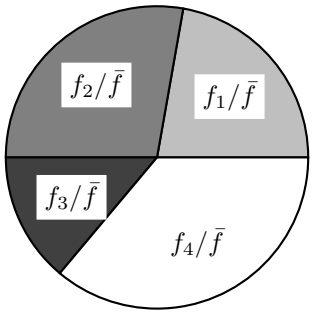

Figure 3.3 The Roulette wheel used in the selection process

Upon configuring a proper fitness function, it is essential that the fitness represents the competence of the chromosome. Otherwise, the algorithm may be trapped into a local maximum or simply yields a meaningless solution.

3.1.3 Selection

Following the rule of natural selection, survival of the fittest, the selection process determines which chromosomes have a better chance to proceed with the further processes, based on the results of fitness evaluation. Here, a "gene pool" is established, which contains the chromosomes proceeding with the crossover process.

A basic selection scheme is the "proportionate selection scheme", which states that the probability of being selected into the gene pool is proportional to the fitness f_i of each chromosome [19,21]. The simplest implementation of the proportionate selection scheme is the "Roulette wheel selection", as shown in 3.3. The probability that the chromosome with a fitness of f_i is chosen into the gene pool is $f_i/\bar{f}$, where $\bar{f}$ denotes the summation of the fitnesses of the whole population. In each selection, a random number is generated and the chromosome in the corresponding area of the wheel is selected.

Although the Roulette wheel selection is straightforward, it could be noisy. There is chance that the gene pool is occupied with chromosomes with lower fitness values and those with higher fitness values are not selected.

Besides the proportionate selection scheme, another common selection scheme is the "Rand-based selection scheme", such as the "Tournament selection", the "Truncation selection" and the "$\mu - \lambda$ selection" [23]. Unlike the Roulette wheel selection, these methods guarantee that the chromosomes with the highest fitness value are selected into the gene pool and are less noisy compared with the Roulette wheel selection.

3.1.4 Crossover

In GAs, the operators which take the responsibility to "search" solutions are the "crossover" and the "mutation" operators. The crossover operator provides a mechanism for chromosomes to exchange genetic information to the others. Together with the above selection process, the "good genes" possessed by the "elite" chromosomes can spread out to the whole population.

There are many crossover operators being studied, such as the one-point crossover, the multi-point crossover [24], the mask crossover and the uniform crossover [25]. Figure 3.4 shows an example of the simplest "one-point crossover" mechanism. First, a "crossover point" is randomly determined. When the crossover point is set between the second and the third genes, for example, the parental chromosomes are divided into two parts at the crossover point. Then, the child chromosomes are formed by exchanging the corresponding parts of the parents. The exchange of the genetic information in certain areas specified by crossover points is the common implementation for all the crossover mechanisms.

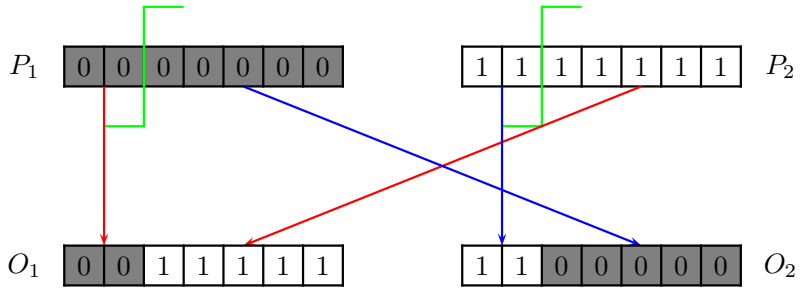

Figure 3.4 Single point crossover. P_1 and P_2 denote the parents; O_1 and O_2 denote the offspring.

3.1.5 Mutation

Since the crossover operator provides a mechanism for chromosomes to exchange genetic information which in turn starts a search in the solution space, no new information is generated in the population. Therefore, the population may finally be saturated with some certain gene strings and the evolved solution will be trapped into some local maxima. Similar to SAs, the mutation operator provides a mechanism to prevent from being trapped into the local maxima. The mutation operator introduces new information into the population and ignite a search toward the global maximum in the solution space. Figure 3.5 shows the relation between the local maximum and the global maximum in the solution space.

Depending on the number of pixels to mutate, there are “one-point mutation” and “multi-point mutation”. Figure 3.6 shows an example of the “one-point mutation” mechanism. Unlike the crossover operators, In practice, the mutation pixels are determined by the so-called mutation clock which is calculated by the pre-determined mutation rate. The reason the mutation clock is used instead of testing every gene by the mutation rate is to reduce the computation time in the random number generation. In Fig. 3.6, the 5th gene

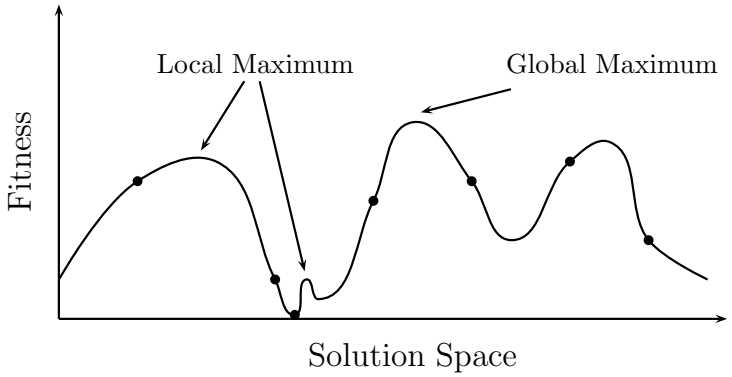

Figure 3.5 Solution space

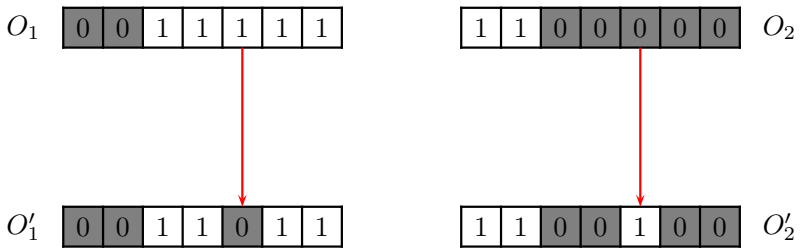

Figure 3.6 Single point mutation. O_1 and O_2 denote the offspring before mutation; O'_1 and O'_2 denote the offspring after mutation.

is chosen and replaced by its complementary digit for binary coding.

3.1.6 Replacement and Termination

The crossover and mutation operators produce the child chromosomes from the parental population. The offspring population is generated by replacing the parent chromosomes with the new ones. In general, there are three kinds of replacement methods.

- All replacement

The whole offspring population consists of only the child chromosomes.

- Tournament replacement

  As the selection process, the fitness values of the child chromosomes are evaluated. The parent and the child chromosomes compete with each other in the tournament and the ones with higher fitness survive in the offspring population.

- Elitism replacement

  A pre-determined number of chromosomes in the parent population which have the highest fitness survive in the population. Others are replaced by the child chromosomes.

After the replacement, the termination condition will determine whether the algorithm should finish or not. There are usually three kinds of termination condition:

- The solution with the desired fitness is found.
- A total generation number is reached.
- The fitness values of the population converged.

3.2 The Design of DOEs using GAs

Many studies have been conducted on the basic theory [26, 27] of GAs, as well as the applications of GAs to a wide variety of different fields [28]. Recently, the new implementation of GAs for better performance is also emerging as a study topic. In the design of DOEs, the applications of GAs have been reported [11, 12]. Nevertheless, studies on the mechanisms of the basic operators and new implementations with improved performance in DOE design are still needed.

The study on the mutation rate [14] shows that, albeit a satisfactory performance of DOE, the computation expense is an issue. However, in the study on the crossover mechanism [13], the different regular exchange crossover patterns only yield DOEs with unacceptable performance when the mutation mechanism is excluded from the algorithm.

To improve the performance of resulted DOEs while boosting the computing speed of the algorithm, an algorithm similar to that used in [13] is adopted in this thesis. That is, the mutation mechanism is not used and a new crossover operator is proposed to improve the performance of the algorithm. The novel crossover operator is introduced in the next chapter. Prior to that, the algorithm used in [13] is described and its results are summarized in this section.

3.2.1 The Algorithm

In [13], the GA was used in the design of two-dimensional binary phase DOEs. The input was assumed to be monochromatic coherent light with uniform amplitude and the output diffractive field was assumed to be in the far field of the DOEs. The DOE was discretized into 64×64 pixels and encoded into a matrix, as illustrated in Chapter 2. The target pattern was a cross sign which was also encoded into a 64×64-pixel matrix. In summary, the simulation parameters in [13] are listed below.

- Phase-only DOEs (64×64 pixels) for monochromatic coherent light with uniform

  transmittance $t(\xi, \eta) \equiv e^{j\phi(\xi,\eta)}$,

- Binary phase DOEs

$$\phi(x, y) \equiv \begin{cases} \pi \\ \\ 0 \end{cases},$$

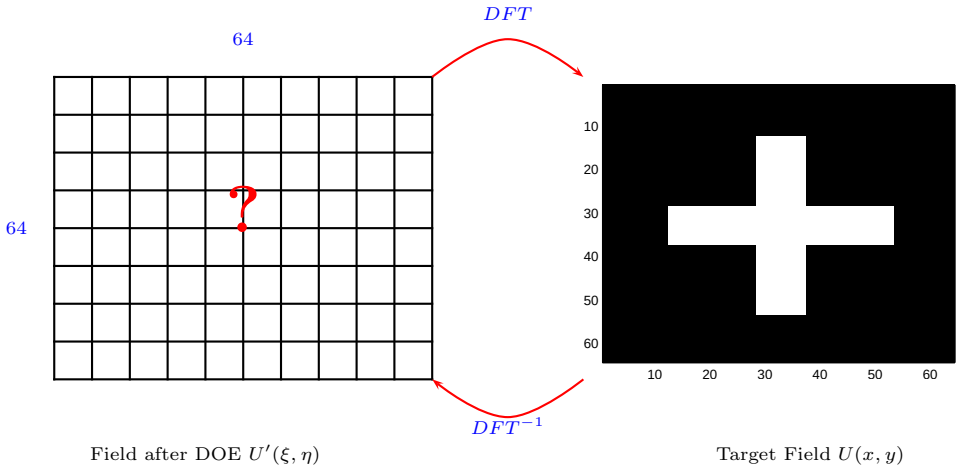

Figure 3.7 Simulation setup

- Cross target pattern.

These parameters are related as in an operation loop as shown in Fig. 3.7. Note that since the incident wave was a coherent uniform one, the amplitude $|U_i(\xi, \eta)|$ was unity and the phase $\phi_i(\xi, \eta)$ was a constant, which was assumed to be zero without losing the generality. Therefore, the field immediately after the DOE was exactly the same as the transmittance function $e^{j\phi(\xi,\eta)}$. The cases of the multi-level and the continuous phase DOE design are very similar to the binary one, except the setting of the phase ϕ. For example, in the four-level case, the phase ϕ can be 0, $\pi/2$, π and $3\pi/2$; and in the continuous case, ϕ can be any real number between 0 and 2π.

The algorithm used in [13] is shown in Fig. 3.8. The algorithm follows a SGA except for the absence of the mutation mechanism since the effect of the crossover mechanism is the core issue of that research.

The initial population consisted of 33 chromosomes. The fitness function was defined by linearly combining the Efficiency, the RMSERR and the SNR as follows,

$$\text{Fitness} \equiv 1 \times \text{Efficiency} + 10 \times \text{SNR} + 10 \times \frac{1}{\text{RMSERR}}, \tag{3.1}$$

where the Efficiency, the RMSERR and the SNR are defined in (2.18) to (2.20). The elitism selection was used to construct a gene pool, in which the chromosomes with the highest efficiency, fitness and SNR got two copies in the pool. The other chromosomes had their own copies in the pool. This resulted in the gene pool with 36 chromosomes, in which the chromosomes with highest efficiency, fitness and SNR had chance twice as much as others to be involved in the crossover process.

In the crossover stage, 16 crossovers were performed and the crossover pattern changed every 100 generations. Each crossover involved two chromosomes, randomly selected from the gene pool and produced two child chromosomes.

After the crossover process, the replacement was conducted to form the offspring population. Unlike the initial population, the offspring population consisted of 36 chromosomes, in which 32 chromosomes were produced by the 16 crossovers, 3 chromosomes were the "elite" with the highest efficiency, fitness and SNR determined previously, and the last one was randomly selected from the gene pool.

It should be noted that there was no mutation process in the algorithm. Without mutations, after a number of generations, the population lost its diversity and the performance did not grow as the population evolves. This phenomenon was called the population purification. In such a circumstance, a "Population Restoration" mechanism was introduced to prevent the population from purification. The mechanism is, at some pre-determined generation, except the three chromosomes with the highest efficiency, fitness and SNR, the whole population will be replaced with the original population, i.e. the original 33

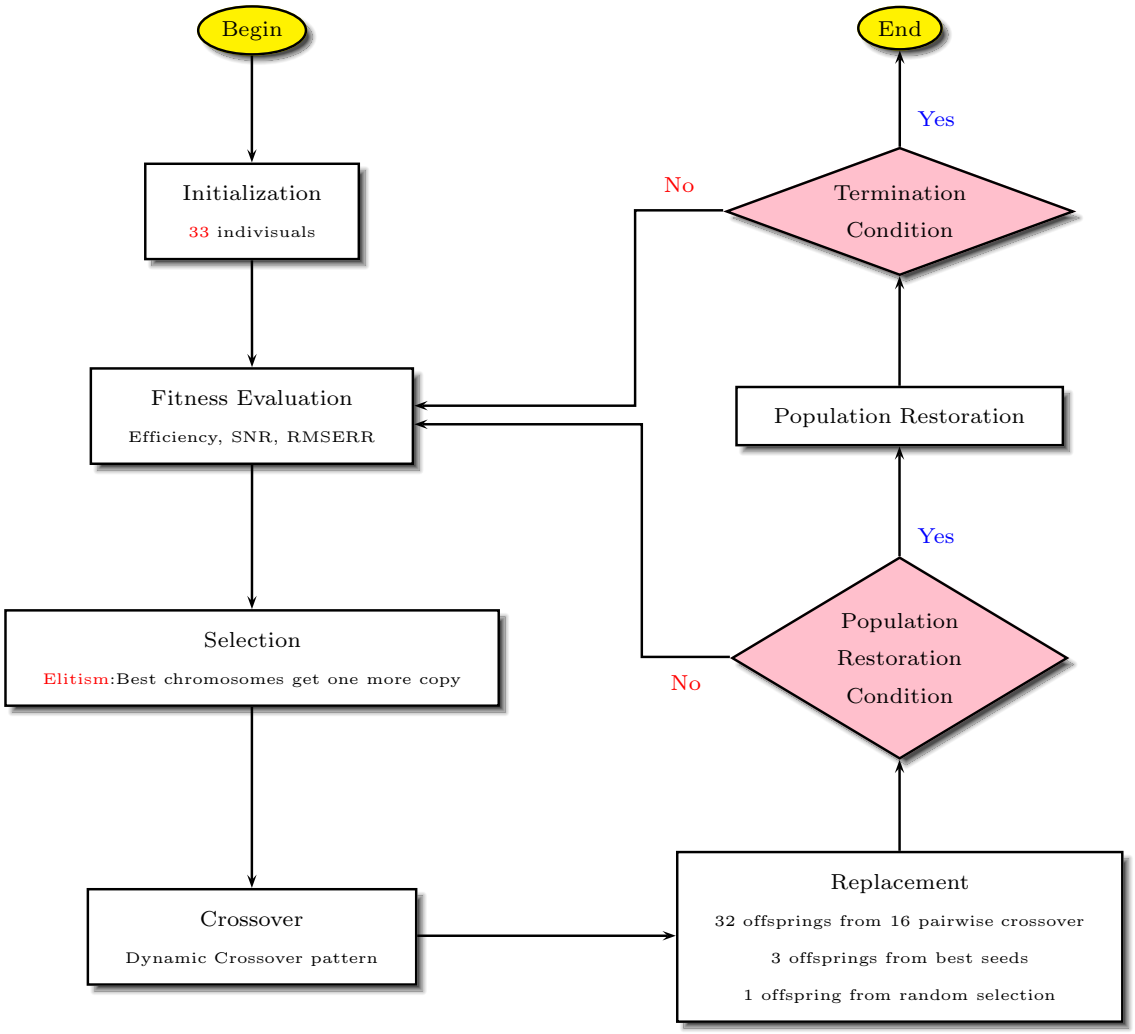

Figure 3.8 Algorithm used in DOE design

chromosomes.

The loop of the evolution proceeded until the pre-determined generation number was reached. The total generation number and the population restoration generation number were determined experimentally.

