MODELING AND PERFORMANCE EVALUATION OF OPTICAL SYSTEMS USING SKEW RAY TRACING

Te-Tan Liao[1] and Jing-Fung Lin[2]
[1]Department of Mechanical Engineering, Far East University, Tainan, Taiwan, R. O. C.
[2]Department of Computer Application Engineering, Far East University, Tainan, Taiwan, R. O. C.
Contact: ttliao@cc.feu.edu.tw

Received November 2005, Accepted May 2007
No. 05-CSME-61, E.I.C. Accession 2913

ABSTRACT
This paper applies a computational geometric algebra approach based on a 4 x 4 homogeneous transformation matrix to model optical systems and to evaluate their performance. In the proposed approach, the directions of the refracted/reflected rays at each boundary in the optical system are determined using skew ray tracing based upon Snell's law. The differential changes in the image coordinates caused by optical aberrations are derived for both polychromatic and monochromatic light by applying a sensitivity analysis approach. Finally, a merit function is constructed comprising five individual defect items in order to evaluate the overall performance of a generic optical system. The proposed analytical approach provides a comprehensive and robust approach for the modeling and evaluation of optical systems.

MODELAGE ET ÉVALUATION DE LA PERFORMANCE DE SYSTÈMES OPTIQUES AU MOYEN D'UN LANCER DE RAYON D'OBLIQUITÉ

RÉSUMÉ
Cet article applique une approche algébrique, géométrique et informatique basée sur une matrice de transformation homogène 4 x 4 pour modeler des systèmes optiques et évaluer leur performance. Dans le cadre de l'approche proposée, les directions des rayons réfractés et réfléchis à chaque périmètre du système optique sont déterminées au moyen d'un lancer de rayon d'obliquité basé sur la loi de Snell. Les changements différentiels aux coordonnées de l'image entraînés par des aberrations optiques sont dérivés pour les lumières polychromatique et monochromatique en appliquant une approche d'analyse de sensibilité. Finalement, une fonction de mérite est établie, laquelle est constituée de cinq points de défaut individuels pour évaluer le rendement général d'un système optique générique. L'approche analytique proposée fournit une option exhaustive et robuste pour le modelage et l'évaluation de systèmes optiques.

1. INTRODUCTION

Practical optical systems form a perfect image only in the paraxial region. However, in most finite aperture optical systems, the field of view extends far beyond the limits of the paraxial region. Since current mathematical techniques do not provide a convenient means of determining the ray path or the parameter sensitivity of light, optical devices and systems are generally designed and evaluated by observing the actual behavior of light as it travels through the optical system in question. In practice, the light path can be determined by applying a ray tracing technique, in which the optical laws of reflection or refraction are applied at each boundary encountered by the light ray [1, 2]. To expedite the design and integration of the various components in an optical system, Lin and Liao [3] reformulated the fundamental optical laws governing skew ray tracing and sensitivity analysis in terms of revolution geometry. Various examples of the state-of-the-art optical designs for mechanical engineering applications are reported in [4,5,6]. All these design procedures are based on the principles of geometrical optics, which assume that light travels along a straight path in a homogeneous medium [7,8]. In general, the light rays in an optical system can be classified as either axial rays, meridional rays, or skew rays. Axial rays and meridional rays can be traced using relatively simple trigonometric formulas, or even graphically if a low precision is acceptable. Meridional rays in the paraxial region of an optical system can also be crudely traced via the successive application of matrix production [7]. A skew ray, which is the most general ray path, is much more difficult to trace. Nevertheless, skew ray tracing is essential when modeling an optical system and evaluating its performance. Consequently, when designing optical systems, it is conventional to trace rays using actual light sources and the optical components in question.

Liao and Lin [9] performed a sensitivity analysis to determine the aberrations of monochromatic and polychromatic light as they passed through an optical element with a flat boundary surface. Lin and Sung [10] developed a novel general matrix method for paraxial skew ray tracing in systems with non-coplanar optical axes containing spherical and flat boundary surfaces. In their approach, a first-order Taylor series expansion was used to approximate the skew ray tracing equations in a simple repetitive linear matrix form. Image orientation in optical systems is difficult in traditional ray tracing. Although many modern computer optical programs are capable of tracing rays in numerical form, such programs have limited capability for modeling systems in algebraic form. Tsai and Lin [11] overcame the difficulty of the traditional trial-and-error method by using a merit function to determine the change in orientation of an image as it was reflected / refracted by a series of flat boundary surfaces. As part of the same study, the authors also proposed a prism design with a minimum number of flat reflective surfaces.

