INCREASED SPEED OF WAVEFRONT ADAPTIVE CONTROL OF LIGHT BEAMS USING THE MULTIDITHER ALGORITHM. PART II. PRACTICAL REALIZATION OF THE ALGORITHM

V.A. Trofimov

M.V. Lomonosov State University, Moscow
Received December 3, 1989

Ways of achieving practical realization of the optimal principles of variation of the control constants obtained in Part I of this article are examined. The possibility of organizing control of the beam wavefront without the use of gradient methods is discussed. Principles for varying the control constants in multichannel adaptive systems are proposed.

1. Let us consider practical ways of realizing the optimal principles of varying the control constants suggested in Part I of this article as well as some control algorithms distinct from the gradient method. It is easy to see that in estimating the quality of compensation of nonlinear distortions by means of functionals of the peak intensity (or equivalent functionals, e.g.. functionals of the power) the control effected by the rule

$$\Theta_{N+1} = \Theta_N + \frac{\gamma}{J^2 L^2} \frac{\partial J}{\partial \Theta_N}. \quad (1)$$

for $\gamma = \gamma_0/2$, is equivalent to adjusting the focusing along the x and y axes with optimal variation of the control constant. Consequently, Eq. (1) represents a practical way of realizing beam focus control having the maximum speed. In general, in the control of the wavefront S of a light beam, instead of Eq. (1), we shall obtain

$$S_{N+1} = S_N + \frac{\gamma}{J^2 L^2} \frac{\partial J}{\partial S_N}. \quad (2)$$

This algorithm substantially increases the speed of the adaptive system during beam focusing in an immobile medium, or in the course of a preliminary correction of the beam shift in a mobile medium during wavefront control based on the algorithms belonging to the class of gradient methods. The numerical experiments performed in Ref, 1 both for optimization of the focusing of the beam and for optimization of the beam wavefront confirm the conclusion reached here.

If it is difficult to obtain data about the location of the receiver. then it becomes advantageous to use an algorithm of the form

$$S_{N+1} = S_N + \frac{\gamma}{J} \frac{\partial J}{\partial S_N}. \quad (3)$$

Its convergence, stability, and speed are determined by the initial beam power and do not depend on the distance to the receiver. Thus, in the course of tuning according to the criterion of peak intensity for the monotonic regime of adjustment of the optimal parameters, It is sufficient that the condition $2\gamma/(1 + \alpha) < 1$ be fulfilled.

Another important advantage of control based on formula (3) is the small probability of the appearance of stochastic regimes of operation of the adaptive system.

The construction of an algorithm analogous to either algorithm (1) or (2) and the realization of the optimal principles of varying $\gamma^{(x)}_{N+1}$ (Eqs. (8′) and (8″) in Part I of this article) in the correction of the beam center shift in a thick layer of a nonlinear medium is a complicated problem which will be solved in the near future. Here, we shall consider some possible approaches to the construction of such an algorithm. Thus, the convergence of the following iteration process for Θ^x_N in correcting the beam center shift

$$\Theta^{(x)}_{N+1} = \Theta^{(x)}_N + \frac{\tilde{\gamma}^{(x)}_{N+1}}{JL^2} \frac{\partial J}{\partial \Theta^{(x)}_N} \quad (4)$$

In the case in which the other control parameters are fixed (focusing, etc.) does not depend explicitly on the distance to the receiver, nor on the location of the beam center, which is intrinsic to algorithm (7) in Part I of the present article. For algorithm (4), the exponential dependence of the optimal values of $\gamma^{(x)}_{N+1}$ on the beam center location are realized automatically, which significantly improves the adaptation conditions, in particular, the choice of the optimum value of $\tilde{\gamma}^{(x)}_{N+1}$ for algorithm (4), which depends only on f^2_N. That is why the advantage of algorithm (4) in comparison with the algorithm in conventional use is obvious. However, in correcting the beam shift according to algorithm (4) it is necessary that the focusing remain constant while controlling the tilt. If data on the maximum radiation intensity at the receiver

0235-6880/90/01 062–4 $02.00

(J_m) are available, then controlling $\Theta_N^{(x)}$ according to the rule

$$\Theta_{N+1}^{(x)} = \Theta_N^{(x)} + \frac{\gamma_0^{(x)}}{JJ_m L^2} \frac{\partial J}{\partial \Theta_N^{(x)}} \quad (5)$$

and controlling $\Theta_N^{(x)}$ according to rule (7) in Part I of the present article with an optimal choice $\gamma_{N+1}^{(x)}$ (Eqs. (8′) and (8″) in Part I of the present article) are equivalent (this cam easily be shown if one takes into account that $J_m \sim 1/f_N^2$). Consequently, algorithm (5) is optimal from the viewpoint of the maximum speed of reaching the optimal value of $\Theta_{opt}^{(x)}$ in the class of gradient methods during the correction of the beam center shift.

