Design of two-dimensional zero reference codes by means of a global optimization method

José Saez-Landete

Departamento de Teoría de la Señal y Comunicaciones – Escuela Politécnica Superior – Universidad de Alcalá. Campus Universitario – (28805) Alcalá de Henares – Madrid – Spain.
jsaez@fis.ucm.es

José Alonso, Eusebio Bernabeu

Departamento de Óptica – Facultad de Ciencias Físicas – Universidad Complutense. Ciudad Universitaria – (28040) Madrid – Spain.
j.alonso@fis.ucm.es, ebernabeu@fis.ucm.es

Abstract: A method to obtain the absolute measure of the position is by means of the autocorrelation of two zero reference marks. In one-axis measurement systems one dimensional mark are used and the design of these marks is relatively complex. When the movement is in two-axes, two dimensional reference marks are required and they are even harder to design. We report a method of global optimization to calculate the optimal two dimensional zero reference marks which generate the reference signal with the highest central peak. This method proves to be a powerful tool for solving this problem.

OCIS codes: (050.2770) Gratings; (120.0120) Instrumentation, measurement, and metrology; (120.3940) Metrology; (220.0220) Optical design and fabrication; (230.0230) Optical devices.

References and Links

1. M. C. King and D. H. Berry, "Photolithografic mask alignment using moiré techniques" Appl. Opt. **11**, 2455-2459 (1972).
2. V. T. Chitnis and Y. Uchida, "Moiré signals in reflection" Optics Communications **54**, 207-211 (1985).
3. Xiangyang Yang and Chunyong Yin, "A new method for the design of zero reference marks for grating measurement systems" J. Phys. E Sci. Instrum. **19**, 34-7 (1986).
4. Li Yajun, "Autocorrelation function of a bar code system" J. Mod. Opt. **34**, 1571-5 (1987).
5. Li Yajun, "Optical valve using bar codes" Optik **79**, 67-74 (1988).
6. J. Sáez-Landete, J. Alonso, E. Bernabeu, "Design of zero reference codes by means of a global optimization method" Op. Ex. **13**, 195-201 (2005), http://www.opticsexpress.org/abstract.cfm?URI=OPEX-13-1-195.
7. Y. Chen, W. Huang and X. Dang, "Design and analysis of two-dimensional zero-reference marks for alignment systems" Review of Scientific Instruments, **74**, 3549-53 (2003).
8. D. R. Jones, C. D. Perttunen, and B. E. Stuckman. "Lipschitzian Optimization without the Lipschitz Constant" J. Optim. Theory Appl. **79**, 157-181 (1993).
9. Donald R. Jones. *DIRECT Global optimization algorithm*. Encyclopedia of Optimization. (Kluwer Academic Publishers, Dordrecht, 2001).
10. Bjorkman, Mattias and Holmstrom, Kenneth. "Global Optimization Using the DIRECT Algorithm in Matlab" Advanced Modeling and Optimization, **1**, 17-37 (1999).
11. Daniel E. Finkel and C. T. Kelley. "Convergence analysis of the DIRECT algorithm" Optimization Online (2004).
12. J. M. Gablonsky. *DIRECT Version 2.0 User Guide*. (CRSC Technical Report, Raleigh, 2001).

1. Introduction

The absolute measure of the position is especially important in precision engineering, nanoscience and nanotechnology. Actually, the increasing demand for high resolution in

#7276 - $15.00 US Received 26 April 2005; revised 19 May 2005; accepted 20 May 2005

lithography and mask-alignment has created a strong incentive for design new techniques of alignment. In 1972 King and Berry, Ref. [1], were the first to use the moiré technique for optical lithographic mask alignment. The technique consists of passing a light beam through a pair of gratings. The variation of light intensity in some diffracted order depends on the lateral displacement of a grating and this variation is registered in a photodiode. The modulation of the generated signal, named "moire signal", strongly depends on the gap between the gratings. The signal modulation vanishes outside of the Talbot planes. Since then, several authors have used this technique in different forms. In Ref. [2], Chitnis *et al.* used a phase shifted pairs of gratings. In this configuration, the system improves the behavior with regards to small variations in the gap between gratings.