Current commercial software applications generally use some form of skew ray tracing approach to evaluate the performance of optical systems based on the conventional Seidel aberrations, the spot size, and so on. However, the skew ray tracing methods used in these applications are computationally intensive. Accordingly, the current study applies a simple geometric algebra approach based on a 4 x 4 homogeneous transformation matrix and skew ray tracing to develop a simple merit function with which to analyze the performance of generic optical systems. Section 2 of this paper applies a skew ray tracing approach based on Snell's law to trace the paths of reflected / refracted rays at each optical boundary within an optical system.

In optical systems, perfect images are formed only in the paraxial region. However, in practice, optical systems have a finite aperture and the field of view extends far beyond the limits of the paraxial region. As a result, both the position and the size of the optical image are subject to errors; conventionally referred to as optical aberrations. Five distinct classes of optical aberration have been identified for monochromatic light, including field curvature and distortion aberrations, which are image-shaping errors resulting in a discernible difference between the shape of the object and that of the image, and spherical, coma and astigmatism aberrations, which are point-imaging aberrations arising when the rays entering the image space are not concurrent. In principle, skew ray tracing provides the means to evaluate the effect of these various aberrations on the overall performance of an optical system. However, in practice, the application of skew ray tracing to assess the cumulative effect of aberrations at each of the individual optical boundaries within the optical system is a highly complex task requiring a huge computational effort [12]. Optical aberrations are essentially complex nonlinear functions of the constructional parameters of the optical system. Section 3 accounts for the effect of aberrations on the performance of an optical system by developing a systematic set of equations which relate changes in the positions and directions of the refracted / reflected rays to changes in the light source position and the direction of the incident ray, respectively. Furthermore, the concept of a sensitivity matrix is introduced to account for the effects of aberrations at an optical boundary. It is shown that the overall sensitivity matrix of the optical system can be formulated as the product of the sensitivity matrices of each optical boundary in the system.

The application of this sensitivity analysis approach to the modeling of the aberrations of monochromatic light is discussed in Section 4. In practice, however, light sources tend to be polychromatic rather than monochromatic. Importantly, the refractive index of any optical medium varies with the wavelength of the incident light. As a result, when polychromatic light is refracted, every individual monochromatic light contained within it has its own unique path and imaging position. Consequently, a chromatic aberration effect is introduced. Welford [13] discussed the difficulties involved in developing formulae to predict these chromatic aberrations and reported that the resulting analytical expressions are highly complex and cumbersome. Accordingly, Section 5 of this paper extends the sensitivity analysis approach presented in Section 3 to quantify the effects of chromatic aberrations by taking into account the effect of changes in the directions of the reflected and refracted rays caused by wavelength-dependent variations in the refractive index.

Software-based optical design programs typically use a single number, known as the merit function, to indicate the overall quality of an optical system. Essentially, the merit function is given by the sum of the squares of many individual image defects. The precise form of the merit function varies from one program to another. For example, some programs trace a large number of rays from a single object point and use a merit function based on the spot sizes of the resulting images to optimize the design parameters [14,15], while others use a merit function based on the classical optical aberration formulae. However, regardless of the particular method adopted, the merit functions used by these programs represent the summation of a number of suitably weighted defect items and provide an overall indication of the performance of the optical system. Basically, the smaller the value of the merit function, the better the optical performance of the system. To minimize the value of this multi-dimensional merit function, all of the gradients of the function with respect to the system's variables (referred to as the constructional parameters) must be set to zero. However, existing merit functions do not generally offer mathematical expressions of their gradients. Accordingly, Section 6 of this paper defines an analytical merit function comprising five discrete defect items (each of them, a merit function in its own right) to enable the rigorous evaluation of the performance of a generic optical system. The validity of the proposed merit function is demonstrated using an f/2.8 Petzval lens system comprising four lenses.