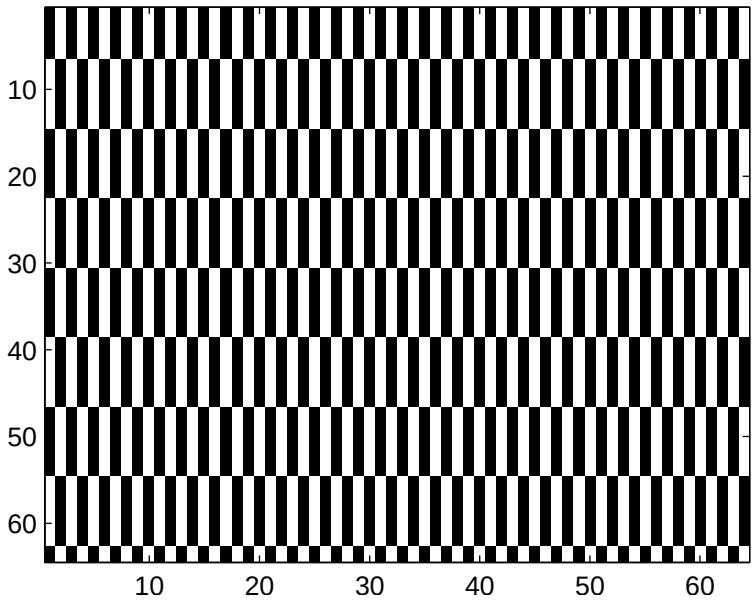

Figure 3.9 Orthogonal sinusoidal crossover pattern

3.2.2 The Results of Regular Exchange Crossover Patterns

In [13], several different crossover patterns were proposed, such as the random crossover pattern, the orthogonal sinusoidal crossover pattern etc. These patterns determined the corresponding pixels of the two parental chromosomes to be exchanged. Figure 3.9 shows an orthogonal sinusoidal crossover pattern. The bright region denotes the pixels to be exchanged and the dark region represents the pixels to be kept. Figures 3.10–3.13 show the results of the algorithm using the orthogonal sinusoidal crossover patterns as shown in Fig. 3.9. Similar results were obtained by using other patterns proposed in [13].

According to these results, it was concluded that although the crossover mechanism could result in evolution which improved the performance of the whole population, the growth speed was too slow and the performance was disappointing. After a total of 500,000 generations, the best efficiency, RMSERR, and SNR were 0.34, 0.32, and 0.035, respectively. The efficiency and RMSERR were only half as good as that obtained in [14] and the SNR even falled into a negligible range. As expected, the best far field pattern

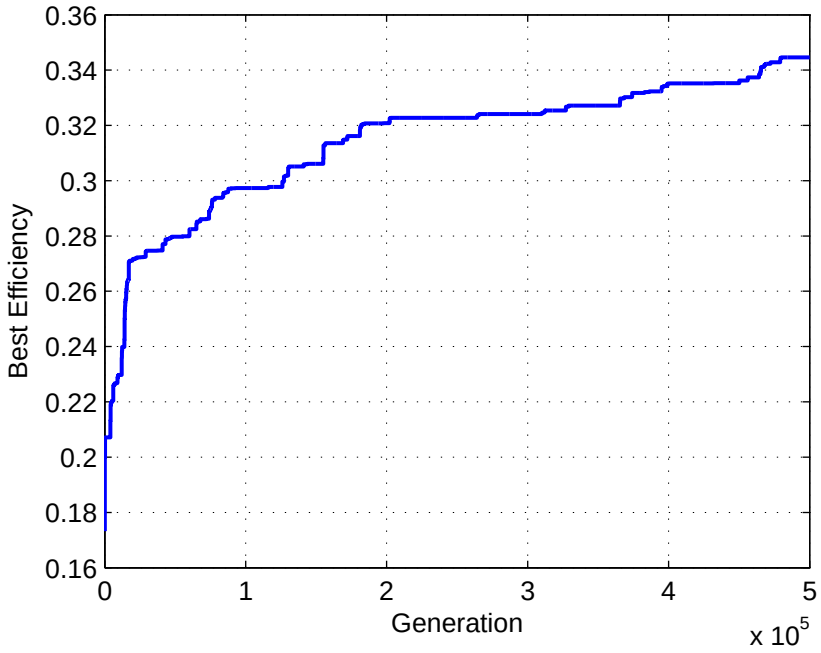

Figure 3.10 Best efficiency of the orthogonal sinusoidal crossover pattern

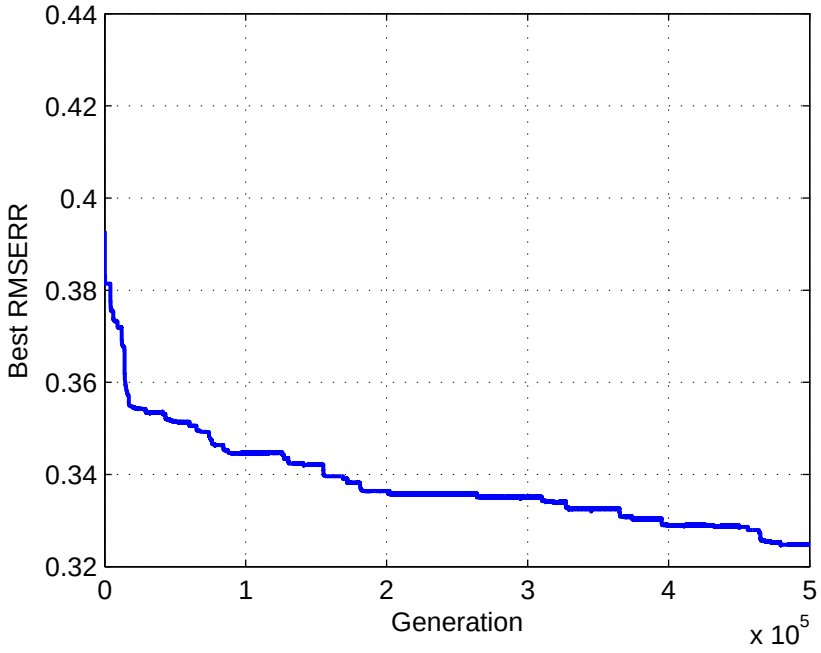

Figure 3.11 Best RMSERR of the orthogonal sinusoidal crossover pattern

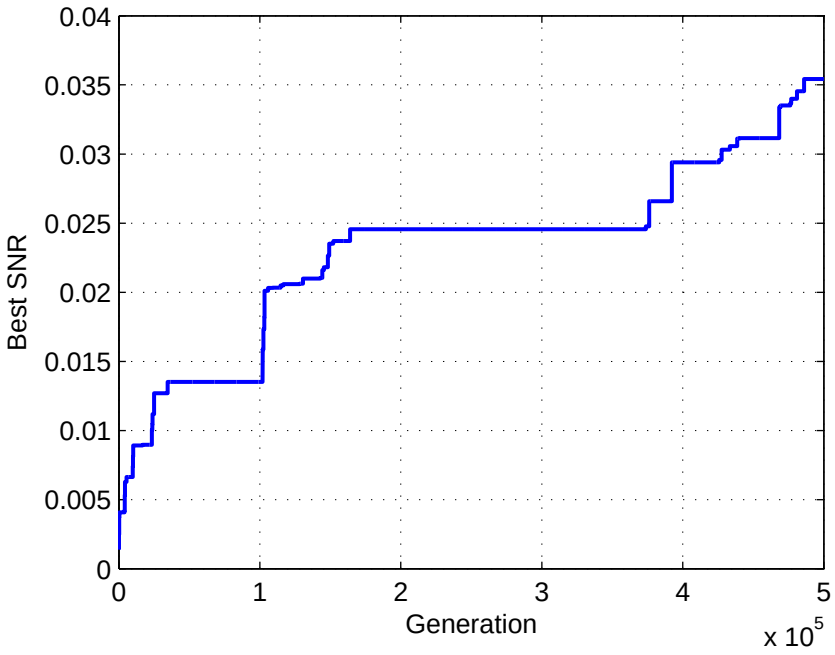

Figure 3.12 Best SNR of the orthogonal sinusoidal crossover pattern

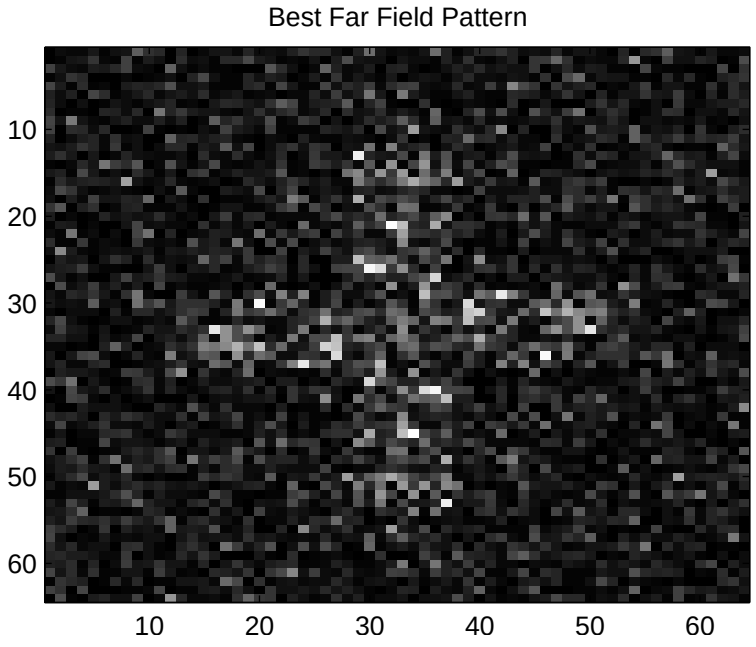

Figure 3.13 Best resulted far field of the orthogonal sinusoidal crossover pattern

shown in Fig. 3.13 possesses a faint shape and is very noisy. Therefore, some stimulation mechanisms may be needed to stimulate the evolution and boost the growth of the performance. The mutation mechanism can provide a stimulation to the evolution and lead the population toward the global maximum. However, the considerable computation expense is inevitable [14].

Motivated by these problems, a novel crossover operator is proposed in the next chapter to increase the growth rate of the performance. By using this crossover operator, the optimization which yields satisfactory performance can be achieved in relatively short time.

Chapter 4

A Novel Crossover Operator—TArget-oriented CrossOver (TACO) Operator

The results and the limitations of the regular exchange crossover patterns used in the algorithm [13] were illustrated in the previous chapter. In this chapter, a novel crossover operator is derived. Then the related parameters used when the operator is introduced into the algorithm are determined.

4.1 Derivation of the Novel Crossover Operator

As illustrated in Fig. 3.7, a phase DOE is to be found such that the output far field is a cross pattern when the DOE is illuminated by a monochromatic coherent light with uniform amplitude. All the fields and the DOE are discretized and encoded into 64-by-64-pixel matrices. The field immediately behind the DOE and the far field can be computed from each other using the discrete Fourier transform (DFT).

4.1.1 Mathematical Expression of the Crossover Operator

The crossover process can be expressed mathematically as

$$O_1 = P_1 \cdot XO_1 + P_2 \cdot XO_2, \tag{4.1}$$

$$O_2 = P_1 \cdot XO_2 + P_2 \cdot XO_1, \tag{4.2}$$

where the matrices P_1 and P_2 represent the parental chromosomes, the matrices XO_1 and XO_2 the crossover patterns, and O_1 and O_2 the child chromosomes. Note that although the chromosomes P_i and O_i represent the DOEs, they also represent the field immediately behind the DOEs. By realizing the relation that XO_1 and XO_2 are complementary to each other, i.e.

$$XO_2 = 1 - XO_1, \tag{4.3}$$

the above (4.1) and (4.2) can be written into

$$O_1 = (P_1 - P_2) \cdot XO_1 + P_2, \tag{4.4}$$

$$O_2 = (P_2 - P_1) \cdot XO_1 + P_1. \tag{4.5}$$

To derive the expression of XO_1 and XO_2, we take advantage of one of the most important characteristic of this optimization problem. This property is that the target of this problem is well-defined—a cross pattern which is encoded into a 64-by-64-pixel matrix. Denoting the target pattern matrix as T, the ultimate goal of the optimization is given by $DFT\{O_1\} = T$ or $O_1 = DFT^{-1}\{T\}$. Substituting this into (4.4) results in

$$\begin{aligned} XO_1 &= \frac{O_1 - P_2}{P_1 - P_2}, \\ &= \frac{DFT^{-1}\{T\} - P_2}{P_1 - P_2}. \end{aligned} \tag{4.6}$$

The above matrix operations are pixel-wise or element-to-element operations and (4.6)

can be written in a pixel-wise sense as

$$xo_1 = \frac{t^{-1} - p_2}{p_1 - p_2}, \tag{4.7}$$

where xo_i, p_i and t^{-1} denote the corresponding pixels in XO_i, P_i and $DFT^{-1}\{T\}$ matrices, respectively. Because of the pixel-wise nature of the above operations, it is possible that the corresponding pixels of P_1 and P_2 are of the same values and, therefore sigulariteis occur in the operations. The sigularity issue can be easily solved if one reflects on the nature of GAs that if the parental chromosomes possess the same gene strings, so do the children chromosomes. Therefore, xo_1 is set to be unity, i.e. $o_1 = p_1 = p_2$, wherever $p_1 = p_2$.

Note that in (4.6), T denotes the matrix of the target pattern and possesses both the amplitude and the phase information. Since only the intensity is detected by the detectors or human eyes, the phase ϕ of T is arbitrary and can be specified at one's convenience.

The crossover operator expressed in (4.6) and (4.7) is derived by taking advantage of the existence of the well-defined target pattern of this optimization problem. Therefore, it is entitled "Target-Oriented CrossOver" operator, and is referred to as the TACO operator hereafter.

4.1.2 Discussions on Several Cases

In this section, the situations of several kinds of the parental chromosomes P_i are discussed.

Case 1. The two parental chromosomes are identical.

$$XO_1 = 1 \quad \Rightarrow \quad O_1 = O_2 = P_1 = P_2.$$

Case 2. None of the corresponding pixels of the two parental chromosomes are the same. For binary phase modulation DOEs, this means that the two parents are complementary matrices, and the two child chromosomes are given by

$$O_1 = DFT^{-1}\{T\},$$

$$O_2 = DFT^{-1}\{T\} + P_2 + P_1.$$

It turns out that O_1 is a Kinoform of T provided the phase of T is given as the phase of $DFT\{P_1\}$.

Case 3. Some of the corresponding pixels of the two parental chromosomes are the same. The identical pixels cause sigularities to occur. In this case, (4.7) is applied for pixels that $p_1 \neq p_2$ and *xo* is set to be unity for pixels in which $p_1 = p_2$. The resulted O_1 is a combination of P_1 and $DFT^{-1}\{T\}$.

It is noted that mostly, Case 3 is the case and Case 1 and Case 2 rarely occurs.

4.2 Determination of Relevant Parameters

To apply the TACO operator to the algorithm as shown in Fig. 3.8, there are two parameters to be determined; one is the condition for population restoration and the other is the termination condition for the simulation. In this section, the generation when the population is restored is determined first. Then the termination condition is determined using this population restoration condition.

4.2.1 Condition for Population Restoration

Ten simulations were conducted to determine the condition for population restoration. To make a clear comparison, the initial population was chosen to be the same as that

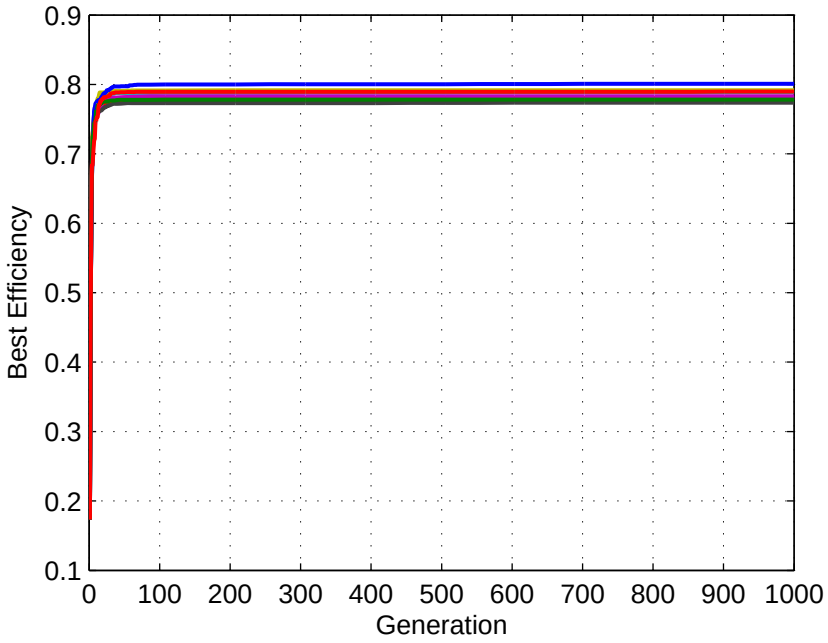

Figure 4.1 Best efficiencies of the population restoration test

used in [13]. The population restoration was not conducted here. The number of the generations of the simulations was set to be 1,000. The phase ϕ of the target pattern matrix T was selected as the phase of $DFT\{P_1\}$.

The simulation results are shown in Figs. 4.1–4.3. Without the population restoration, the highest efficiencies and the lowest RMSERRs soon converged to about 0.8 and 0.15 within 50 generations of evolution, respectively. On the other hand, the SNRs grew more slowly and converged at around generations 100–200. In most cases, the highest SNRs converged to about 0.5 to 1.0 and stagnated after 200 generations of evolution. Compared with the increase of SNRs before generation 100, the SNRs increased more slowly in generations 100–200. Therefore, a number between 100 to 200 may be a reasonable generation for population restoration. In the following simulations, every 100 generations of evolution is selected as the population restoration condition.

