

Optimal design of optical reference signals by use of a genetic algorithm

José Saez-Landete, Sancho Salcedo-Sanz, and Manuel Rosa-Zurera

Departamento de Teoría de la Señal y Comunicaciones, Escuela Politécnica Superior, Universidad de Alcalá, Campus Universitario, 28805 Alcalá de Henares, Madrid, Spain

José Alonso and Eusebio Bernabeu

Departamento de Óptica, Facultad de Ciencias Físicas, Universidad Complutense de Madrid, Ciudad Universitaria, 28040 Madrid, Spain

Received May 24, 2005; revised manuscript received June 20, 2005; accepted June 25, 2005

A new technique for the generation of optical reference signals with optimal properties is presented. In grating measurement systems a reference signal is needed to achieve an absolute measurement of the position. The optical signal is the autocorrelation of two codes with binary transmittance. For a long time, the design of this type of code has required great computational effort, which limits the size of the code to ~ 30 elements. Recently, the application of the dividing rectangles (DIRECT) algorithm has allowed the automatic design of codes up to 100 elements. Because of the binary nature of the problem and the parallel processing of the genetic algorithms, these algorithms are efficient tools for obtaining codes with particular autocorrelation properties. We design optimum zero reference codes with arbitrary length by means of a genetic algorithm enhanced with a restricted search operator. © 2005 Optical Society of America

OCIS codes: 120.0120, 120.3940, 230.0230.

Measuring position is especially important in precision engineering, nanoscience, and nanotechnology. The increasing demand for high-resolution grating measurement systems has created a strong incentive for the design of systems that generate zero reference signals. A zero reference signal is an important addition to an incremental displacement measurement system for achieving absolute measurement, finding the origin of a coordinate, or a machine home position. To acquire a zero reference signal, gratings with adjacent zero reference codes (ZRCs) have been developed since 1986.¹ The ZRC is a group of unequally spaced transparent and opaque slits. Typically the system consists of two opposite ZRCs. The relative displacement between ZRCs produces a change in the overlapping area between them. A parallel ray beam propagates through both codes and the transmitted light is registered by means of a photodiode. The output signal is the reference signal and is the correlation between the two ZRCs. To increase the maximum of this signal, the two codes are identical so that the reference signal becomes the autocorrelation of the ZRC. In Fig. 1 we show the procedure for generation of a reference signal.

The characterization and design of optimum codes to obtain suitable reference signals was studied in Refs. 2 and 3. These works consider the properties of the autocorrelation function and establish the necessary conditions to achieve a suitable signal. The main difficulty in designing ZRCs is that the necessary calculations are laborious and there is not a systematic method for obtaining codes of greater length than 30 elements. Recently, we presented⁴ a new approach to the design of ZRCs based on optimization techniques. The design problem is transformed into a minimization problem with binary variables, and we applied the dividing rectangles (DIRECT) algorithm.⁵ With

this technique we have extended the design capabilities to 100-element codes. To design larger codes, suboptimal solutions can be obtained decreasing the number of objective function evaluations. Computer memory is the limiting factor with the new technique. In this Letter, we propose a genetic algorithm (GA) that definitely breaks these barriers, allowing the optimization of arbitrarily long codes.

In general, a ZRC can be described by the following sequence of binary data $\mathbf{c}=[c_1, c_2, \dots, c_n]$, where n is the length of the ZRC, $c_i=1$ if a transparent slit is located at the i position, and $c_i=0$ elsewhere. The number of transparent slits is n_1 . The sizes of the transparent and opaque regions in the code are integer multiples of the width of a single slit. To calculate the reference signal, we assume that the illuminating light is a parallel ray beam and diffraction effects are negligible. This approach is valid when the gap between ZRCs is close with regard to the width of the code slits and the width of the code slits is greater than the wavelength of the illuminating light. Under these conditions, the light flux that passes through

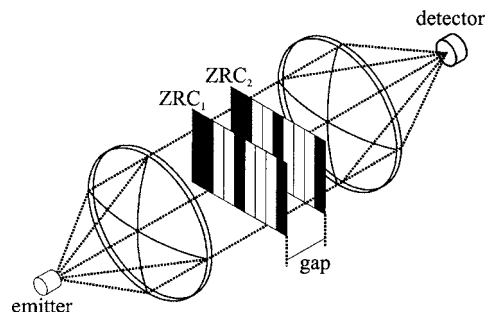


Fig. 1. Procedure for generation of the autocorrelation signal. We consider the gap between ZRCs close with regard to the width of the code slits and the width of the code slits greater than wavelength of the illuminating light.