From a comparison of Eqs. (1)–(5), important practical conclusions follow: first, it is advantageous to compensate the beam center shift and the focusing of the beam onto the receiver separately (see also Ref. 2); second, to realize the maximum speed of attainment of the optimal conditions of concentration of the light energy on the receiver, it is necessary to control the focusing and the wavefront slope of the beam according to different algorithms (e.g., according to algorithms (2) and (5)).

It should also be noted that if the optimal principles of variation of the control constants are not realized, then the iterative process of achieving the optimal distribution of S converges (we assume that the conditions necessary for this are fulfilled) geometrically with denominator q equal to $1 - c\gamma$, where c is a constant, i.e., it converges linearly in q. It is impossible to achieve a faster rate of convergence in the class of gradient methods, since, in essence, they are simple iteration methods for obtaining the roots of the equation

$$\partial J/\partial S = 0. \quad (6)$$

To accelerate the attainment of the optimal values of beam focusing, wavefront tilt, and other parameters (e.g., to converge the corresponding iteration processes with denominator q^2), it is necessary to resort to algorithms that do not belong to the class of gradient methods, e.g., to Newton's method for the solution of nonlinear equations (Refs. 3 and 4), which has a quadratic convergence near the root of the respective equation. Application of Newton's method presupposes, in this case, an organization of control according to the rule

$$S_{N+1} = S_N + \gamma \frac{\partial J/\partial S_N}{|\partial^2 J/\partial S_N^2|}. \quad (7)$$

Note that the presence of the modulus around the second derivative in Eq. (7) is fundamental owing to the specific characteristics of the problem. Analysis shows (see Ref. 5) that algorithm (7) possesses a number of advantages over gradient methods.

2. Let us consider the problem of increasing the speed of multichannel adaptive systems. As is well known, the elastic mirror drives are conventionally arranged in such a way that some of them take part in beam focusing, others remove astigmatism, still others remove coma, etc. This approach is justified for beam focusing in a linear medium, or in examining contributions to the efficiency of compensation of various aberrations. In the case of a strong nonlinear response, it may be more advantageous (from the viewpoint of the number of drives), when the same drive takes part in the compensation of several types of aberration. In this case, the problem arises of their optimal distribution on the mirror, on which the organization of control in the adaptive system and, ultimately, its speed depend. Thus, for the case of a weak mutual influence of the drives. It is possible to organize their parallel control. However, in a sufficiently bad overlapping of the action of separate drives, the quality of control decreases since the desired wavefront of the light beam is not formed. If the action of separate drives overlap strongly, the mirror represents a system with strong coupling between the control channels, and the breakdown of convergence of the optimization iteration process in one channel unavoidably affects the convergence of the process in the other channels. A comparison of different ways of arranging the elastic mirror drives from the viewpoint of their mutual overlapping was carried out in Ref. 6.

It should be emphasized that it is desirable to minimize the number of drives on the elastic mirror in order to increase the speed of the system. Therefore, it may be the case that none of the different ways of arranging them that have been suggested, e.g., in Ref. 6, can be realized owing to an unsufflcient number of drives for their uniform arrangement over the mirror surface. In this situation it is necessary to take into account the amplitude distribution of the beam and to arrange the drives more densely in the region of the maximum intensity, and less densely in the beam periphery. Then, the most intense portion of the beam will be focused better than for a uniform distribution of the drives on the mirror, which may result in a higher concentration of light energy on the receiver, specifically, in its maximum intensity. In my opinion, arrangement of the drives, taking into account the beam profile, will make it possible to considerably reduce their overall number without any significant reduction in the be sun focusing quality.