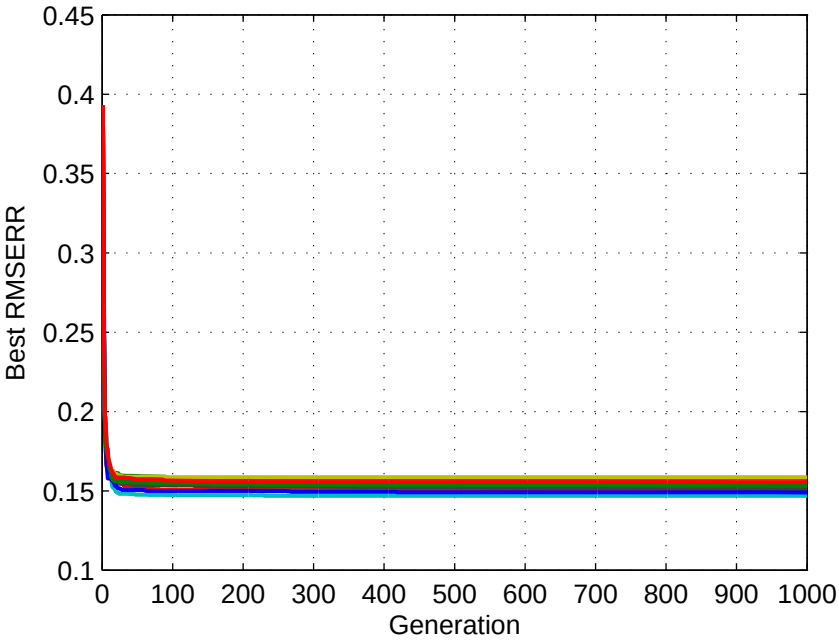

Figure 4.2 Best RMSERRs of the population restoration test

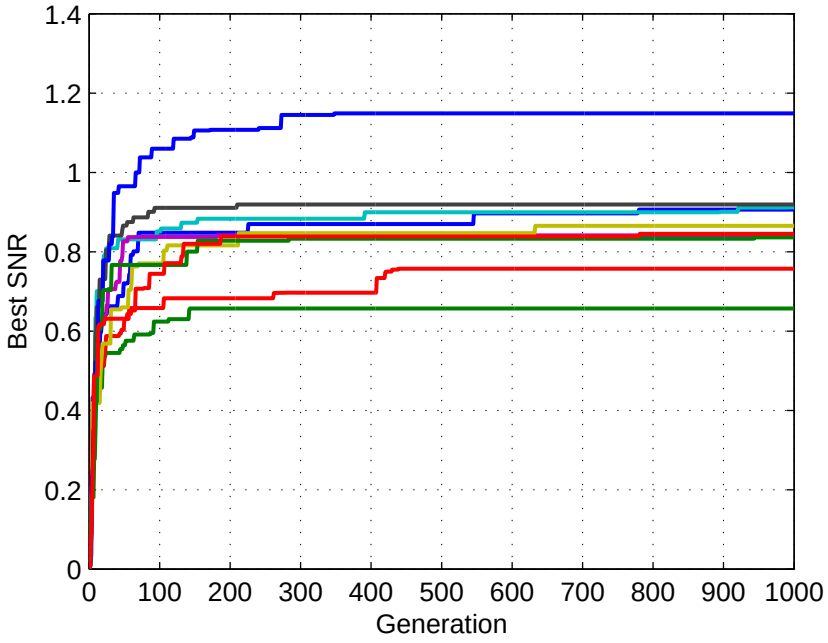

Figure 4.3 Best SNRs of the population restoration test

4.2.2 Condition for Termination of the Simulation

Although the total generations of a simulation may be chosen as large as the computation resource is available in order to yield the desired performance, it is of practice to terminate a simulation in a certain number of generations. After the certain generations of evolution, there exists chromosomes with the desired performance or the performance has converged to a certain value and may not be better anymore. Using the population restoration condition described in the previous section, another 10 simulations were conducted to determine when to finish the simulation. In all 10 simulations, the population restoration was performed every 100 generations of evolution, and a total of 10,000 generations were conducted.

The simulated performance is shown in Figs. 4.4–4.6. As shown in Fig. 4.6, three of ten highest SNRs kept increasing till the termination condition was fulfilled. In the rest of the cases, the SNRs reached a level of 1.2 and the increase was not significant after 5,000 generations of evolution. To reach a compromise between the optimal performance and the computation expense, a total of 5,000 generations will be used to terminate the simulations in the next chapter.

The results shown in Figs. 4.4–4.6 can be compared with that shown in Figs. 4.1–4.3 to illustrate the effect of the population restoration, as described in Section 3.2.1. Taking into account of the values at the 1,000th generation in Figs. 4.4–4.6 and comparing with that in Figs 4.1–4.3, most of the efficiencies and the SNRs were higher and the RMSERRs were lower. This is a prime example of the importance of the population restoration mechanism in the absence of a mutation mechanism in order to introduce new growth stimulation into the population.

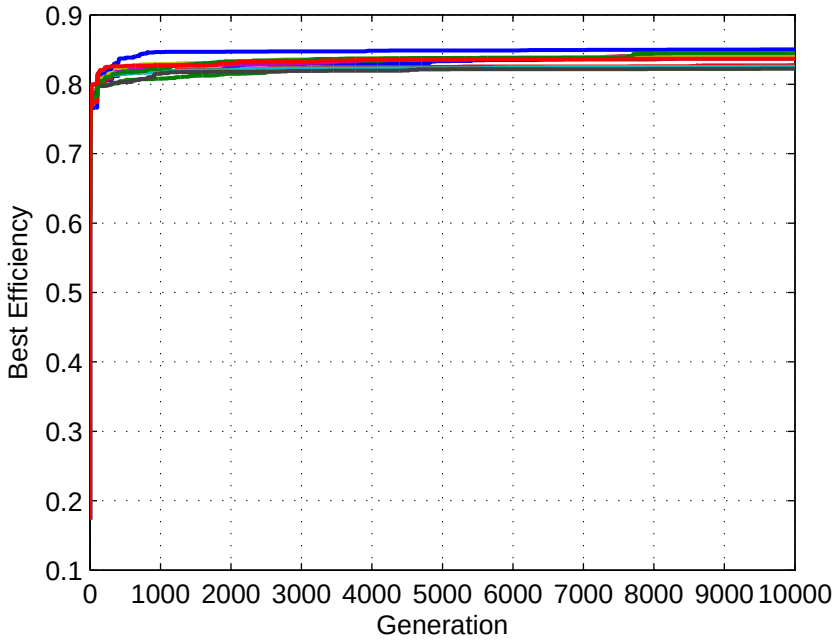

Figure 4.4 Best efficiencies of the total generation test

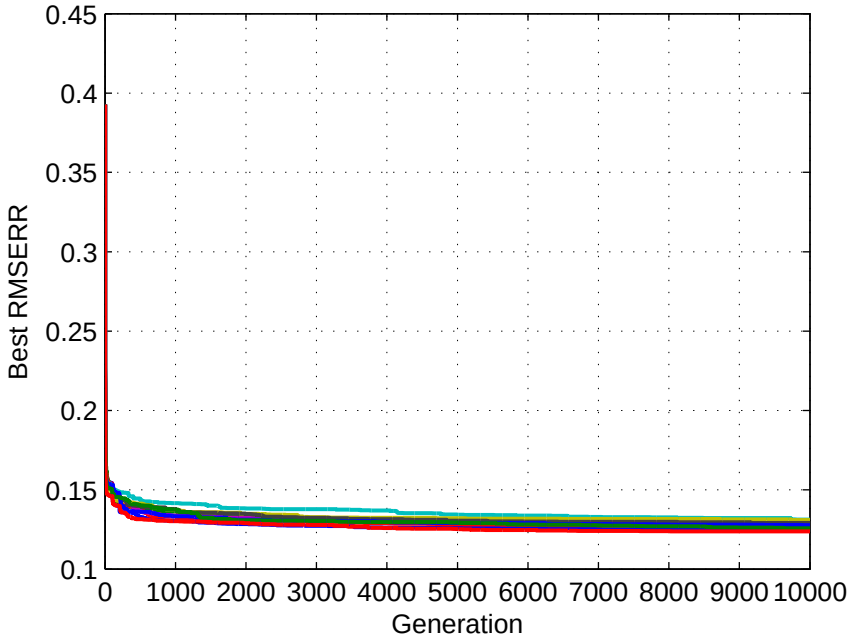

Figure 4.5 Best RMSERRs of the total generation test

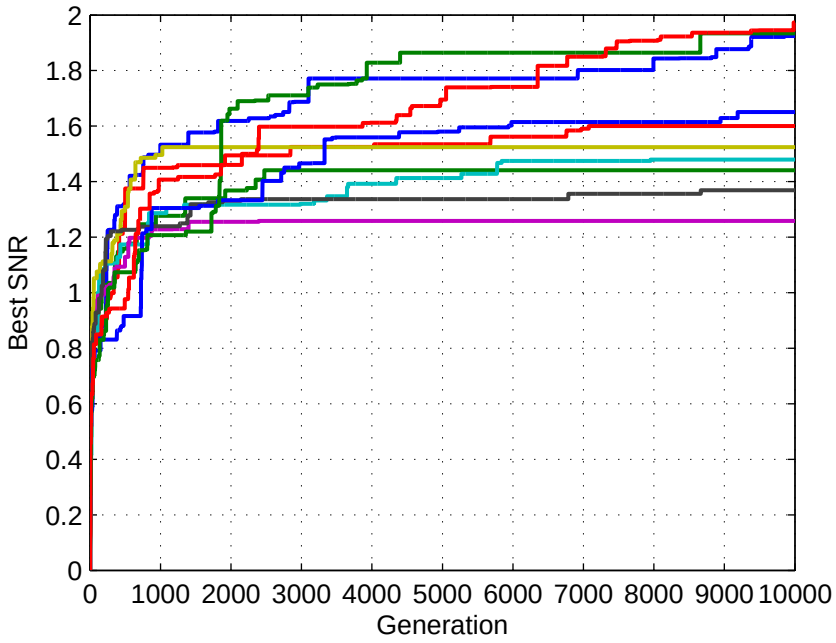

Figure 4.6 Best SNRs of the total generation test

4.2.3 Verification of the Relative Parameters

In the above sections, the population restoration condition and the termination condition for the simulation were determined—i.e. the initial population is restored every 100 generations of evolution and the simulation is terminated when a total of 5,000 generations have evolved. These conditions yield a total of 50 restorations of population.

It is worth noting that, only after 200 generations of evolution does the real convergence of the highest SNR happens, as shown in Fig. 4.3. Therefore, it is possible that the population restoration condition of every 100 generations may be a premature one and affect the performance of the results. To this end, an experiment was conducted to verify the above conditions. Similar to the previous section, 10 simulations were conducted with the same initial population. The population restoration condition was set to be every 200 generations and the simulation was terminated after a total of 10,000 generations. These conditions were selected such that they yield a total of 50 restorations of population, the

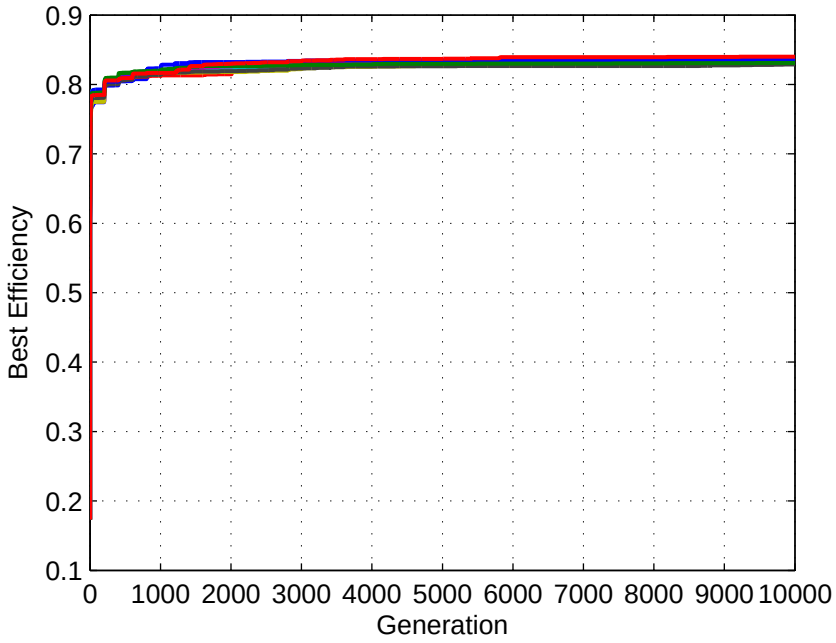

Figure 4.7 Best efficiencies of the simulations with comparable parameters

same as the case to be verified.

The results are shown in Figs. 4.7–4.9. From these figures, it can be concluded that both cases yield comparable performance after 50 population restorations—efficiencies about 0.8–0.85, RMSERRs about 0.12–0.13, and SNRs about 1.1–1.8. This is to be expected since, as indicated in previous sections, the increase in performance in generations 100–200 is not so significant as that before the generation 100. Therefore, the premature population restoration is not an issue if the population restoration condition is set to be every 100 generations of evolution. The parameters—population restoration every 100 generations of evolution; a total of 5,000 generations—will be used in the simulations of the following chapter.

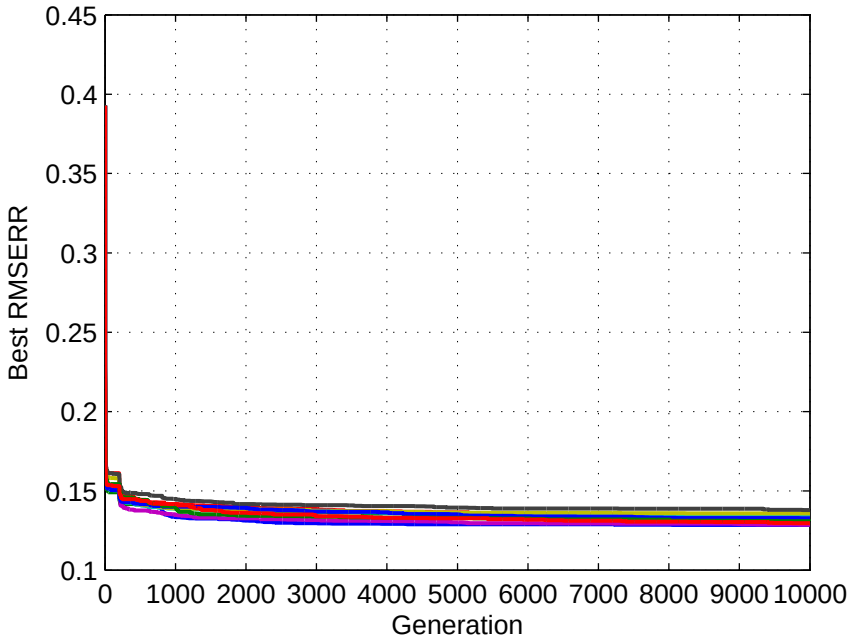

Figure 4.8 Best RMSERRs of the simulations with comparable parameters

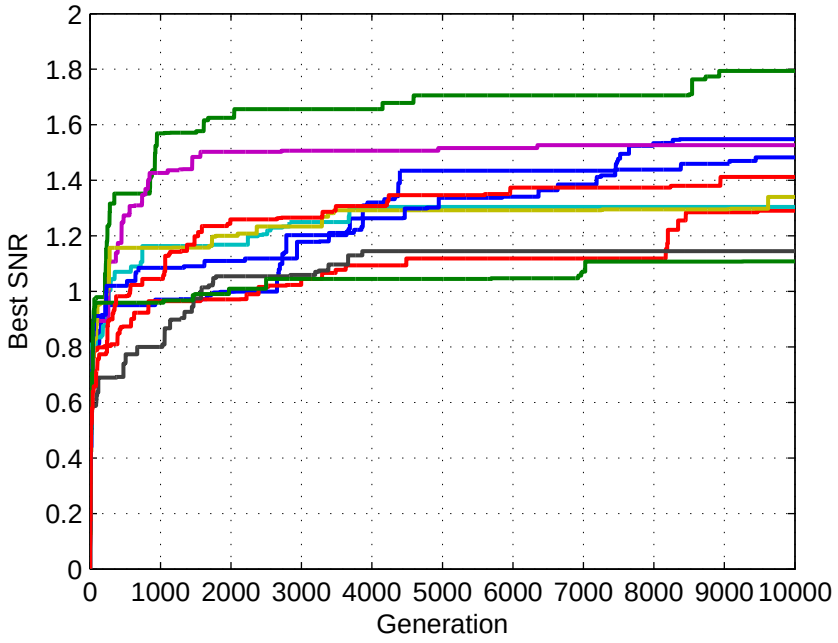

Figure 4.9 Best SNRs of the simulations with comparable parameters

Chapter 5

Simulation Results—Optimization Using GA with the TACO Operator

In this chapter, the TACO operator is applied to the optimization process of the DOE design. The simulations include the robustness test, the comprehensive information of the final population, and the results of using different phases of the target pattern. The TACO operator is also applied to the multi-level DOE design, and is compared with the IFTA.

5.1 Simulation Setup

Before the discussion of the results, the simulation setup is summarized as below.

1. DOE modeling

    - Modulation method: Phase modulation type DOEs.
    - Quantization method: binary, four-level (Section 5.5.1), and eight-level (Section 5.5.2) quantizations.
    - Incident wave: monochromatic coherent incident light with uniform ampli-

tude.

- Target pattern: a cross pattern.
- Discretization method: 64×64-pixel matrices.