On the other hand, in grating measurement systems (displacement transducer), a reference signal (or zero reference signal) is an important addition to an incremental displacement measurement system for achieving absolute position, an origin of a coordinate, or a machine home position. To acquire a zero reference signal, gratings with adjacent Zero Reference Codes (ZRC) have been developed since 1986, Ref. [3]. The ZRC consist of a group of unevenly spaced transparent and opaque slits. When a ZRC moves with respect to another, the two ZRCs overlap. A parallel ray beam propagates through both codes and the total amount of transmitted light is registered by means of a photodiode. The reference signal is the output from this photodiode and it is the correlation between the two ZRCs. The characterization and design of optimum codes to obtain suitable reference signals has been studied in Ref. [4-5]. In these works there is not a systematic method of calculus from which arbitrary length ZRCs could be obtained. In Ref. [6], we have presented an algorithm for the design of ZRC which generate optimum reference signal for arbitrary length of the codes by means of a direct search algorithm.

In Ref. [7], Huang *et al.* proposed a new kind of 2-dimensional binary codes for precise positioning and alignment. These 2D codes are made up of unevenly located, opaque pixels on a transparent substrate. The system operation is similar to the one used with one-dimensional zero reference signals. When the movement takes place on the XY plane, the signal is obtained as the two-dimensional correlation of the ZRCs. Therefore, two-axis alignment can be detected with a simple system. Despite the simplicity of the optical alignment system, high quality two-dimensional ZRCs are much harder to design than its one-dimensional counterparts. In this work we demonstrate that a global optimization algorithm can also be used to quickly obtain two-dimensional ZRCs with optimal correlation properties.

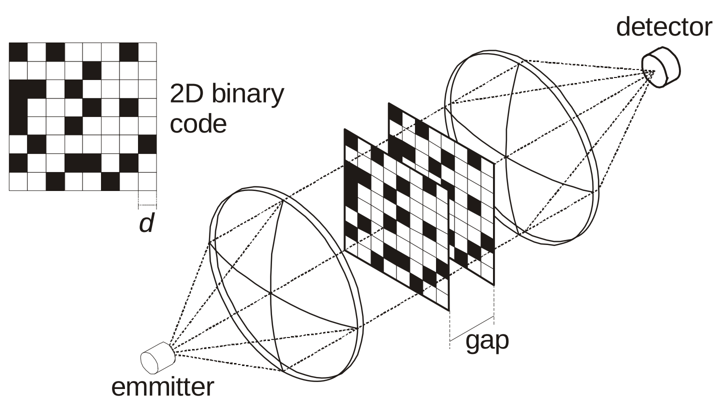

Fig. 1. Two-dimensional alignment system based on two-dimensional ZRCs

2. General considerations

Two dimensional ZRCs consist of a set of specially coded elements that can be implemented as two-dimensional binary codes. In Fig. 1 we show the alignment system. The ZRCs are parallel to each other and at least one of them is set in an XY mobile stage. A collimated beam

#7276 - $15.00 US Received 26 April 2005; revised 19 May 2005; accepted 20 May 2005

passes through them in the perpendicular direction and the total transmitted flux is detected in a photodiode. The output signal depends on the relative displacements along the X and Y directions. In order to increase the maximum of this signal, the two codes are made identical so that the reference signal becomes the autocorrelation of the same repeated ZRC.

In general, the structure of the 2D ZRC can be represented by the following matrix of binary data

$$\mathbf{c} = \left[c_{ij} \right] = \begin{bmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{n1} & \cdots & c_{nn} \end{bmatrix}, \; c_{ij} \in \{0,1\} \,, \tag{1}$$

where n^2 is the total number of elements of the ZRC, c_{ij}=1 if a transparent pixel is located at the *ij*-position, and c_{ij} = 0 elsewhere. The number of transparent pixels is n_1. The sizes of the transparent and opaque regions in the ZRC are integer multiples of the width of a single pixel.

We will 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 small with regard to the size of the pixels in the code and this size is greater than wavelength of the illuminating light.