In the derivations performed in this paper, the position vector $P_{ix}i+P_{iy}j+P_{iz}k$ is written in the form of the column matrix ${}^{j}P_{i}=[P_{ix} \quad P_{iy} \quad P_{iz} \quad 1]^{T}$, where the pre-superscript "j" of the leading symbol ${}^{j}P_{i}$ indicates that the vector is referred with respect to coordinate frame $(xyz)_{j}$. Furthermore, given a point ${}^{j}P_{i}$, its transformation, ${}^{k}P_{i}$, is represented by the matrix product ${}^{k}P_{i}={}^{k}A_{j}\,{}^{j}P_{i}$, where ${}^{k}A_{j}$ is a 4×4 matrix defining the position and orientation (referred to hereafter as the configuration) of frame $(xyz)_{j}$ with respect to another frame $(xyz)_{k}$ [16]. The same notation rules are also applied to the unit directional vector ${}^{j}\ell_{i}=[\ell_{ix} \quad \ell_{iy} \quad \ell_{iz} \quad 0]^{T}$. Note that for vectors referred to the world frame, $(xyz)_0$, the pre-superscript "0" is omitted for reasons of convenience.

2. SKEW RAY TRACING AT OPTICAL BOUNDARY SURFACES

A fundamental feature of many optical elements is that their boundaries are surfaces of revolution. Therefore, when performing a geometrical analysis of the optical performance of a system, it is first necessary to define the various boundary surfaces within the system in terms of their respective revolution geometries. A ray tracing technique can then be used to determine the paths followed by the skew rays as they undergo successive reflection and refraction operations at the various optical surfaces they encounter as they travel through the system. Figure 1 illustrates the skew ray paths at a typical boundary surface in an optical system. The boundary surface ${}^i r_i$ can be defined by rotating its generating curve in the $x_i z_i$ plane, i.e. ${}^i l_i = [R_i C\beta_i \quad 0 \quad R_i S\beta_i \quad 1]^T$, about the optical z_i axis, i.e.

$$ {}^i r_i = Rot(z_i,\alpha_i)\,{}^i l_i = \begin{bmatrix} C\alpha_i & -S\alpha_i & 0 & 0 \\ S\alpha_i & C\alpha_i & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R_i C\beta_i \\ 0 \\ R_i S\beta_i \\ 1 \end{bmatrix} = [R_i C\beta_i C\alpha_i \quad R_i C\beta_i S\alpha_i \quad R_i S\beta_i \quad 1]^T , \tag{1} $$

where $Rot(z_i,\alpha_i)$ is the rotation transformation matrix about the z_i axis, and C and S indicate cosine and sine functions, respectively.

The unit normal ${}^i n_i$ to this boundary surface is given by ${}^i n_i = s_i \left(\frac{\partial\, {}^i r_i}{\partial \beta_i} \times \frac{\partial\, {}^i r_i}{\partial \alpha_i}\right) \Big/ \left|\frac{\partial\, {}^i r_i}{\partial \beta_i} \times \frac{\partial\, {}^i r_i}{\partial \alpha_i}\right|$, where the value of s_i is specified as either +1 or −1 such that the cosine of the incident angle has a positive value, i.e. $C\theta_i > 0$. In other words, the unit normal at each boundary surface ${}^i r_i$ can be expressed as

$$ {}^i n_i = -s_i [C\beta_i C\alpha_i \quad C\beta_i S\alpha_i \quad S\beta_i \quad 0]^T . \tag{2} $$

Note that ${}^i r_i$ and ${}^i n_i$ are both expressed with respect to the boundary coordinate frame $(xyz)_i$. The configuration of the world frame $(xyz)_0$ with respect to the boundary coordinate frame is given by

$$ {}^i A_0 = A_{i0} = \begin{bmatrix} I_{ix} & J_{ix} & K_{ix} & t_{ix} \\ I_{iu} & J_{iy} & K_{iy} & t_{iy} \\ I_{iz} & J_{iz} & K_{iz} & t_{iz} \\ 0 & 0 & 0 & 1 \end{bmatrix} . \tag{3} $$