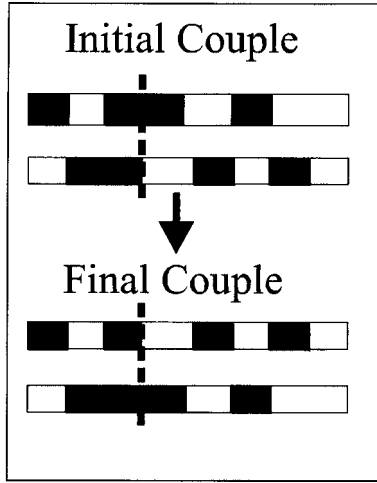


Fig. 2. Example of crossover operation, in which the initial couple swaps its genes from position 3.

the two codes when one of them is shifted by k units is

$$S_k = S_{-k} = \sum_{i=1}^{n-k} c_j c_{j+k}, \quad k = 0, 1, \dots, n-1, \quad (1)$$

where S_k are the components of symmetric vector \mathbf{S} . The maximum of the autocorrelation signal is S_0 and it is equal to the number of transparent slits in the code, $S_0 = \sum_{j=1}^n c_j^2 = n_1$. The second maximum of the signal is named σ and it is $\sigma = \max\{S_1, \dots, S_n\}$. When the number of elements and the number of slits in the ZRC are fixed, there is a theoretical limit in the minimum value of the second maximum. In Ref. 3, Yajun established a conservative lower bound for the second maximum,

$$\sigma \geq \frac{(2n+1) - \sqrt{(2n+1)^2 - 4n_1(n_1-1)}}{2}. \quad (2)$$

Although there are some simple cases in which this bound is reached, up to now to our knowledge there has been no evidence that for any values of n and n_1 at least one ZRC could be found for which the equality sign holds.

In metrological applications, the most important parameters that characterize a zero reference signal are the width and the height of the central maximum and the height of the second maximum. If the slits of the ZRC are rectangles, the central maximum is a triangle whose width in the base is twice the width of the slits of the ZRC. A good zero reference signal must be a single and distinct peak, so the secondary maxima must be as low as possible. Usually, the period of the gratings of the displacement measurement systems determines the width of the reference signal and therefore the width of the slits of the ZRC. Actually, ultrahigh-resolution systems impose the use of very narrow slits (a few micrometers). A significant signal level requires a large number of slits, and then optimal codes with a large number of elements are needed. Also, the width and the number of elements in the ZRC determine the minimum diameter of the

light beam. According with these working requirements, we fix the number of elements of the code and the number of slits. The objective is to minimize the second maximum of the signal, σ . The objective function is

$$\begin{aligned} \min_{\mathbf{c}} f(\mathbf{c}) \\ f(\mathbf{c}) = \max\{S_1, \dots, S_{n-1}\}, \quad S_k = \sum_{j=1}^{n-k} c_j c_{j+k}, \\ k = 1, \dots, n-1, \end{aligned} \quad (3)$$

where \mathbf{c} is a binary vector subject to the constraint, $S_0 = \sum_{j=1}^n c_j = n_1$.

GAs are robust problem solving techniques based on the principles of natural evolution and selection.⁶ GAs are population-based algorithms, in which a set of potential solutions to the problem are evolved by application of several operators. To apply a GA to a given optimization problem requires an objective function and a set of genetic operators that will be applied to the population of solutions to improve it.

The GA that we implement for finding optimum binary ZRCs starts with a randomly generated initial population of binary codes (individuals) of length n . Each binary code represents a possible solution to the problem that must be evaluated to obtain an objective function value associated with it, also known as fitness. The population of binary codes evolves through the successive application of the genetic operators, basically selection, crossover, and mutation. Selection is the process by which individuals are randomly sampled with probabilities proportional to their fitness values, which, in this case, is the value of σ . An elitist strategy, consisting of passing the highest fitness codes to the next generation, is applied to preserve the best solution encountered thus far in the evolution. The selected set, which is the same size as the initial population, is subjected to the crossover operation with a probability P_c . Figure 2 shows an example of a crossover procedure: first, the

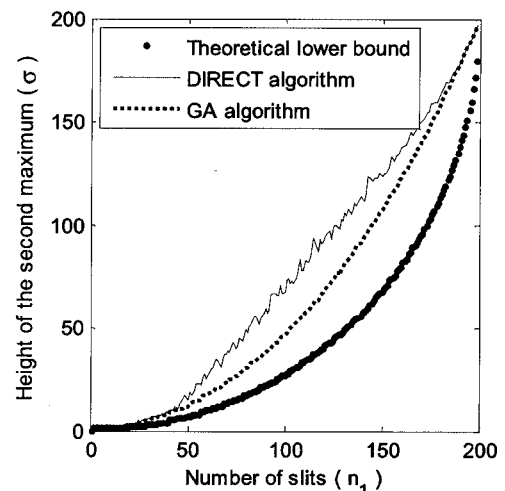


Fig. 3. Comparison of the second maximum reached with the DIRECT algorithm, the GA, and the theoretical lower bound. The code length is 200, and n_1 varies from 1 to 199.