When the two ZRC have relative displacements of k and l units in the X and Y directions respectively, the signal registered in the photodiode is proportional to

$$S_{kl} = \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} c_{i+k,\, j+l} \,, \tag{2}$$

where $k, l = -n+1, ..., n-1$, and the signal S_{kl} is the autocorrelation matrix of the two ZRC defined in Eq. (1). S_{00} is the signal obtained when the relative displacement between the ZRCs is zero. It is the central maximum and is equal to the number of transparent pixels, n_1

$$S_{00} = \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij}^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} = n_1 \,. \tag{3}$$

The secondary maximum of the signal is

$$\sigma = \max_{k^2+l^2 \neq 0} \left[S_{kl} \right] \tag{4}$$

where $k^2 + l^2 \neq 0$ mean that $k \neq 0$ and $l \neq 0$ at the same time.

The most important parameter that characterize a zero reference signal is the ratio between the secondary and the main maximum, $K = \sigma / S_{00}$. A good zero reference signal must be a single and well distinct peak, so the secondary maxima of the correlation signal must be low. The smaller K value, the higher the sensitivity and robustness of the zero reference signal.

In absence of diffraction, the size of the pixels of the ZRC defines the width of the central peak of the reference signal and this width is the resolution of the alignment system. The diameter of the light beam limits the number of pixels in the ZRC and in turn, the sensitivity of the photodiode determines the minimum value for the central maximum of the signal, that is, the number of transparent pixels of the ZRC. According with these working requirements, we have n and n_1 predetermined and we have to minimize the second maximum of the signal, σ.

In the following section we calculate a new lower bound for the second maximum of the signal. In sections 4 and 5 we propose a computation method to obtain the optimum two dimensional ZRC with the lowest second maximum.

3. The lower bound of the second maximum

In one dimension displacement, a lower bound of the second maximum has been obtained by Yajun in Ref. [5]. In two dimensional displacements, Chen *et al.* Ref. [7] obtained a lower bound for two dimensional ZRC's, assuming that all n_1 transparent pixels are concentrated in one corner of the ZRC. This lower bound is very conservative, and it is given by:

#7276 - $15.00 US Received 26 April 2005; revised 19 May 2005; accepted 20 May 2005

$$\sigma \geq \frac{n_1(n_1+3)(n_1-1)}{4(n^2-1)}. \tag{5}$$

We will obtain a new value for this bound by means of a generalization of the one dimensional bound from Yajun, Ref. [5]. The two dimensional autocorrelation satisfy the following properties:

i. When a code moves with respect to the other a full run along the x and y directions, the integrated light intensity registered in the photodiode is:

$$\sum_{k=-n+1}^{n-1}\sum_{l=-n+1}^{n-1} S_{kl} = \sum_{k=-n+1}^{n-1}\sum_{l=-n+1}^{n-1}\sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}c_{i+k,j+l} = \sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}\sum_{k=-n+1}^{n-1}\sum_{l=-n+1}^{n-1} c_{i+k,j+l} = n_1^2 \tag{6}$$

ii. The autocorrelation of a ZRC is symmetric, $S_{k,l} = S_{-k,-l}$.

iii. The value of S_{kl} is equal to the number of coincident transparent pixels when a code is displaced k and l units in the x and y directions respectively. In this case, a transparent pixel (c_{ij} = 1) coincide with other transparent pixel ($c_{i+k,,j+l}$ = 1) if both elements are in the code, that is, if there is a substructure with dimension (k+1, l+1) in the ZRC,

$$\underbrace{\begin{bmatrix} X & \dots & 1 \\ \vdots & \dots & \vdots \\ 1 & \dots & X \end{bmatrix}}_{l+1} \left.\vphantom{\begin{bmatrix} X \\ \vdots \\ 1 \end{bmatrix}}\right\} k+1 \tag{7}$$

where $X \in \{0,1\}$. Moreover, the value of S_{kl} is the number of these type of substructures in the ZRC with dimension $(|k|+1, |l|+1)$. Being the code dimension $n\times n$, the maximum number of substructures contained in the code are $(n-|k|)(n-|l|)$, so

$$S_{kl} \leq (n-|k|)(n-|l|). \tag{8}$$

For large values of $|k|$ and $|l|$ (approaching n), the number of substructures contained in the ZRC is small and this upper bound is also small, $S_{kl} \leq 0$ for $k=n$ and $l=n$. On the other hand, when the displacements $|k|$ and $|l|$ are small (approaching 0), the upper bound is large and very rough, $S_{kl} \leq n^2$ for $k=0$ and $l=0$.