Where vectors $[I_{ix} \quad I_{iy} \quad I_{iz} \quad 0]^T$, $[J_{ix} \quad J_{iy} \quad J_{iz} \quad 0]^T$ and $[K_{ix} \quad K_{iy} \quad K_{iz} \quad 0]^T$ describe the orientation of the three unit vectors of frame $(xyz)_0$ with respect to frame $(xyz)_i$. Vector $[t_{ix} \quad t_{iy} \quad t_{iz} \quad 1]^T$ is the position vector of the origin of frame $(xyz)_0$ with respect to frame $(xyz)_i$. The unit normal with respect to the world frame, i.e. n_i, can be obtained as

$$n_i = \begin{bmatrix} n_{ix} & n_{iy} & n_{iz} & 0 \end{bmatrix}^T = A_{0i}\,{}^i n_i = A_{i0}^{-1}\,{}^i n_i$$
$$= -s_i \begin{bmatrix} I_{ix} C\beta_i C\alpha_i + I_{iy} C\beta_i S\alpha_i + I_{iz} S\beta_i \\ J_{ix} C\beta_i C\alpha_i + J_{iy} C\beta_i S\alpha_i + J_{iz} S\beta_i \\ K_{ix} C\beta_i C\alpha_i + K_{iy} C\beta_i S\alpha_i + K_{iz} S\beta_i \\ 0 \end{bmatrix} \quad (4)$$

Figure 1 shows the general case where a light ray originating at point $P_{i-1} = [P_{i-1x} \;\; P_{i-1y} \;\; P_{i-1z} \;\; 1]^T$ and directed along a unit directional vector $\ell_{i-1} = [\ell_{i-1x} \;\; \ell_{i-1y} \;\; \ell_{i-1z} \;\; 0]^T$ is reflected / refracted at an optical medium boundary surface, ${}^i r_i$. The incident point P_i, refracted ray ℓ_i, and reflected ray $\underline{\ell}_i$ are given by the following formulae [17]:

$$P_i = [P_{i-1x} + \ell_{i-1x}\lambda_i \;\;\; P_{i-1y} + \ell_{i-1y}\lambda_i \;\;\; P_{i-1z} + \ell_{i-1z}\lambda_i \;\;\; 1]^T \;, \quad (5)$$

$$\ell_i = \begin{bmatrix} \ell_{ix} \\ \ell_{iy} \\ \ell_{iz} \\ 0 \end{bmatrix} = \begin{bmatrix} -n_{ix}\sqrt{1 - N_i^2 + (N_i C\theta_i)^2} + N_i(\ell_{i-1x} + n_{ix} C\theta_i) \\ -n_{iy}\sqrt{1 - N_i^2 + (N_i C\theta_i)^2} + N_i(\ell_{i-1y} + n_{iy} C\theta_i) \\ -n_{iz}\sqrt{1 - N_i^2 + (N_i C\theta_i)^2} + N_i(\ell_{i-1z} + n_{iz} C\theta_i) \\ 0 \end{bmatrix} , \quad (6)$$

$$\underline{\ell}_i = \begin{bmatrix} \underline{\ell}_{ix} \\ \underline{\ell}_{iy} \\ \underline{\ell}_{iz} \\ 0 \end{bmatrix} = \begin{bmatrix} \ell_{i-1x} + 2n_{ix} C\theta_i \\ \ell_{i-1y} + 2n_{iy} C\theta_i \\ \ell_{i-1z} + 2n_{iz} C\theta_i \\ 0 \end{bmatrix} , \quad (7)$$

where λ_i is the magnitude of vector P_iQ_i, $N_i = \xi_{medium,i-1}/\xi_{medium,i}$ is the relative refractive index of medium i-1 with respect to medium i, and θ_i is the incident angle. Following reflection / refraction at the boundary surface, the light ray proceeds along its new path with point P_i as its new point of origin and ℓ_i ($\underline{\ell}_i$) as its new unit directional vector(s).

Figure 2 provides a schematic illustration of a generic optical system comprising various spherical and flat boundary surfaces. Under the assumption that the light source located at P_0 radiates light uniformly over an imaginary hemispherical surface area, the unit directional vector $\ell_0 = [\ell_{0x} \;\; \ell_{0y} \;\; \ell_{0z} \;\; 0]^T$ of any ray originating from this light source can be expressed in terms of the coordinates of spherical angles Φ and Ψ as $\ell_0 = [S\Phi \;\; C\Phi C\Psi \;\; C\Phi S\Psi \;\; 0]^T$. As a result, any intermediate point along an arbitrary ray emitted from P_0 has the form

$$Q_1 = P_0 + \ell_0\lambda = [P_{0x} + \lambda S\Phi \;\;\; P_{0y} + \lambda C\Phi C\Psi \;\;\; P_{0z} + \lambda C\Phi S\Psi \;\;\; 1]^T . \quad (8)$$