In the use of elastic mirrors, it is also advantageous to introduce damping. This is effectively achieved by introducing constraints on the deviation of the shape of the mirror, e.g., from a planar profile.[7] It is important that a regularization of the adaptation process occurs In this case as a result of which the functional being minimized becomes convex, while its dependence on the parameters being optimized will be single-valued. Some increase in the speed of the adaptive system (by approximately a factor of two) also occurs.[7]

Another aspect of controlling elastic (or segmented) mirrors is the following. As numerical experiments have shown, the speed of attainment of the optimal drive perturbations for a constant value (e.g., the optimal one from the viewpoint of convergence of the control algorithm) of the control constants (let us denote it by γ_{opt}) over each channel depends on the location of the drive with respect to the beam center: those closer to it attain their optimal value earlier. Therefore, to increase the speed of a multichannel adaptive system, it is advantageous to perform the control over each channel with different constants $\gamma_p = \gamma_{opt}\tilde{\gamma}_p$, where p is the channel number. Let us obtain the dependence of $\tilde{\gamma}_p$ on the initial beam profile $A_0(x, y)$ and on the response function $\Phi_p(x, y)$ of the drives. Towards this end, let us consider the iteration process of optimization of the perturbation Θ_p of the drives.

As is known, in elastic mirrors their required is created by means of perturbations applied to some points on the mirror:

$$S(x, y) = \sum_{p=1}^{M} \Theta_p \Phi_p(x, y), \tag{8}$$

where M is the total number of drives. We assume that with increase in the drive number the distance from it to the beam center does not decrease. Optimization of Θ_p according to the gradient method with the purpose of maximizing the chosen criterion J is performed according to the rule

$$(\Theta_p)_{N+1} = (\Theta_p)_N - (\gamma_p)_{N+1} \frac{\partial J}{\partial (\Theta_p)_N}, \quad p = 1 - M. \tag{9}$$

Making use of the standard way of calculating the functional derivative (see, e.g., Ref. 8) by solving the equation adjoint to the quasi-optical equation in the complex beam amplitude $A(L, x, y)$ it can easily be shown that

$$\frac{\partial J}{\partial (\Theta_p)_N} = 2IM \int\int \Phi_p(x, y) \left\{ A_0(x, y)\, e^{iS(x,y)}\, \psi^*(0, x, y) \right\} \times dx\, dy, \tag{10}$$

where IM means that the imaginary part of the integral is taken, the asterisk on the function ψ denotes the complex conjugate, ψ is the solution of the adjoint equation, and S is defined by Eq. (8).

From Eq. (10) it follows that with increase in the drive number the value of the integral and, consequently, that of the functional derivative varies like

$$\eta_p = \int\int \Phi_p(x, y)\, A_0(x, y)\, dx\, dy, \tag{11}$$

and decreases for hyper-Gaussian beams, as the distance from the beam center to the next drive increases resulting in a slowing down of the adaptation process. Therefore, in order to equalize the increments over each channel, it is necessary to take account in Eq. (10) of the value of n_p:

$$\gamma_p = \gamma_{opt}/\eta_p = \gamma_{opt}\tilde{\gamma}_p. \tag{12}$$

The control constants in the p-channel should be multiplied by $\gamma_p = 1/n_p$. In this case, the speeds of attainment of the optimal drive perturbations are equalized.

3. Conclusions. It has thus been shown that the variation of the control constants over different channels should be performed according to their own principles. Several ways of modernizing the gradient method, which make it possible to eliminate many of the difficulties associated with the optimization of the beam wavefront according to the given algorithm, in particular, to eliminate the dependence of its convergence on the distance to the target, the initial beam power, and so on, have been considered.

REFERENCES

1. A.P. Sukhorukov and V.A. Trofimov, Kvant. Elektron. **12**, No. 8. 1617 (1985).
2. I.N. Kozhevnikova, A.P. Sukhorukov and V.A. Trofimov, Izv. Vyssh. Uchebn. Zaved. Ser. Fiz., No. 2, 13 (1985).
3. A.A. Samarskiĭ, *Introduction to Numerical* Methods (Nauka, Moscow, 1987).
4. A.A. Samarskiĭ, *A Theory of Differential Schemes* (Nauka, Moscow, 1983).
5. A.P. Sukhorukov and V.A. Trofimov, Kvant. Elektron. **14**, No. 11. 133 (1987).
6. A.P. Sukhorukov and V.A. Trofimov, Izv. Akad. Nauk SSSR, Ser. Fiz. **52**, No. 2, 377 (1988).
7. Yu.N. Karamzin, A.P. Sukhorukov and V.A. Trofimov, Kvant. Elektron. **11**, No. 4, 693 (1984).
8. F.P. Vasil'ev *Numerical Methods of Solution of Extremal Problems* (Nauka, Moscow, 1980).