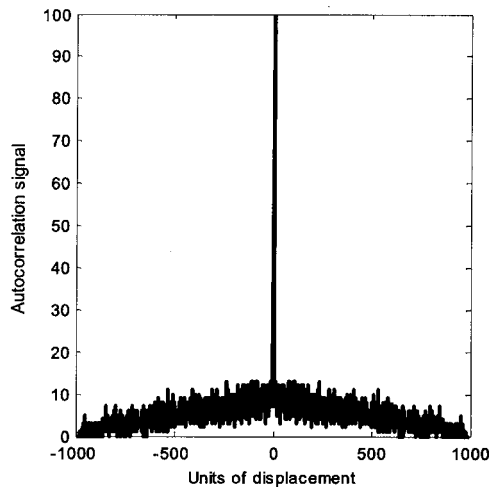


Fig. 4. Autocorrelation signal generated with an optimum code. The length of the code is 1000, and $n_1=100$. The value of the second maximum is 13.

binary codes are coupled with a predetermined probability. Second, for each pair of codes, an integer position along the code is selected uniformly at random. Two new codes are then composed by swapping all elements between the selected position and the end of the code. The last operator in the standard GA is the mutation operator. By this operation, every element in every code of the population may be changed from 1 to 0, or vice versa, with a very small probability named P_m . This operator prevents the GA from prematurely converging to suboptimal solutions. The GA evolution stops when a given stop criterion is fulfilled, usually the number of generations (number of times that the genetic operators are applied).

Note that the GA described above needs an extra operator to fix the number of ones in the binary codes. This operator is known as a restricted search operator and has been used before in the literature.⁷ After the application of the crossover and mutation operators of the GA, there will be a number of individuals that, in general, will be different from n_1 . The restricted search operator randomly adds or removes the necessary ones to keep n_1 constant. The restricted search operator reduces the size of the search space⁷ from 2^n to

$$\binom{n}{n_1}.$$

The GA was run with $P_c=0.6$, $P_m=0.01$, 500 generations, and a population of 100 individuals. A comparison between the theoretical lower bound of σ given by inequality (2), the value of the second maximum reached with the DIRECT algorithm and the GA is shown in Fig. 3. The optimizations were done with a fixed number of elements, $n=200$, and with a variable number of slits in the interval from 1 to 199. From this figure, it can be seen that in a few cases it is possible to reach the theoretical lower bound, and this takes place when the number of slits is small. For ZRCs with more than 100 elements, the DIRECT solutions are worse than GA solutions since the DIRECT solutions are suboptimal.

To demonstrate the power of the method we have calculated various cases with large codes, in particular, a ZRC with $n=1000$ elements and $n_1=100$. With these parameters the value of the second maximum is 13. Figure 4 shows the optimal autocorrelation signal with respect to the relative displacement between the codes. This problem is unmanageable with previous techniques.

This research was supported by the National Research Program project DPI2004-7334 and by a project of the Universidad de Alcalá (UAH PI2005/078). J. Saez-Landete's e-mail address is jsaez@fis.ucm.es.

References

1. X. Yang and C. Yin, *J. Phys. E* **19**, 34 (1986).
2. L. Yajun, *J. Mod. Opt.* **34**, 1571 (1987).
3. L. Yajun, *Optik (Stuttgart)* **79**, 67 (1988).
4. J. Sáez-Landete, J. Alonso, and E. Bernabeu, *Opt. Express* **13**, 195 (2005).
5. D. R. Jones, in *Encyclopedia of Optimization*, Vol. 1 (Kluwer Academic, 2001), pp. 431–440.
6. D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning* (Addison-Wesley, Reading, Mass., 1988).
7. S. Salcedo-Sanz, G. Camps-Vals, F. Pérez-Cruz, J. Sepúlveda-Sanchís, and C. Bousoño-Calzón, *IEEE Trans. Syst. Man Cybern.* **34**, 749 (2005).