Without loss of generality, the following analysis traces a light ray originating at $P_{i-1} = [P_{i-1x} \;\; P_{i-1y} \;\; P_{i-1z} \;\; 1]^T$ and having a unit directional vector of $\ell_{i-1} = [\ell_{i-1x} \;\; \ell_{i-1y} \;\; \ell_{i-1z} \;\; 0]^T$ $(i = 1,2,3,\ldots,n-1)$ which is refracted / reflected at an optical boundary surface ${}^i r_i$. To trace the path of the light ray as it travels through the optical system, it is first necessary to label the individual boundary surfaces within the system sequentially from 1 to n, where n is the total number of boundary surfaces in the system.

Any ray can then be traced via the application of a conventional skew ray tracing algorithm at each boundary surface $i=1, i=2, i=3, \ldots n-1$. The point of input ray path to the system's sensor plane r_n is given by

$$Q_n = [P_{n-1x} + \ell_{n-1x}\lambda \quad P_{n-1y} + \ell_{n-1y}\lambda \quad P_{n-1z} + \ell_{n-1z}\lambda \quad 1]^T \ . \tag{9}$$

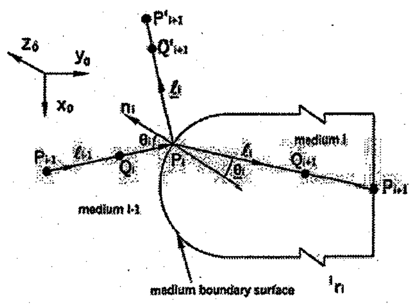

Figure 1: Skew ray tracing at medium boundary surface ${}^i r_i$

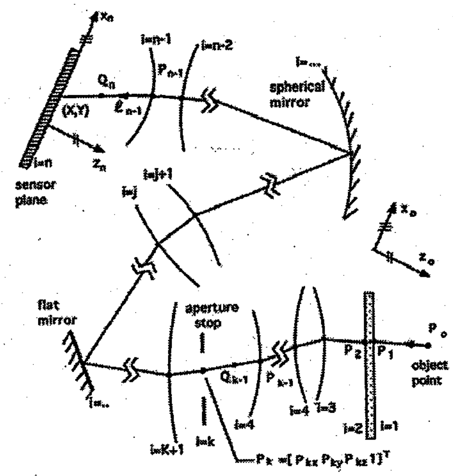

Figure 2: General notation for ray tracing through optical system with n boundary surfaces

In Eq. (9), Q_n is referred to the world frame $(xyz)_0$. Its expression with respect to the sensor frame $(xyz)_n$ is given by ${}^nQ_n = A_{n0}Q_n$, in which

$$A_{n0} = \begin{bmatrix} 1 & 0 & 0 & t_{nx} \\ 0 & 1 & 0 & t_{ny} \\ 0 & 0 & 1 & t_{nz} \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{10}$$

is the configuration of the world frame $(xyz)_0$ with respect to the sensor frame $(xyz)_n$. The parameter $\lambda = \lambda_n$ can then be obtained as

$$\lambda_n = \frac{-(P_{n-1z} + t_{nz})}{\ell_{n-1z}} \ . \tag{11}$$

The sensor readings $[{}^nP_{nx} \quad {}^nP_{ny}]^T$ of the system are given by

$$\begin{bmatrix} {}^nP_{nx} \\ {}^nP_{ny} \end{bmatrix} = \begin{bmatrix} P_{n-1x} + \ell_{n-1x}\lambda_n + t_{nx} \\ P_{n-1y} + \ell_{n-1y}\lambda_n + t_{ny} \end{bmatrix}. \tag{12}$$

Equation (12) provides the basis for determining the sensor readings $[{}^nP_{nx} \quad {}^nP_{ny}]^T$ using a ray tracing approach.

3. SENSITIVITY ANALYSIS

Sensitivity analysis provides the means to establish the respective contribution of each individual optical boundary surface along the ray path to the overall resolution of the optical system. Essentially, sensitivity analysis relates changes in the unit directional vectors and incident points of the refracted / reflected rays, i.e. $\Delta\ell_i$ and ΔP_i, to changes in the unit directional vector of the incident ray and the position of the light source, $\Delta\ell_{i-1}$ and ΔP_{i-1}.