2. Optimization algorithm

- Initial population: the 33 chromosomes the same as [13].
- Crossover mechanism: the TACO operator with population restoration.
- Mutation mechanism: none.
- Population restoration: every 100 generations of evolution.
- Termination of simulation: after a total of 5,000 generations.

Figure 5.1 shows the flowchart used in the following simulations.

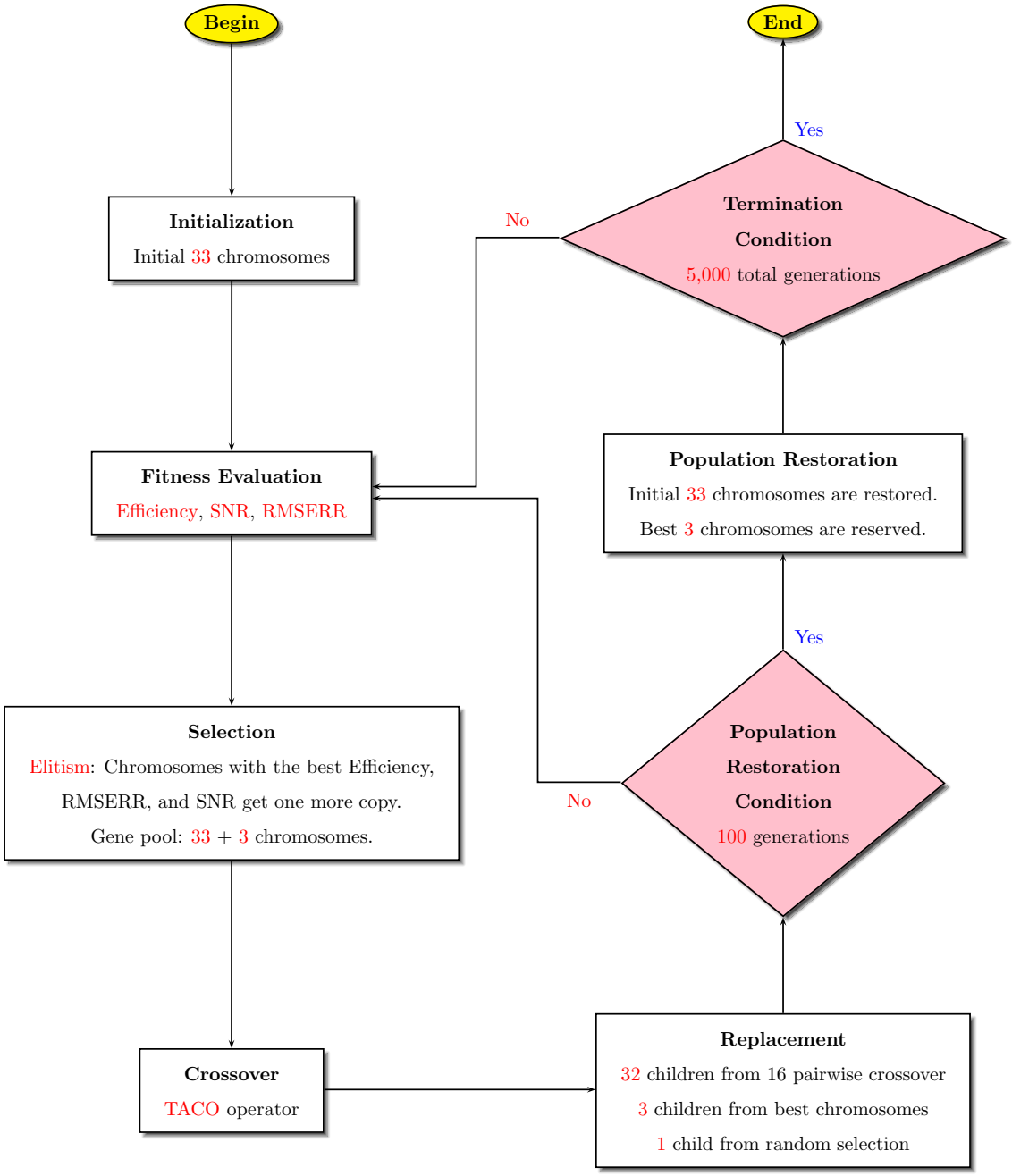

Figure 5.1 Flowchart of the algorithm with TACO used in the simulation

5.2 Robustness Test

Some algorithms, such as the IFTA, yield results with performance of a wide range, depending on the different initial conditions. On the other hand, there are algorithms which are insensitive to the initial conditions, such as the simulated annealing (SA) algorithm. A typical GA sorts mainly to the random process. Ideally, it is insensitive to the initial condition, provided the random process is "random enough". Since the proposed method is based on the GA, the insensitivity to the initial condition is to be expected.

To test the robustness of the TACO operator, another five random binary populations, each of which consisted of 33 chromosomes, were generated and used in the algorithm. The five additional populations are numbered random binary 1 to 5. Including the original population used in [13], each of the six populations was used in two simulations as discussed in the following. The results were then compared with each other to verify the robustness of the TACO operator.

The results of the first simulation are shown in Figs. 5.2–5.4. It can be observed from the figures that all the six populations had almost the same efficiencies and RMSERRs. In contrast, the SNRs vary from 1.2 to about 1.7—the random binary 4 had the highest SNR and the random binary 5 had the lowest one.

One possible reason for the difference in SNR between each population is that there are populations consisting of chromosomes superior to other populations. That is, in the present simulation result, the chromosomes of the random binary 4 might be superior to that of the random binary 5. To this end, the second simulation was conducted with exactly the same six initial populations. If a certain initial population possesses superiority over others, the results of the second simulation will resemble that of the first simulation in the way that the random binary 4 has the highest SNR and the random binary 5 has the

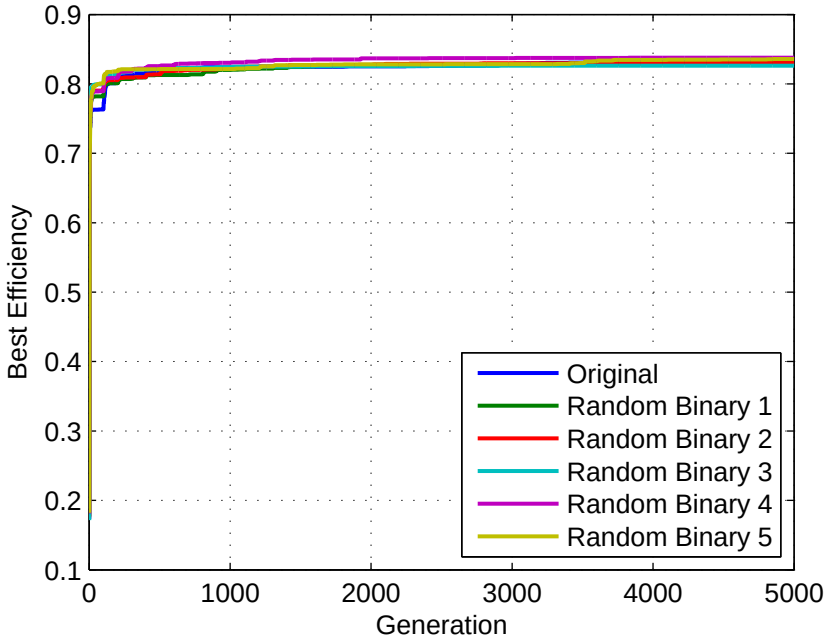

Figure 5.2 Best efficiencies of the robustness test 1

lowest one.

The results of the second simulation are shown in Figs. 5.5–5.7. Similar to that of the first simulation, the efficiency and the RMSERR of each population were comparable while the SNRs differed from population to population. As shown in Fig. 5.7, the original population had the highest SNR and the random binary 2 had the lowest one—a result different from the first simulation.

A ranking of the SNR of each population for each simulation is shown in Table 5.1. It is clear that no specific population is certainly superior to others. Each population has its chances to evolve to a solution with a higher or a lower SNR, depending on the random process. Note that, albeit a somewhat lower value, the lowest SNR was still higher than 1.2. Therefore, it is concluded that the TACO operator is robust against the initial population.

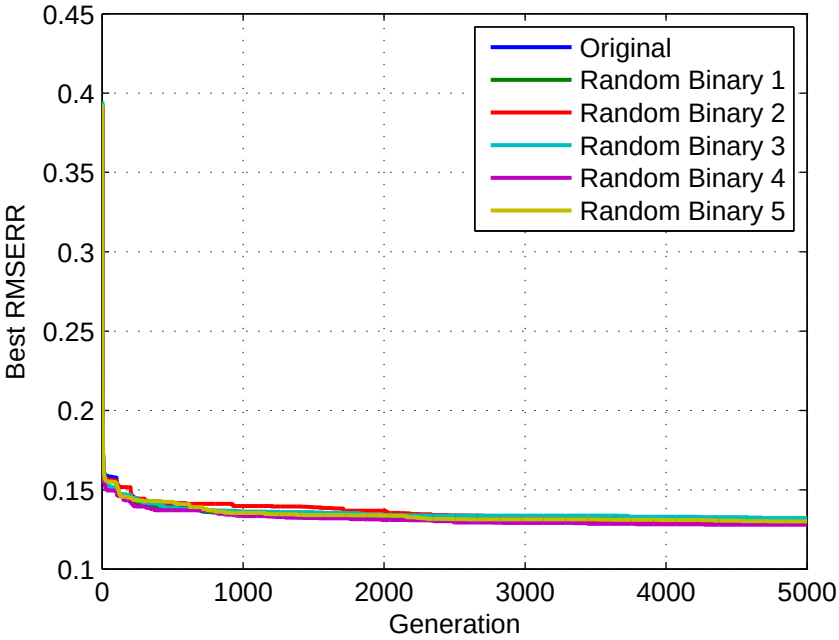

Figure 5.3 Best RMSERRs of the robustness test 1

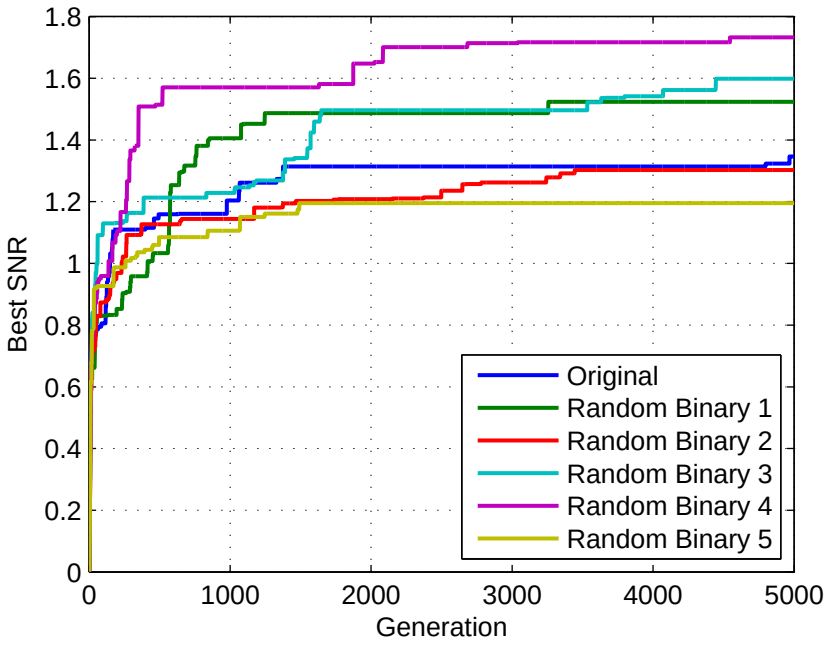

Figure 5.4 Best SNRs of the robustness test 1

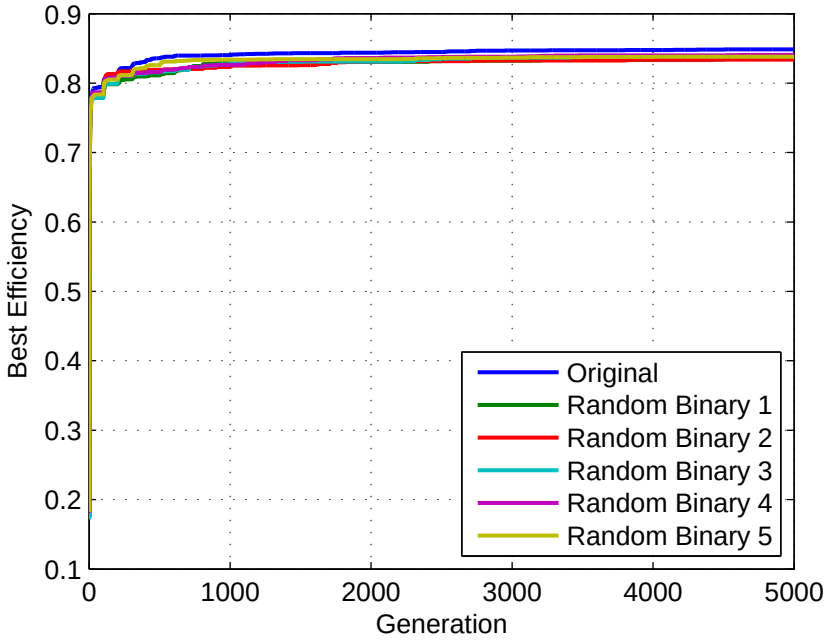

Figure 5.5 Best efficiencies of the robustness test 2

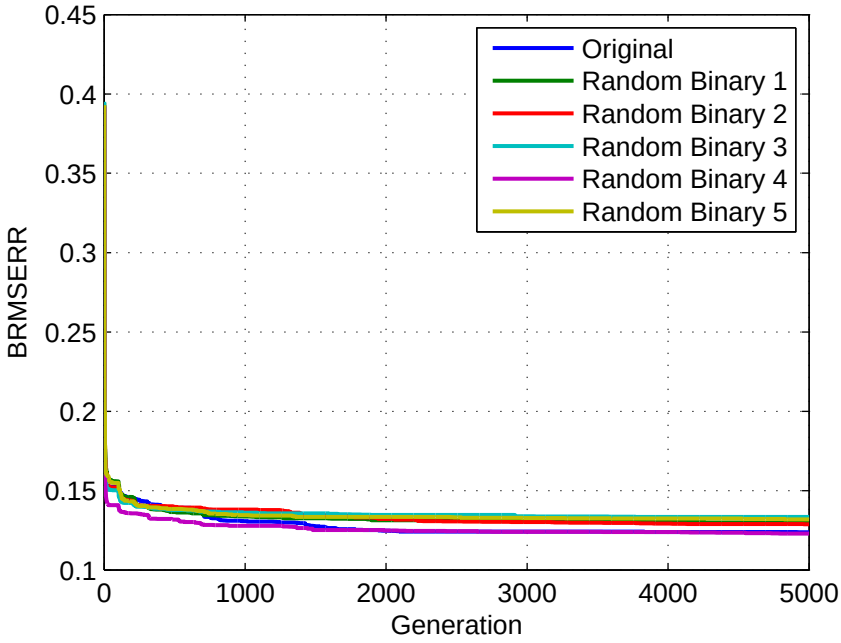

Figure 5.6 Best RMSERRs of the robustness test 2

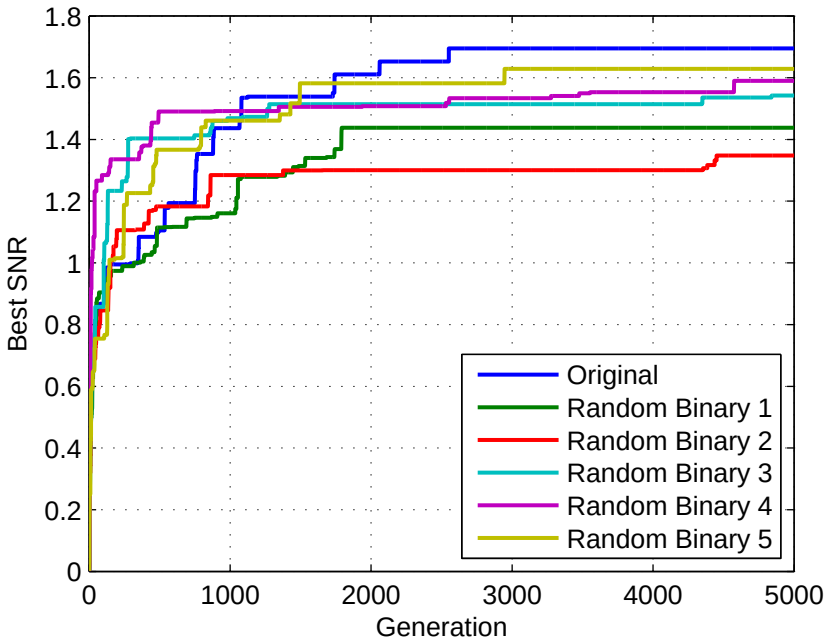

Figure 5.7 Best SNRs of the robustness test 2

SNR Ranking	Robustness Test 1 Population No.	Robustness Test 2 Population No.
1	4	original
2	3	5
3	1	4
4	original	3
5	2	1
6	5	2

Table 5.1 SNR ranking of the robustness test

5.3 Comprehensive Performance of a Single Population

In the selection process of the algorithm shown in Fig. 5.1, the "Elitism" was adopted. To construct the gene pool for the crossover process, the chromosomes with the highest efficiency, the lowest RMSERR, and the highest SNR had one more copy than other chromosomes. However, in order to keep the versatility of gene information, a rule was forced on the selection process so that the highest-efficiency, the lowest-RMSERR and the highest-SNR chromosomes are different ones. Therefore, in the results of previous sections, it is very likely that no chromosomes possessed the highest efficiency and SNR and the lowest RMSERR simultaneously. In this section, the efficiency, the RMSERR, and the SNR of each chromosome in the population after the evolution are investigated to determine the chromosome with the best performance.