iv. Considering the first maximum S_{00} in Eq. (3) and the second maximum σ in Eq. (4), we can establish another upper bound for the signal,

$$\underset{k^2+l^2\neq 0}{S_{kl}} \leq \sigma. \tag{9}$$

This upper bound is a constant and obviously, $\sigma = n_1 \leq n^2$.

For k and l close to n, the lowest bound is given by Eq. (8), and for k and l close to 0, Eq. (9) establishes the lowest bound. A limit between two regions can be established when two bounds are equal.

$$|l_1| = n - \frac{\sigma}{n-|k_1|}. \tag{10}$$

Then, when $|l| \leq |l_1|$ the lowest bound is the Eq. (9), whereas if $|l| > |l_1|$ it is the Eq. (8). For some values of k_1 and σ, the Eq. (10) can be negative and we only have the bound given in Eq. (8). These values are for

$$|k| > n - \frac{\sigma}{n}. \tag{11}$$

#7276 - $15.00 US Received 26 April 2005; revised 19 May 2005; accepted 20 May 2005

Finally, if $|k| \le n - \sigma/n$, S_{kl} has two upper bounds,

$$S_{kl} \le \begin{cases} \sigma & 0 \le |l| \le n - \dfrac{\sigma}{n-|k|} \\ (n-|k|)(n-|l|) & n - \dfrac{\sigma}{n-|k|} \le |l| < n \end{cases} \tag{12}$$

On the other hand, if $|k| > n - \sigma/n$ we only consider an upper bound,

$$S_{kl} = (n-|k|)(n-|l|) \qquad 0 \le |l| \le n-1 \,. \tag{13}$$

In order to estimate a lower bound for σ, we replace the signal S_{kl} by the upper bound in the Eq. (6). For this purpose, by means of property (ii), we split the Eq. (6) in three terms:

$$\sum_{k=-n+1}^{n-1} \sum_{l=-n+1}^{n-1} S_{kl} = S_{00} + 2\sum_{l=1}^{n-1} S_{0l} + 2\sum_{k=1}^{n-1} \sum_{l=-n+1}^{n-1} S_{kl} \tag{14}$$

In order to calculate the first sum at the right of Eq. (14), we split it using Eq. (12),

$$\sum_{l=1}^{n-1} S_{0l} = \sum_{l=1}^{n-\frac{\sigma}{n}} \sigma + \sum_{l=n-\frac{\sigma}{n}+1}^{n-1} n(n-l) = \frac{1}{2}\left(-\frac{\sigma^2}{n} + \sigma(2n-1) \right). \tag{15}$$

We also split the second sum at the right of Eq. (14) according to the values of k and l, and using the inequalities given in Eq. (8) and Eq. (12),

$$\begin{aligned} \sum_{k=1}^{n-1} \sum_{l=-n+1}^{n-1} S_{kl} &= \sum_{k=1}^{n-\frac{\sigma}{n}} \sum_{l=-n+1}^{n-1} S_{kl} + \sum_{k=n-\frac{\sigma}{n}+1}^{n-1} \sum_{l=-n+1}^{n-1} S_{kl} \le \sum_{k=1}^{n-\frac{\sigma}{n}} \sum_{l=-n+1}^{-\left(n-\frac{\sigma}{n-k}\right)} (n-k)(n+l) + \sum_{k=1}^{n-\frac{\sigma}{n}} \sum_{l=-\left(n-\frac{\sigma}{n-k}\right)+1}^{n-\frac{\sigma}{n-k}} \sigma + \\ &+ \sum_{k=1}^{n-\frac{\sigma}{n}} \sum_{l=n-\frac{\sigma}{n-k}+1}^{n-1} (n-k)(n-l) + \sum_{k=n-\frac{\sigma}{n}+1}^{n-1} \sum_{l=-n+1}^{0} (n-k)(n+l) + \sum_{k=n-\frac{\sigma}{n}+1}^{n-1} \sum_{l=1}^{n-1} (n-k)(n-l) = \\ &= \frac{1}{2}\left(-3\sigma^2 + \sigma n(4n-1) - 2\sigma \sum_{k=1}^{n-\frac{\sigma}{n}} \frac{\sigma}{n-k} \right) \end{aligned} \tag{16}$$