As the test conducted in determining the termination conditions in the previous chapter, 10 simulations were conducted. Then the best and the worst results were used to investigate the performance of the whole final population. Since TACO operator is insensitive to the initial population, as shown previously, the original population in [13] was used as the initial population.

The results of the 10 simulations are shown in Figs. 5.8–5.10. Since the efficiency and the RMSERR of each simulation were almost the same, the SNR was used to determine which results were used to investigate the performance. As shown in Fig. 5.10, the 7th simulation resulted in the highest SNR while the 6th simulation yielded the lowest one.

Tables 5.2 and 5.3 show the performance data of the best chromosomes of the 6th and 7th simulation results, respectively. It can be seen from these tables that the SNR of the chromosome with the highest efficiency was relatively low. In particular, the SNR was only as low as a half of the highest SNR. On the other hand, the efficiency of the

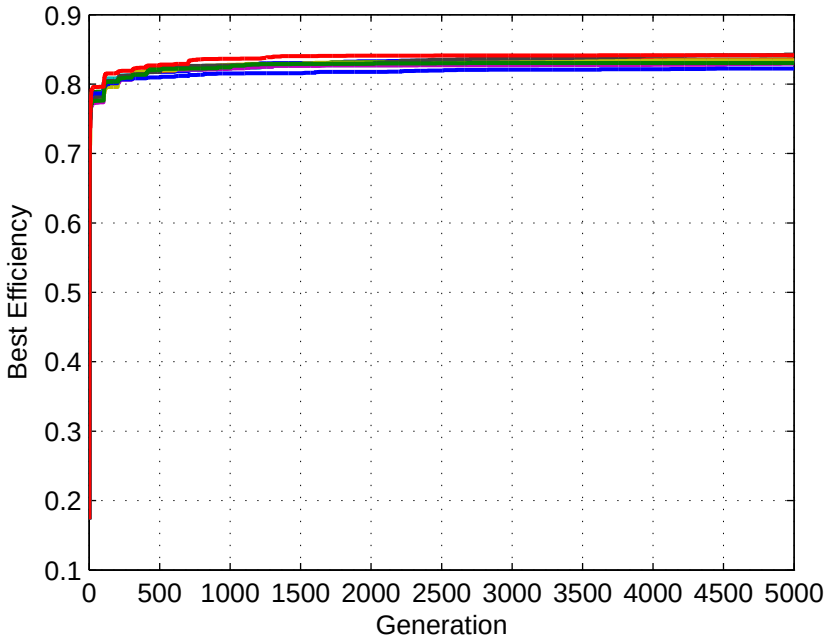

Figure 5.8 Best Efficiencies of the 10 simulations used in population performance investigation

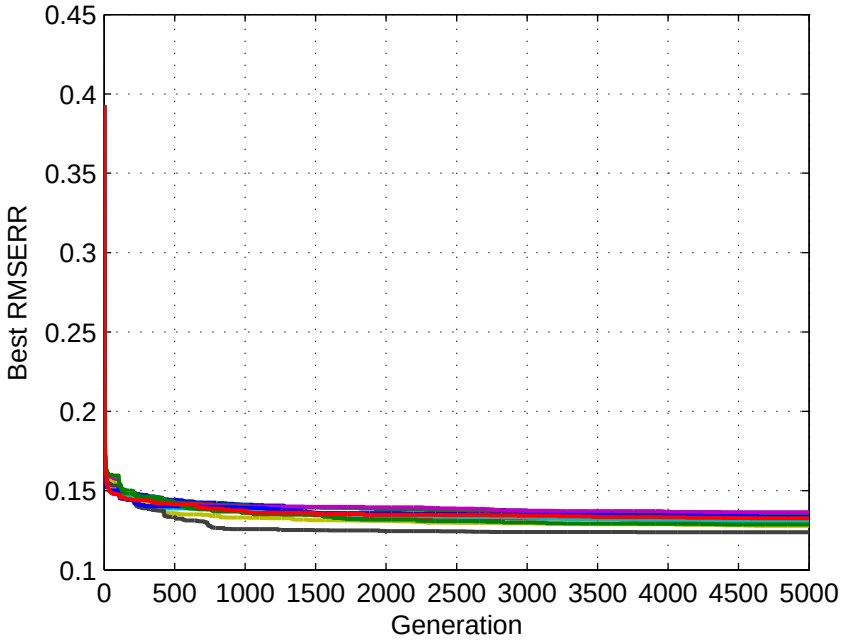

Figure 5.9 Best RMSERRs of the 10 simulations used in population performance investigation

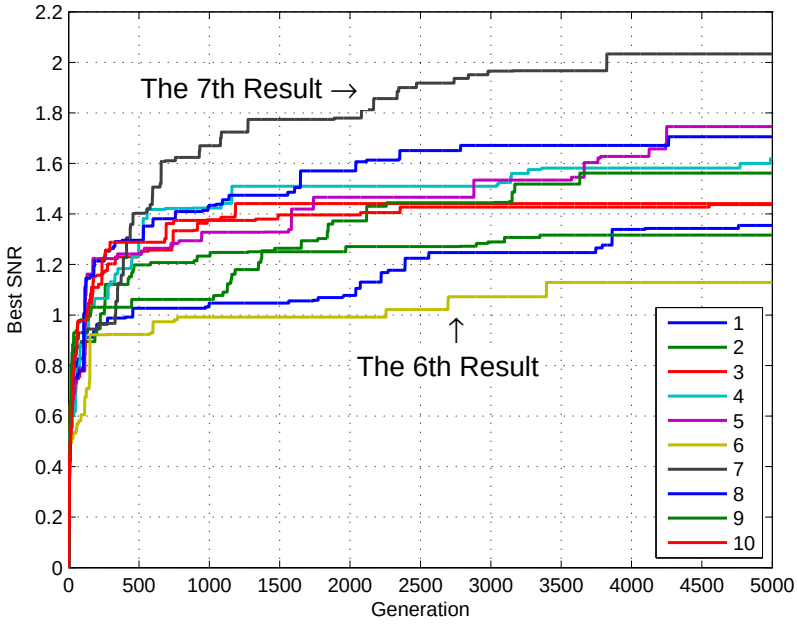

Figure 5.10 Best SNRs of the 10 simulations used in population performance investigation

chromosome with the highest SNR, albeit somewhat lower, was still higher than 0.75. Specifically, the efficiency was only 5% to 10% lower for the chromosomes with the highest SNR than the highest efficiency.

Figures 5.11–5.13 show the comprehensive information of the whole population of the 6th simulation result. Figures 5.14–5.16 show the comprehensive information of the whole population of the 7th simulation result. According to these figures, the following facts were observed:

- chromosomes with the highest efficiencies had relatively low SNRs, as low as a half of that of the highest SNRs;
- chromosomes with the highest SNRs had moderate efficiencies and RMSERRs; and
- the average SNR of the whole population was not so high, although chromosomes with high SNRs existed.

The 6th Test	Efficiency	RMSERR	SNR
Best Efficiency	0.8341	0.1458	**0.4265**
Best RMSERR	0.8005	0.1278	**0.4780**
Best SNR	0.7576	0.1602	1.1293

Table 5.2 Performance of the best chromosomes of the 6th simulation in population performance investigation

The 7th Test	Efficiency	RMSERR	SNR
Best Efficiency	0.8425	0.1466	**0.9052**
Best RMSERR	0.8096	0.1237	**1.1936**
Best SNR	0.8068	0.1358	2.0333

Table 5.3 Performance of the best chromosomes of the 7th simulation in population performance investigation

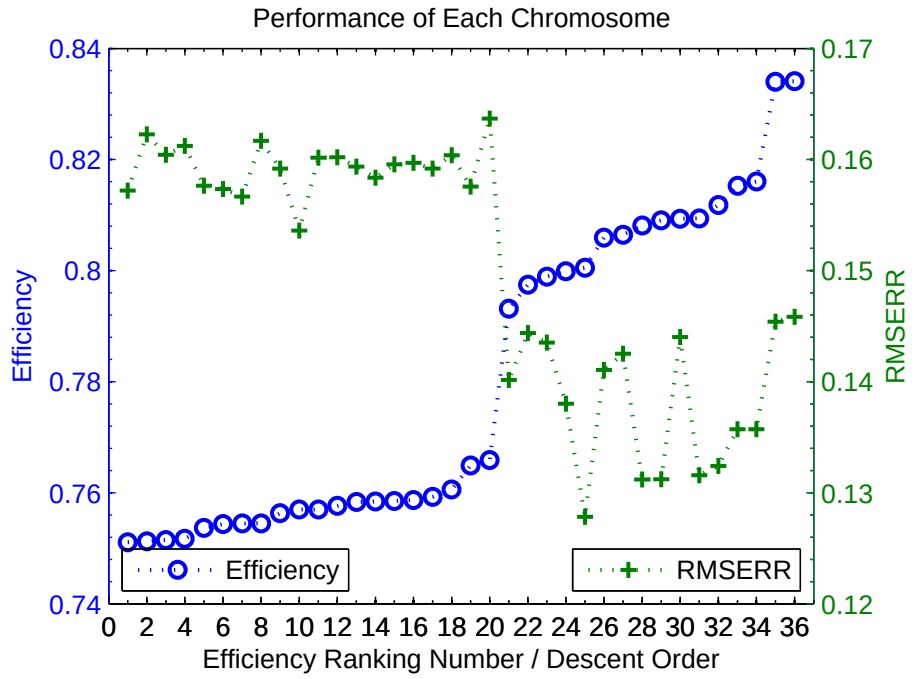

Figure 5.11 Efficiency *v.s.* RMSERR of each chromosome of the 6th simulation in population performance investigation

The trade-off between the efficiency and the SNR may be understood if one reflects on the definitions of these parameters. By definition, the efficiency becomes higher as long as more energy is diffracted into the signal region. In contrast, it is not easy to increase the SNR. As defined in (2.20), only the ratio of the smallest intensity in the signal region to the largest one in the noise region is larger does the SNR increase. Therefore, in the optimization process, chances are that more and more diffracted energy is concentrated into certain pixels of the signal region while the smallest intensity in this region does not become larger or even reduces. As a result, the chromosomes with the highest efficiencies may still possess relative low SNRs.

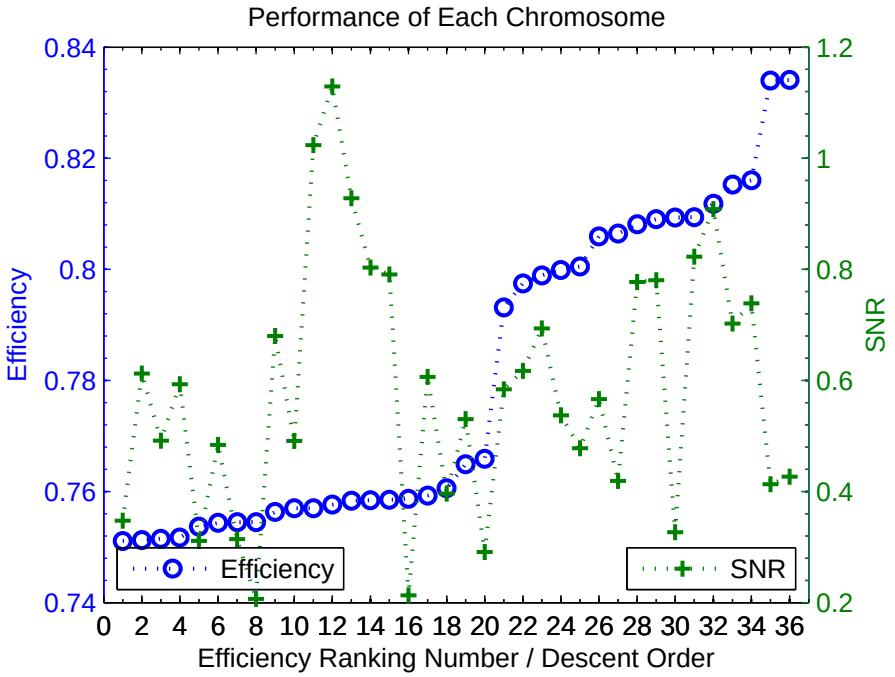

Figure 5.12 Efficiency *v.s.* SNR of each chromosome of the 6th simulation in population performance investigation

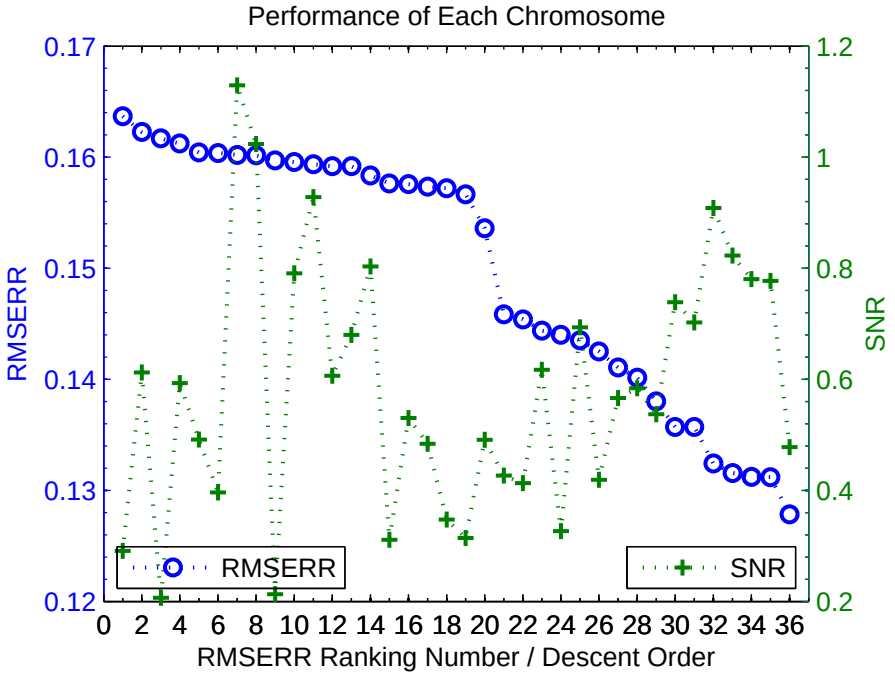

Figure 5.13 RMSERR *v.s.* SNR of each chromosome of the 6th simulation in population performance investigation

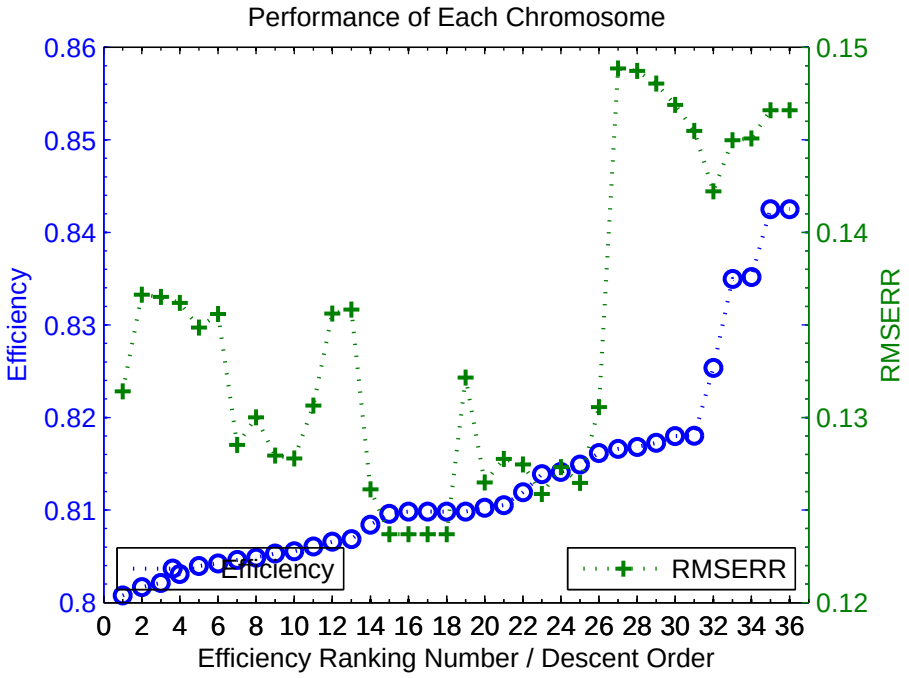

Figure 5.14 Efficiency *v.s.* RMSERR of each chromosome of the 7th simulation in population performance investigation

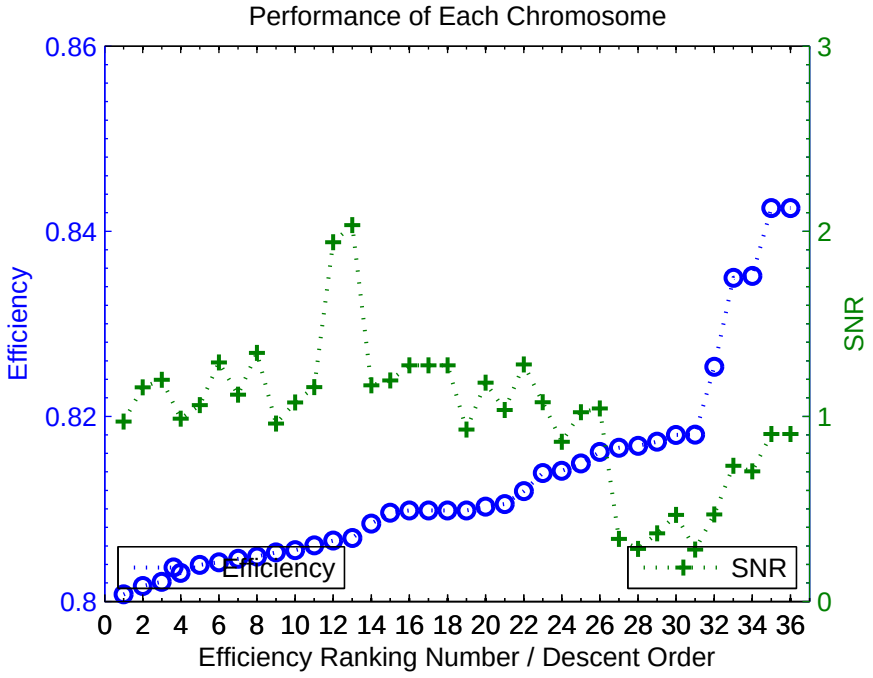

Figure 5.15 Efficiency *v.s.* SNR of each chromosome of the 7th simulation in population performance investigation

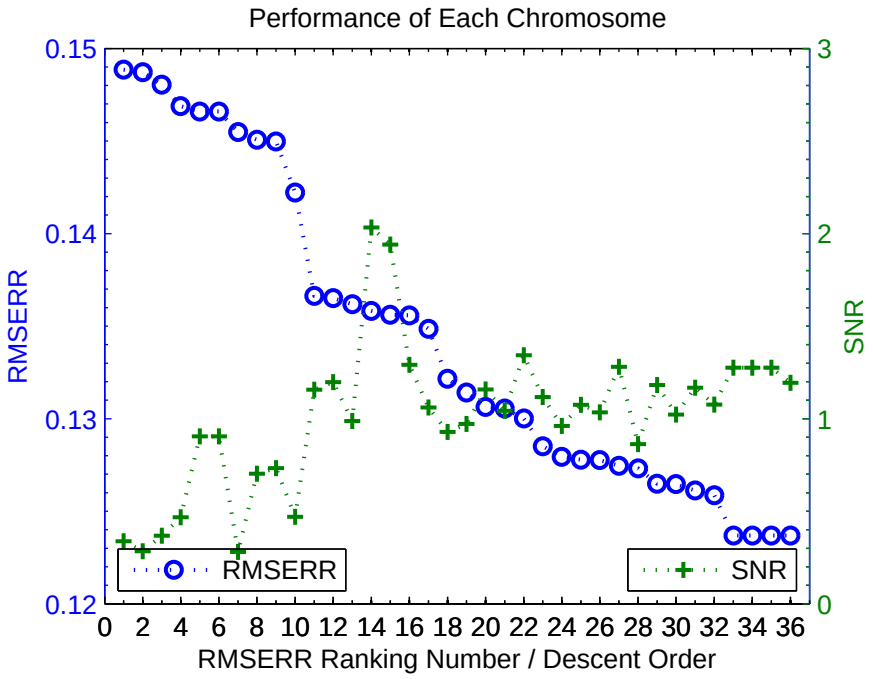

Figure 5.16 RMSERR *v.s.* SNR of each chromosome of the 7th simulation in population performance investigation

5.4 The Phase of the Target Pattern in the TACO Operator

In the simulations conducted in the previous sections, the phase ϕ of the target pattern T was chosen as the phase of $DFT\{P_1\}$. Other choices of ϕ, such as a random phase or a constant phase, are used in this section to make the comparisons.

5.4.1 Random Phase

For the random phase case, ϕ was randomly assigned upon each crossover process. As shown in Figs. 5.17–5.19, the best efficiency, RMSERR and, SNR were about 0.46, 0.3, and 0.024 respectively. The best resulted far field is shown in 5.20. For a comparison, the best far field resulted from another simulation, which used the phase of $DFT\{P_1\}$ as

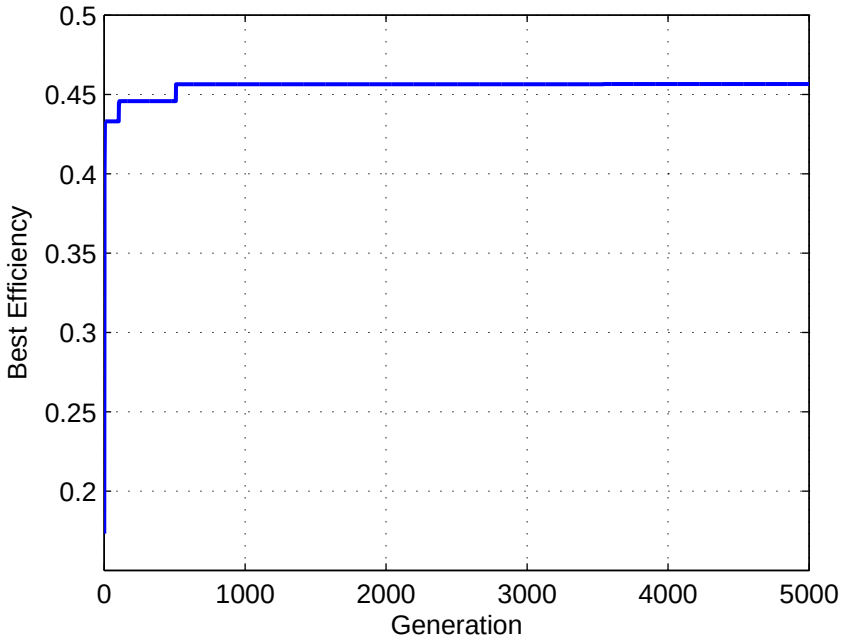

Figure 5.17 Best efficiency of DOEs for random target pattern phase

the phase ϕ, is shown in Fig. 5.21. The unsatisfactory results may be attributed to the lack of the phase retrieval process such as the use of the phase of $DFT\{P_1\}$ which provides an effect equivalent to that of IFTA.

5.4.2 Constant Phase

For the constant phase case, ϕ was selected as 0 without loss of generality. The results are shown in Figs. 5.22–5.25. Under the same number of total generations, it is obvious that the constant phase yielded worse results than the random phase. The best far field is not even distinguishable after 5,000 generations of evolution.

The low performance compared with the above random phase case may be explained as follows. The target pattern with a constant phase can be viewed as a single optimal point in the solution space and may be difficult for the population to approach in the optimization process. The target pattern with the randomly generated phase, produces

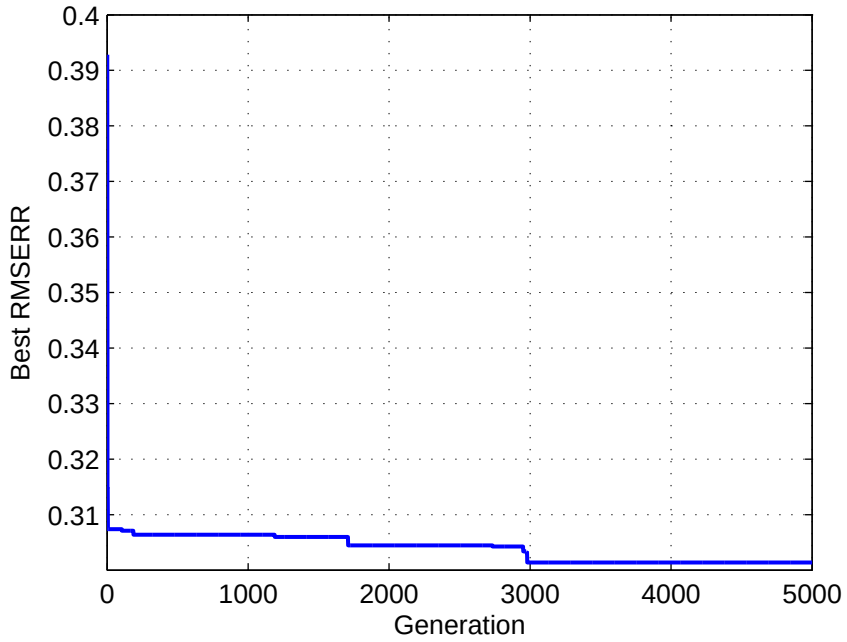

Figure 5.18 Best RMSERR of DOEs for random target pattern phase

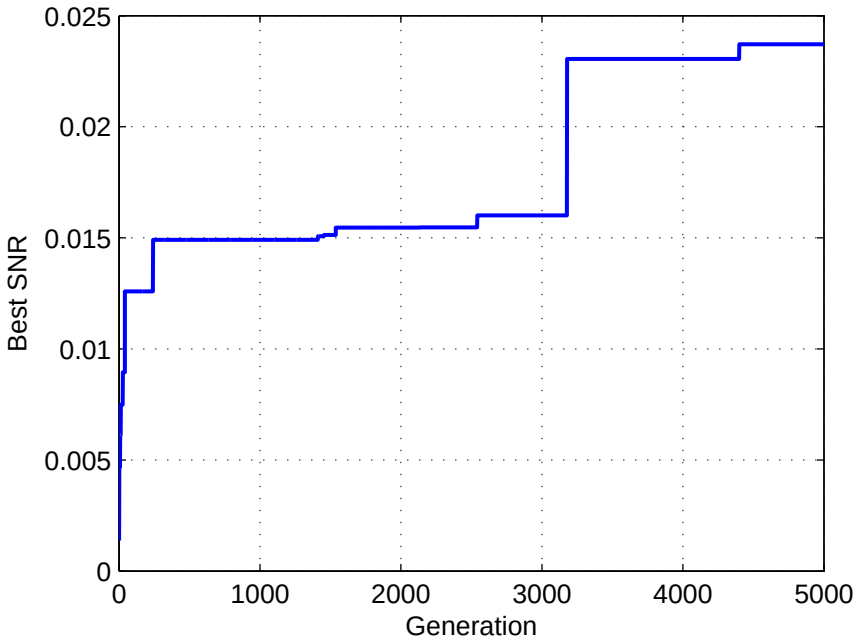

Figure 5.19 Best SNR of DOEs for random target pattern phase

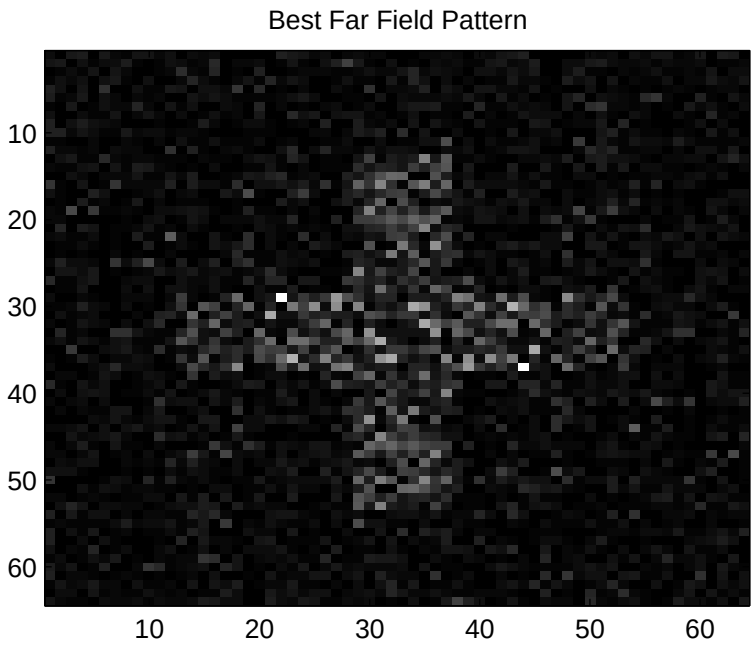

Figure 5.20 Best far field of DOEs for random target pattern phase

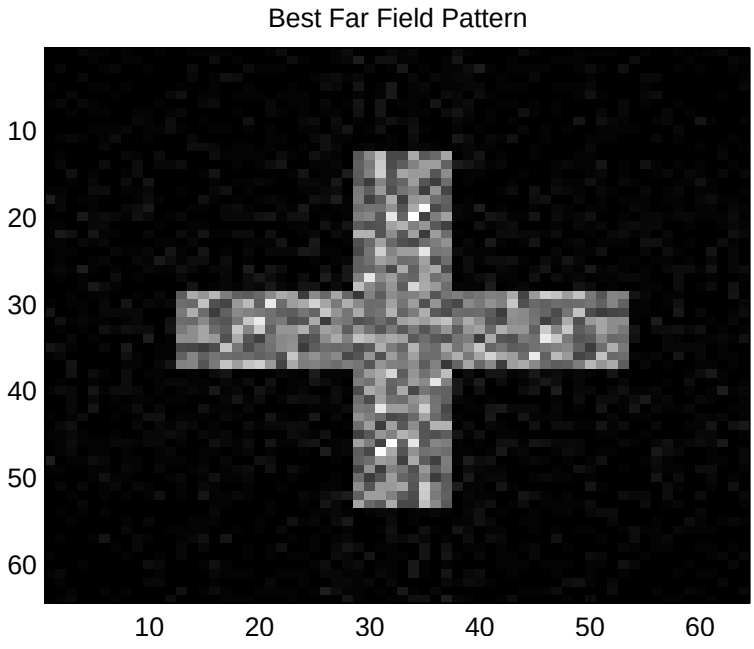

Figure 5.21 Best far field of the simulation using the phase of $DFT\{P_1\}$ as the phase ϕ of the target pattern

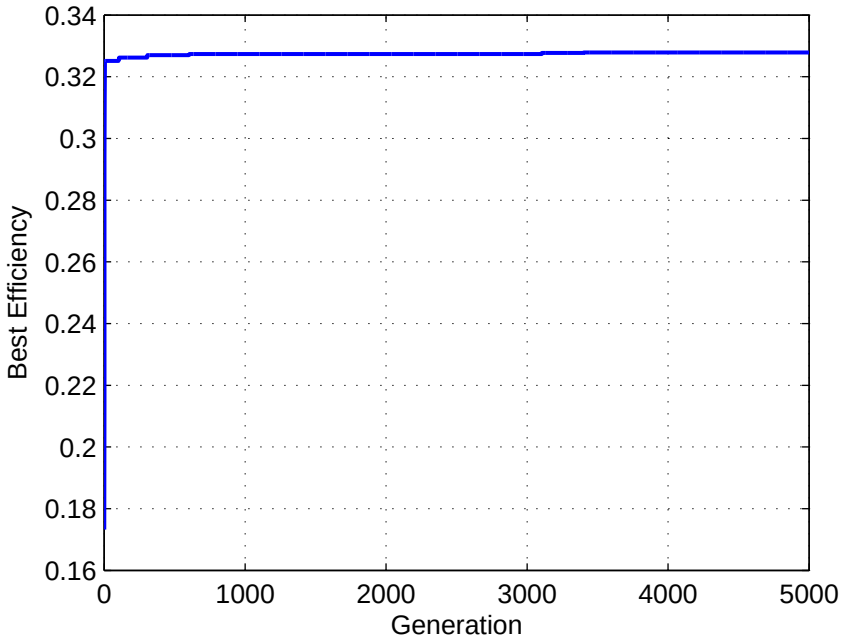

Figure 5.22 Best efficiency of DOEs for constant target pattern phase

more possible optimal points in the muti-dimensional solution space. Therefore, during the optimization process, the population may evolve toward the optimal points whenever is possible. A simulation of a total of 500,000 generations was also conducted and resulted in an efficiency of 0.38, an RMSERR of 0.32, and an SNR of 0.035. Therefore, about 100 times of the total generations were needed for the constant phase case to result in performance comparable to the random phase case.