The last sum is bounded for $\sum_{k=1}^{n-\sigma/n} \frac{\sigma}{n-k} \ge n\left(n - \frac{\sigma}{n}\right)$. Substituting the Eq. (16) and Eq. (15) into Eq. (14) and using Eq. (3) and Eq. (6) we obtain:

$$0 \ge n_1(n_1 - 1) + \sigma(2n^2 + n - 1) - \sigma^2\left(1 + \frac{1}{n}\right). \tag{17}$$

The second maximum of the signal must fulfill the inequality

$$\sigma_1 \le \sigma \,, \tag{18}$$

where σ_1, the lower bound for the second maximum of the signal, it is one of the solutions of the second order equation, Eq. (17). In particular,

$$\sigma_1 = \frac{-(2n^2+n-1) + \sqrt{(2n^2+n-1)^2 + 4\left(1+\frac{1}{n}\right)n_1(n_1-1)}}{-2\left(1+\frac{1}{n}\right)} \,. \tag{19}$$

Although there are some simple cases in which this bound is reached, we have no evidence that for any values of n and n_1, at least one ZRC could be found for which the equality sign holds.

Now, the objective is the calculation the ZRCs whose autocorrelation has the minimum second maximum, for this, we will use the DIRECT algorithm.

#7276 - $15.00 US Received 26 April 2005; revised 19 May 2005; accepted 20 May 2005

4. Direct algorithm

DIRECT is an algorithm developed by Donald R. Jones *et al.* Ref. [8] for finding the global minimum of a multivariate function subject to simple bounds, using no derivative information. DIRECT is a sampling algorithm, that is, the algorithm samples points in the domain of the function, and uses the information it has obtained to decide where to search next. The name DIRECT comes from "DIviding RECTangles", which describes the way the algorithm moves towards the minimum. The first step in the algorithm is to transform the search space in a unit hypercube. The function is then sampled at the center of this hypercube. The hypercube is then divided into smaller hyperrectangles whose center-points are also sampled. Normally, the Lipschitz constant of the objective function is utilized for determining the next rectangle to sample. As the Lipschitz is not known, DIRECT identifies a set of "potentially optimal" rectangles corresponding to all the combinations of Lipschitz constants and rectangles sizes in each iteration. All "potentially optimal" rectangles are further divided into smaller rectangles whose centers are sampled. This method does not have a natural way of defining convergence and DIRECT typically terminates when a user-supplied limit of function evaluations is reached. Unless the global minimum value is known, there is no way to assure that it is reached. In practice, if the optimal value does not change when increasing the number of evaluations we can consider it a global minimum.

5. Description of the modeling and results

We will apply DIRECT algorithm to the design of an optimum ZRC. For this application, the objective function is,

$$\min_{\mathbf{c}\in binary} f\left(\mathbf{c}\right), \quad f(\mathbf{c}) = \max_{k^2+l^2\neq 0}\left\{S_{kl}\right\}, \quad S_{kl} = \sum_{i=1}^{n-k}\sum_{j=1}^{n-l} c_{ij}c_{i+k,\,j+l} \tag{20}$$

where $f(\mathbf{c}) = \sigma$ is the second maximum of the autocorrelation signal, $\mathbf{c}$ is a binary matrix and the constraint is

$$\sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij} = n_1\,, \tag{21}$$

the constraint is the maximum of the signal which is equal to the number of transparent pixels in the ZRC, Eq. (3).