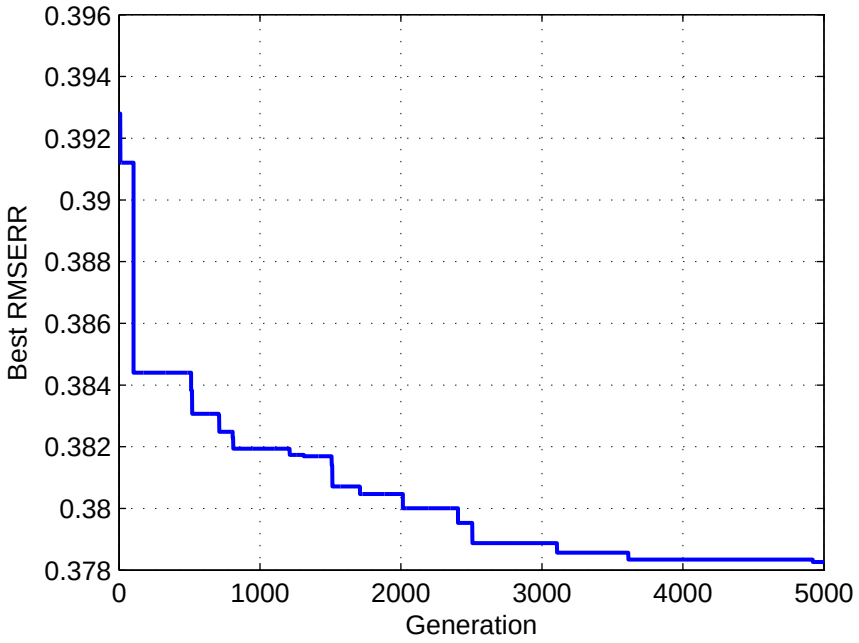

Figure 5.23 Best RMSERR of DOEs for constant target pattern phase

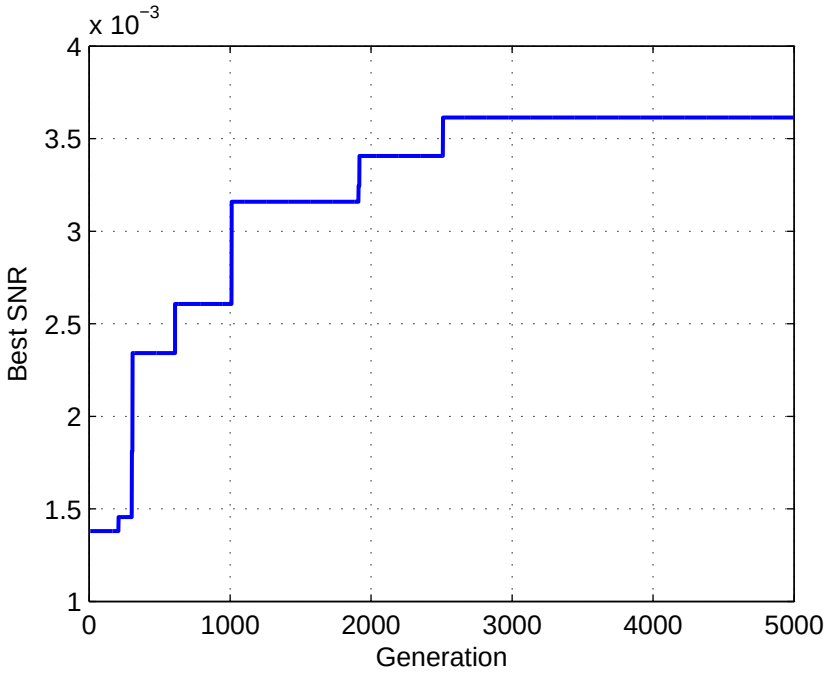

Figure 5.24 Best SNR of DOEs for constant target pattern phase

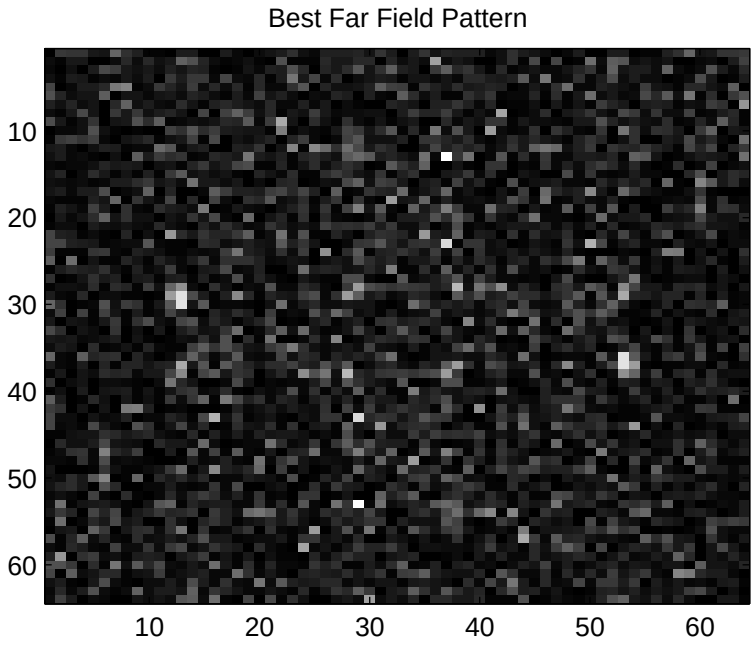

Figure 5.25 Best far field of DOEs for constant target pattern phase

5.5 Application to Design of Multi-Level DOEs

To verify the application of the algorithm with the TACO operator to the design of multi-level DOEs, the four- and eight-level cases were conducted.

5.5.1 Four-level DOEs

Figure 5.26 shows the four-level quantization process. The continuous phase is quantized to 0, $\pi/2$, π, and $3\pi/2$. Similar to that of the binary cases, 10 simulations were conducted. The results of the 10 simulations of the four-level DOE design are shown in Figs. 5.27–5.29.

The first impression of the results may be unfavorable since there is little difference between the four-level cases and the binary cases. However, it is easily understood if one keeps in mind that a binary phase DOE reconstructs both the original and the conjugate images which are symmetric about the axis [18]. For a target with on-axis symmetry, the

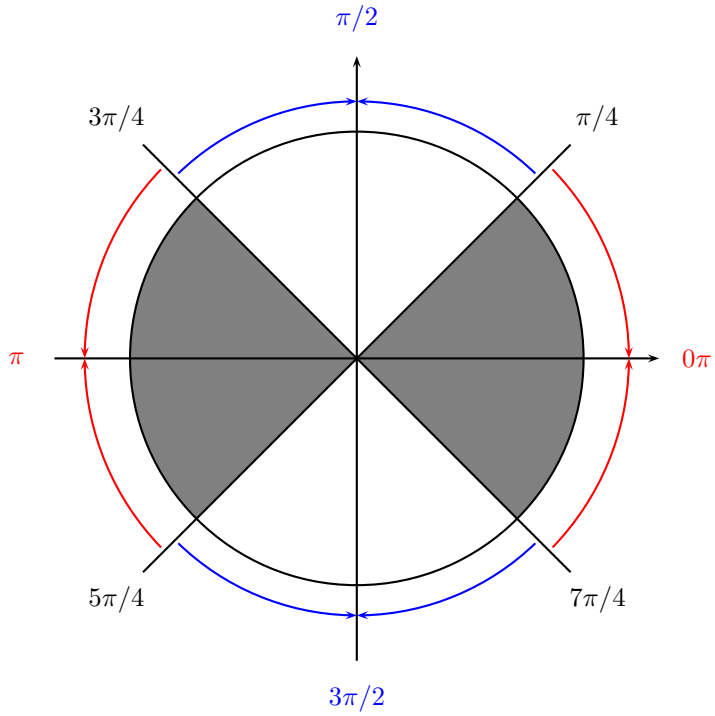

Figure 5.26 Four-level quantization

original and the conjugate images merge and there seems to be only one reconstructed image. The cross pattern used in the simulations is symmetric about the axis. Consequently, the efficiencies, the RMSERRs, and the SNRs calculated in the binary cases were contributed by both of the original and the conjugate images. For example, the efficiency of the original image was only a half of that of the merged image. That is, the efficiencies were about 0.4 for the binary cases, which make the results of the four-level DOEs reasonable.

5.5.2 Eight-level DOEs

Figure 5.30 shows the eight-level quantization process. The continuous phase is quantized to 0, $\pi/4$, $\pi/2$, $3\pi/4$, π, $5\pi/4$, $3\pi/2$, and $7\pi/4$.

Figures 5.31–5.33 show the results of the 10 simulations of the eight-level DOE de-

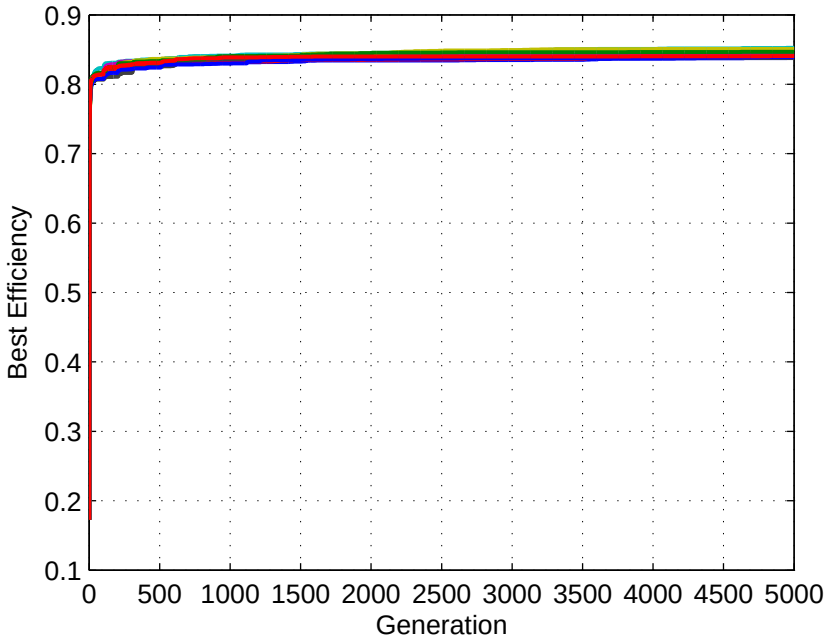

Figure 5.27 Best efficiencies of the four-level DOEs

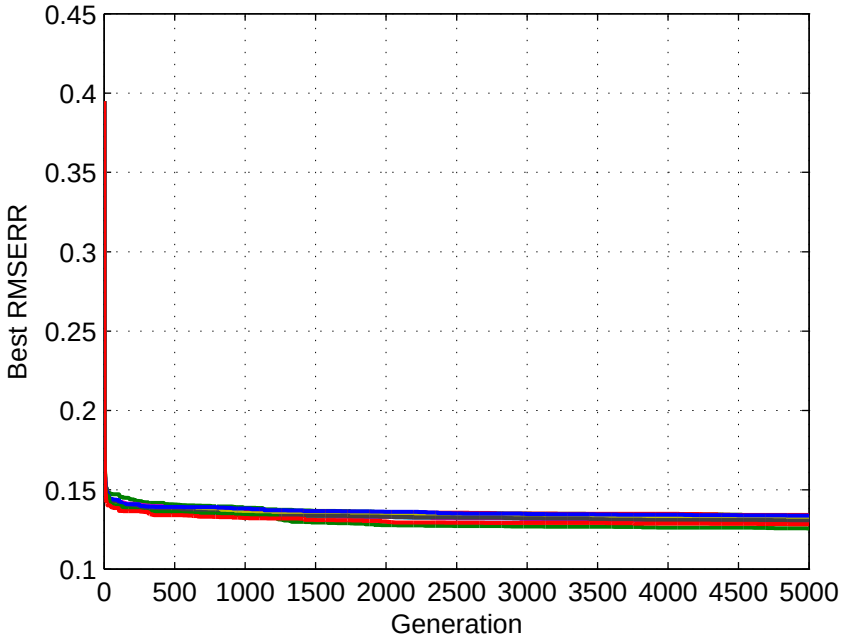

Figure 5.28 Best RMSERRs of the four-level DOEs

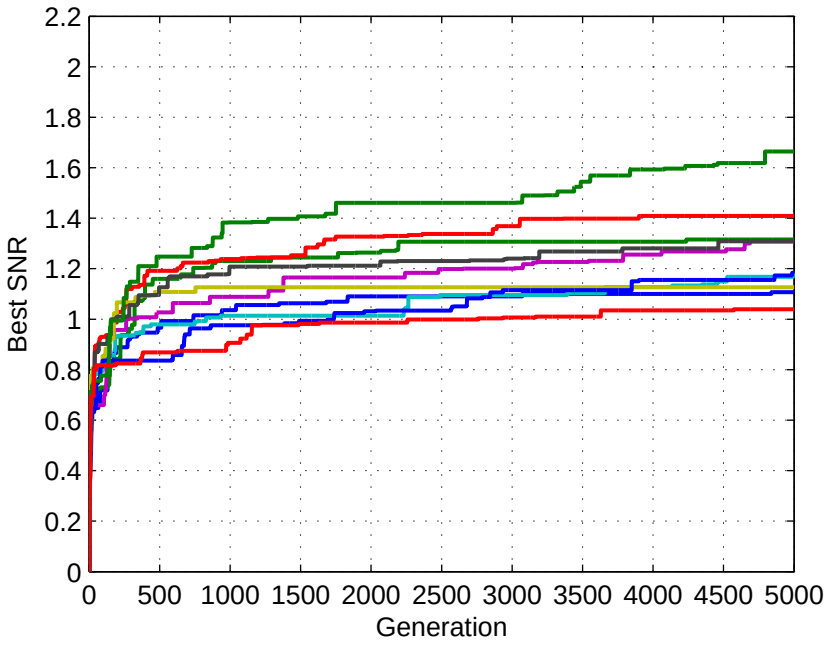

Figure 5.29 Best SNRs of the four-level DOEs

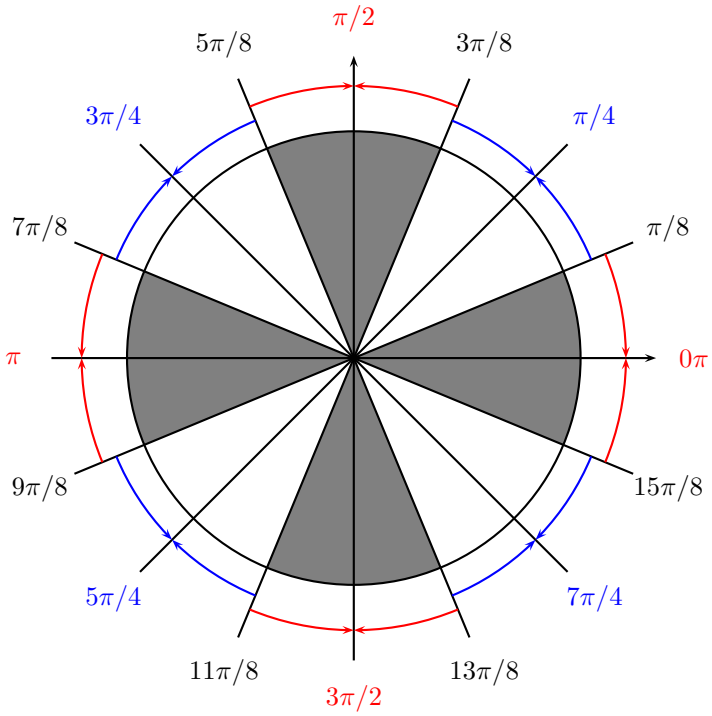

Figure 5.30 Eight-level quantization

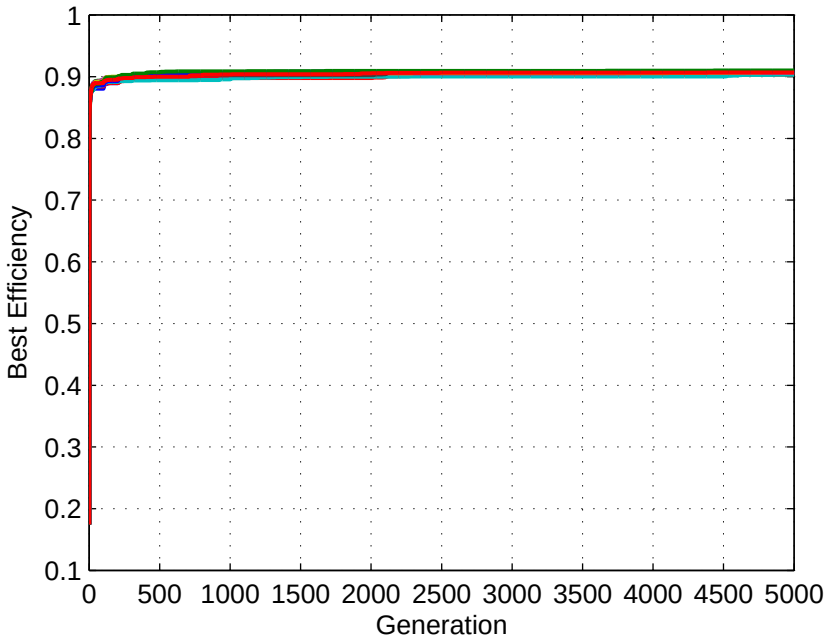

Figure 5.31 Best efficiencies of the eight-level DOEs

sign. The best efficiencies, RMSERRs, and SNRs were as good as 0.9, 0.1, and 4.4 respectively. Compared with the binary and the four-level cases, the eight-level DOEs yielded superior performance, as expected.

Therefore, from the results of the four- and eight-level DOEs, it is concluded that the TACO operator is applicable to the design of multi-level DOEs. The application to design of continuous DOEs was also implemented with superior performance—an efficiency of 0.95, an RMSERR of 0.07, and an SNR of 7.

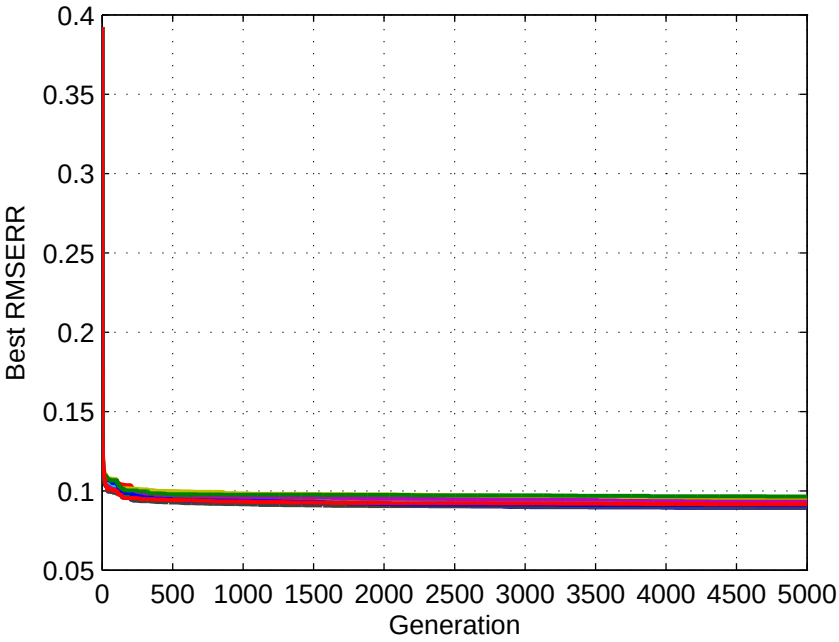

Figure 5.32 Best RMSERRs of the eight-level DOEs

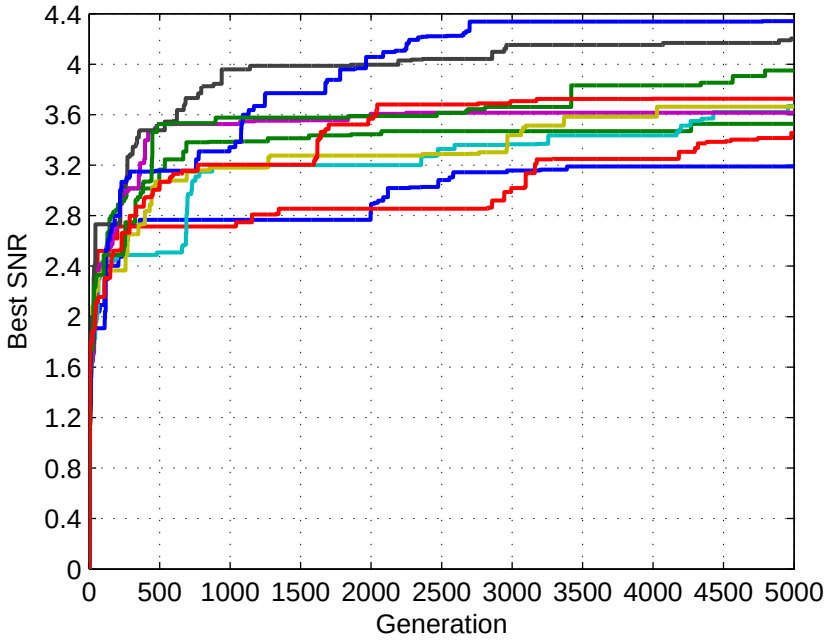

Figure 5.33 Best SNRs of the eight-level DOEs

5.6 Comparisons with Other Optimization Methods

In this section, the proposed method is compared with other optimization methods. First, the results of the iterative Fourier transform algorithm (IFTA) are compared with our results as illustrated in the previous sections. Then the results of the simple genetic algorithm (SGA) used in [14] are briefly described and compared with our results.

5.6.1 IFTA with Iterative Quantization Method

The theory and the application of IFTA has been widely studied since it was proposed by Gerchberg and Saxton in 1971 [10]. Figure 5.34 shows the procedure of IFTA. There are many implementations of IFTA to optimize the performance of the results. In this section, the iterative quantization method [29] was used. To compare with the GA with the TACO operator, the simulation of IFTA used the following conditions:

- the binary phase DOEs were designed;
- the initial guesses were the original 33 chromosomes used in the previous sections; and
- a total 18 steps were used in the quantization process, i.e. 10 degrees per step.

As shown in Table 5.4, the resulted DOEs had efficiencies ranging from 0.8501 to 0.8674, RMSERRs ranging from 0.1549 to 0.1327, and SNRs ranging from 0.1695 to 0.8867. The average efficiency, RMSERR, and SNR are 0.8615, 0.1432 and, 0.5267, respectively. A comprehensive performance of the 33 DOEs are listed in Appendix A.

To compare with the results of the algorithm with the TACO operator, the performance of DOEs with the best efficiency, RMSERR and SNR was considered, which is shown in Table 5.5. In particular, the DOE with the highest efficiency and SNR was selected as the

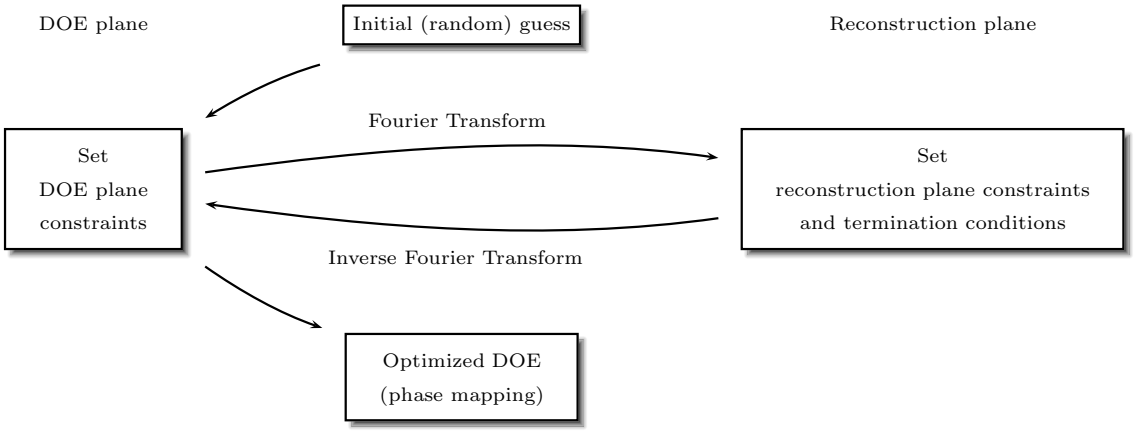

Figure 5.34 The schematic diagram of IFTA

IFTA Results	Best	Worst	Average
Efficiency	0.8674	0.8501	0.8615
RMSERR	0.1327	0.1549	0.1432
SNR	0.8867	0.1695	0.5267

Table 5.4 Results of IFTA (Note: resulted DOEs with best efficiency, RMSERR and SNR may not be the same ones; the same is true for the worst cases.)

IFTA Results	Efficiency	RMSERR	SNR
Best Efficiency	0.8647	0.1375	0.8867
Best RMSERR	0.8538	0.1327	0.7671
Best SNR	0.8647	0.1375	0.8867

Table 5.5 Performance of Best DOEs resulted from IFTA Simulations

optimal solution of the IFTA. The results of the 6th and the 7th simulations in Section 5.3 were used in the comparison. As shown in Table 5.2, it is reasonable to choose the chromosome with the highest SNR to be the final optimal DOE since a 2.5 times of SNR can compensate for the 10% loss of efficiency. Similarly, the optimal DOE in the 7th simulation was selected in the same way. The best solutions are listed in Table 5.6.

Compared with the IFTA, the GA with the TACO operator yielded DOEs with lower efficiencies, comparable RMSERRs, and higher SNRs. Specifically, the efficiencies were about 7.5% to 10% lower and the SNRs were about 1.3 to 2.3 times higher. Therefore, with a little decrease of the efficiency, the GA with the TACO operator results in DOEs with superior performance. Also, the computation time was comparable for both methods to yield the optimal results listed in Table 5.6. Under our computing environment, which is a laptop installed with the Intel® Core™Duo L2400 1.66 GHz processor and the MATLAB® R2007b computing package, the time duration was about a half of an hour.

5.6.2 The GA using the Mutation Mechanism

The method used in [14] was an SGA. Different mutation rates were employed to find out a best one which resulted an optimized solution. Note that the initial population used

	Efficiency	RMSERR	SNR
IFTA	**0.8647**	0.1375	0.8867
The 6th of TACO	0.7576	0.1602	1.1293
The 7th of TACO	0.8068	**0.1358**	**2.0333**

Table 5.6 Performance of Best DOEs resulted from IFTA and GA with TACO operator

in the 6th and 7th simulation results shown in Table 5.6 is the same as that used in [14].

Using a mutation rate of 0.04% and a random crossover pattern to design the binary phase DOEs, the method used in [14] resulted in a DOE with an efficiency of 0.8497, an RMSERR of 0.1274, and an SNR of 2.3739, after 100,000 generation of evolution. The performance of the DOE was comparable to the 7th simulation result of our method as shown in Table 5.6. However, the method in [14] was far more computation expensive than our method. Specifically, the total number of generations used in [14] was 20 times of that used in our method. In terms of time duration, our method resulted in the 7th simulation result within a half of an hour, while eight to ten hours were needed for the evolution of 100,000 generations in [14].

Chapter 6

Conclusion

6.1 Conclusion on the algorithm with the TACO operator

In this thesis, a method that introduced the novel TACO operator into the GA has been proposed to enhance the performance of the optimization in the DOE design. The TACO operator was derived by taking the advantage of the existence of the well-defined target pattern of the optimization problem.

The performance of the proposed method used in the phase DOE design was studied. The method yielded DOEs with satisfactory efficiencies, RMSERRs, and SNRs. The computation time needed for the method was effectively reduced to achieve results comparable with the SGA using the mutation mechanism [14]. Specifically, the number of generations needed in the method was only 1/20 of that in [14]. Compared with the IFTA using iterative quantization, the method resulted in binary-phase DOEs with higher SNRs and comparable RMSERRs. The applications of the proposed method to the multi-level and the continuous DOE design were also implemented with satisfactory performance.

The characteristics of the TACO operator were investigated, too. The robustness of

the TACO operator against the initial population was confirmed. Depending on the choice of the phase of the target pattern in the TACO operator, the resulted DOE may possess different performance. The use of the diffractive field phase resulted in superior performance while the random phase and the constant phase provided limited increase in performance. We believe the reasons are the lack of the effect of the phase retrieval process similar to the IFTA and the lack of the approachable optimal points in the solution space. The performance of the whole population after the final evolution was evaluated. The trade-off between the SNR and the efficiency was also observed in the whole population.

6.2 Improvement of the Algorithm

The performance of the proposed algorithm may be further improved by using several methods. For example, the iterative quantization technique [29, 30] may be adopted to pursue higher performance. By introducing the constraints in a stepwise way iteratively, more design degree of freedom can be provided. Therefore, the stagnation of the performance may be avoid and the movement toward the global maximum may be more smoothly. In particular, the efficiency may be expected to increase to the level as that of the IFTA described in Section 5.6.1. However, the computation expense and the time duration of the simulation may still be an issue for a noticeable improvement.

References

[1] A. Sommerfeld, *Optics: Lectures on Theoretical Physics*, vol. IV. Academic Press, Inc., New York, 1954.

[2] S. Sinzinger and J. Jahns, *Microoptics*. Wiley-VCH, 2003.

[3] F. Dickey, S. Holswade, and D. Shealy, *Laser beam shaping applications*. Taylor & Francis, Boca Raton, FL, 2005.

[4] G.-Y. Yoon, S. Matsuoka, T. Jitsuno, M. Nakatsuka, and Y. Kato, "Wave-front design algorithm for shaping a quasi-far-field pattern," *Appl. Opt.*, vol. 37, no. 8, pp. 1386–1392, 1998.

[5] J. S. Liu and M. R. Taghizadeh, "Iterative algorithm for the design of diffractive phase elements for laser beam shaping," *Opt. Lett.*, vol. 27, no. 16, pp. 1463–1465, 2002.

[6] M. A. Seldowitz, J. P. Allebach, and D. W. Sweeney, "Synthesis of digital holograms by direct binary search," *Appl. Opt.*, vol. 26, no. 14, p. 2788, 1987.

[7] M. Clark, "Enhanced direct-search method for the computer design of holograms using state variables," in *Proc. SPIE*, vol. 2689, (San Jose, CA, USA), pp. 24–34, SPIE, May 1996.

[8] N. Yoshikawa and T. Yatagai, “Phase optimization of a kinoform using simulated annealing,” *Appl. Opt.*, vol. 33, no. 5, p. 863, 1994.

[9] M. S. Kim and C. C. Guest, “Simulated annealing algorithm for binary phase only filters in pattern classification,” *Appl. Opt.*, vol. 29, no. 8, p. 1203, 1990.

[10] R. Gerchberg and W. Saxton, “A practical algorithm for the determination of phase from image and diffraction plane pictures,” *Optik*, vol. 35, no. 2, pp. 237–246, 1972.

[11] N. Yoshikawa, M. Itoh, and T. Yatagai, “Quantized phase optimization of two-dimensional fourier kinoforms by a genetic algorithm,” *Opt. Lett.*, vol. 20, no. 7, p. 752, 1995.

[12] G. Zhou, Y. Chen, Z. Wang, and H. Song, “Genetic local search algorithm for optimization design of diffractive optical elements,” *Appl. Opt.*, vol. 38, no. 20, pp. 4281–4290, 1999.

[13] T.-C. Hou, “The study of crossover mechanism of genetic algorithm for diffractive optical elements,” Master’s thesis, National Taipei University of Technology, July 2007.

[14] C.-M. Lin, “Study of mutation system of genetic algorithm for optimization design of diffractive optical elements,” Master’s thesis, National Taipei University of Technology, July 2007.

[15] J. W. Goodman, *Introduction to Fourier Optics*. Roberts & Company Publishers, 2005.

[16] H. Haus, *Waves and fields in optoelectronics*. Prentice-Hall, Englewood Cliffs, NJ, 1984.

[17] R. N. Bracewell, *The Fourier Transform and its Applications*. McGraw-Hill, 3rd ed., 1999.

[18] B. Kress and P. Meyrueis, *Digital diffractive optics: An Introduction to Planar Diffractive Optics and Related Technology*. Wiley, 2000.

[19] J. H. Holland, *Adaptation in natural and artificial systems*. The University of Michigan Press, Ann Arbor, MI, 1975.

[20] D. E. Goldberg, *Computer-aided gas pipeline operation using genetic algorithms and rule learning*. PhD thesis, University of Michigan, Ann Arbor, Michigan, 1983.

[21] M. Srinivas and L. M. Patnaik, "Genetic algorithms: a survey," *Computer*, vol. 27, no. 6, pp. 17–26, Jun 1994.

[22] 周鵬程, ***遺傳演算法原理與應用–活用*** *Matlab*. 全華科技圖書股份有限公司, 2nd ed., 2005.

[23] B. L. Miller and D. E. Goldberg, "Genetic algorithms, selection schemes, and the varying effects of noise," *Evolutionary Computation*, vol. 4, no. 2, pp. 113–131, 1996.

[24] K. A. De Jong and W. M. Spears, "A formal analysis of the role of multi-point crossover in genetic algorithms," *Annals of Mathematics and Artificial Intelligence*, vol. 5, pp. 1–26, Mar. 1992.

[25] G. Sywerda, "Uniform crossover in genetic algorithms," in *Proceedings of the third international conference on Genetic algorithms*, (San Francisco, CA, USA), pp. 2–9, Morgan Kaufmann Publishers Inc., 1989.

[26] W. Spears, “Crossover or mutation,” in *Proceedings of the Foundations of Genetic Algorithms Workshop*, Morgan Kaufmann, 1992.

[27] K. De Jong and W. Spears, *Parallel Problem Solving from Nature*, vol. 496 of *Lecture Notes in Computer Science*, ch. An analysis of the interacting roles of population size and crossover in genetic algorithms, pp. 38–47. Springer Berlin / Heidelberg, 1991.

[28] K. Tang, K. Man, S. Kwong, and Q. He, “Genetic algorithms and their applications,” *Signal Processing Magazine, IEEE*, vol. 13, no. 6, pp. 22–37, 1996.

[29] W.-F. Hsu, “Backward iterative quantization methods for designs of multilevel diffractive optical elements,” *Opt. Express*, vol. 13, no. 13, pp. 5052–5063, 2005.

[30] F. Wyrowski, “Iterative quantization of digital amplitude holograms,” *Appl. Opt.*, vol. 28, no. 18, pp. 3864–3870, 1989.

Appendices

Appendix A

Comprehensive Results of IFTA

The efficiency, the RMSERR, and the SNR of the 33 resulted DOEs optimized by the IFTA as described in Section 5.6 are shown in the following tables. The initial guesses of the 33 IFTA results are the 33 chromosomes of the original population used in the GA in Chapter 5.

No.	Efficiency	RMSERR	SNR	No.	Efficiency	RMSERR	SNR
1	0.8627	0.1430	0.2284	9	0.8538	**0.1327**	0.7671
2	0.8608	0.1371	0.3667	10	0.8572	0.1419	0.5732
3	0.8588	**0.1549**	0.3874	11	0.8601	0.1459	**0.1695**
4	0.8585	0.1467	0.2786	12	0.8531	0.1437	0.5497
5	0.8578	0.1393	0.5664	13	**0.8674**	0.1375	**0.8867**
6	0.8581	0.1410	0.7785	14	0.8589	0.1440	0.5329
7	0.8644	0.1464	0.5414	15	0.8565	0.1422	0.2693
8	0.8642	0.1417	0.5772	16	0.8657	0.1489	0.5405

Table A.1 The comprehensive IFTA results

No.	Efficiency	RMSERR	SNR
17	0.8624	0.1410	0.3826
18	0.8572	0.1377	0.6336
19	0.8569	0.1409	0.3722
20	0.8619	0.1457	0.5182
21	0.8520	0.1417	0.8631
22	0.8638	0.1462	0.2915
23	0.8638	0.1473	0.4098
24	0.8607	0.1420	0.4739
25	0.8554	0.1386	0.3548
26	0.8572	0.1480	0.3725
27	0.8624	0.1457	0.3962
28	0.8623	0.1442	0.3962
29	0.8646	0.1491	0.6289
30	**0.8501**	0.1400	0.7545
31	0.8665	0.1400	0.6032
32	0.8576	0.1415	0.8588
33	0.8627	0.1430	0.2284

Table A.2 The comprehensive IFTA results (contd.)