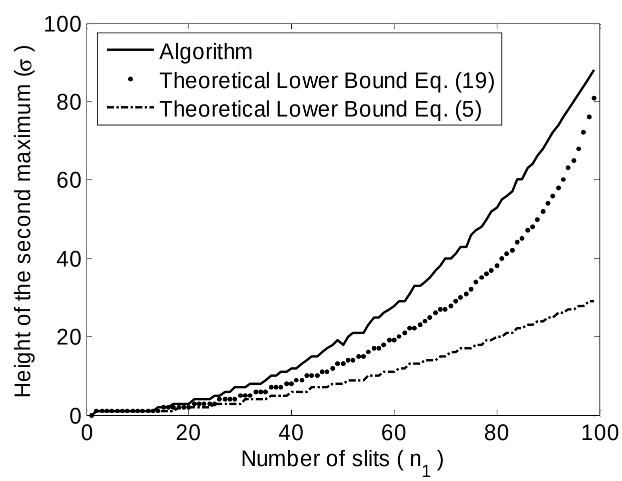

Fig. 2. Height of the second maximum of the autocorrelation with n=10. The continuous graph is the reached with DIRECT, the dotted one is a lower bound calculated theoretically in Eq. (19) and the dash-dot one is the bound showed in Eq. (5).

#7276 - $15.00 US
Received 26 April 2005; revised 19 May 2005; accepted 20 May 2005

A comparison between the theoretical lower bound of σ given by the Eq. (19), the conservative lower bound given by Chen *et al*. (Eq. (5)), and the value of the second maximum reached with DIRECT is shown in Fig. 2. The optimizations were done with 100 elements, $n = 10$, and a variable number of slits in the interval from 1 to 99. From this figure, it can be seen that in few cases it is possible to reach the theoretical lower bound and this takes place when the number of transparent pixels is small. In Fig. 3, we show the optimum autocorrelation signal for a ZRC with 10x10 elements and 50 slits.

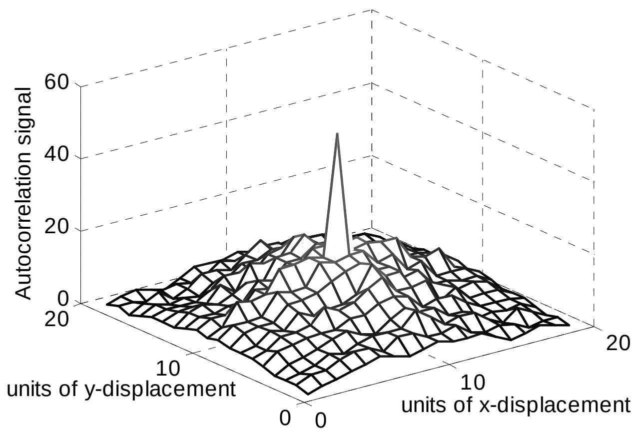

Fig. 3. Optimum reference signal for n=10 and n_1=50.

Another advantage of the DIRECT method is that the algorithm finds some equivalent solutions with the same second maximum. Among these solutions, it is possible the selection of the ZRCs by taking into consideration other secondary (optical) criteria: sensibility to diffraction, sensibility to non-uniformity of the light beams, Ref. [6].

5. Conclusions

The design of optimum 2D zero reference marks is a very demanding task in term of computing, due to the quadratic growth of the number of pixels as a function of the code dimension. At the present, there is not a systematic method to calculate zero reference marks which generate the optimum reference signal. We have obtained a new theoretical lower bound for the amplitude of secondary peak in the autocorrelation signal, which is a measure of the quality of the ZRC. In one dimension, the previous techniques may tackle the design of one dimensional ZRC's up to 25 elements. In two dimensional codes this capacity establishes a limit for the dimension of the ZRC to just 5x5 elements. The application of the DIRECT algorithm allows expanding the design capability for greater values of 10x10 elements. We have demonstrated that the technique is also valid for two dimensional ZRC, which are necessary for two-dimensional alignment. We demonstrate that the algorithm obtains many two dimensional ZRC with the same optimal value of the objective function. This allows for a selection in order to obtain codes with other optimal properties.

Acknowledgments

This research was supported by the national research program, Project No. DPI2004-7334

Received 26 April 2005; revised 19 May 2005; accepted 20 May 2005