The changes in the incident point can be obtained by differentiating Eqs. (5), (6) and (7) [16], i.e.

$$\Delta P_i = \begin{bmatrix} \Delta P_{ix} \\ \Delta P_{iy} \\ \Delta P_{iz} \end{bmatrix} = \begin{bmatrix} \Delta P_{i-1x} \\ \Delta P_{i-1y} \\ \Delta P_{i-1z} \end{bmatrix} + \lambda_i \begin{bmatrix} \Delta\ell_{i-1x} \\ \Delta\ell_{i-1x} \\ \Delta\ell_{i-1x} \end{bmatrix} + \begin{bmatrix} \ell_{i-1x} \\ \ell_{i-1y} \\ \ell_{i-1z} \end{bmatrix} \Delta\lambda_i = \left[\frac{\partial P_i}{\partial P_{i-1}}\right][\Delta P_{i-1}] + \left[\frac{\partial P_i}{\partial \ell_{i-1}}\right][\Delta\ell_{i-1}] = [M_{PPi} \quad M_{P\ell i}]\begin{bmatrix} \Delta P_{i-1} \\ \Delta\ell_{i-1} \end{bmatrix}. \tag{13}$$

Meanwhile, changes in the unit directional vector of the refracted ray are given by

$$\Delta\ell_i = \begin{bmatrix} \Delta\ell_{ix} \\ \Delta\ell_{iy} \\ \Delta\ell_{iz} \end{bmatrix} = \begin{bmatrix} -n_{ix}N_i^2C\theta_i/\sqrt{1-N_i^2+(N_iC\theta_i)^2}+N_in_{ix} \\ -n_{iy}N_i^2C\theta_i/\sqrt{1-N_i^2+(N_iC\theta_i)^2}+N_in_{iy} \\ -n_{iz}N_i^2C\theta_i/\sqrt{1-N_i^2+(N_iC\theta_i)^2}+N_in_{iz} \end{bmatrix} \Delta C\theta_i + N_i \begin{bmatrix} \Delta\ell_{i-1x} \\ \Delta\ell_{i-1y} \\ \Delta\ell_{i-1z} \end{bmatrix}$$

$$+(-\sqrt{1-N_i+(N_iC\theta_i)^2}+N_iC\theta_i)\begin{bmatrix} \Delta n_{ix} \\ \Delta n_{iy} \\ \Delta n_{iz} \end{bmatrix}$$

$$= \left[\frac{\partial \ell_i}{\partial P_{i-1}}\right][\Delta P_{i-1}] + \left[\frac{\partial \ell_i}{\partial \ell_{i-1}}\right][\Delta\ell_{i-1}] = [M_{\ell Pi} \quad M_{\ell\ell i}]\begin{bmatrix} \Delta P_{i-1} \\ \Delta\ell_{i-1} \end{bmatrix}. \tag{14}$$

Finally, changes in the unit directional vector of the reflected ray can be derived from

$$\Delta\ell_i = \begin{bmatrix} \Delta\ell_{ix} \\ \Delta\ell_{iy} \\ \Delta\ell_{iz} \end{bmatrix} = \begin{bmatrix} \Delta\ell_{i-1x} \\ \Delta\ell_{i-1y} \\ \Delta\ell_{i-1z} \end{bmatrix} + \begin{bmatrix} 2n_{ix} \\ 2n_{iy} \\ 2n_{iz} \end{bmatrix} \Delta C\theta_i + 2C\theta_i \begin{bmatrix} \Delta n_{ix} \\ \Delta n_{iy} \\ \Delta n_{iz} \end{bmatrix} = [M_{\ell Pi} \quad M_{\ell\ell i}]\begin{bmatrix} \Delta P_{i-1} \\ \Delta\ell_{i-1} \end{bmatrix}. \tag{15}$$

Combining Eqs. (13) and (14), $[\Delta P_i \quad \Delta\ell_i]^T$ can be expressed in terms of $[\Delta P_{i-1} \quad \Delta\ell_{i-1}]^T$ as

$$\begin{bmatrix} \Delta P_i \\ \Delta\ell_i \end{bmatrix} = \begin{bmatrix} \dfrac{\partial P_i}{\partial P_{i-1}} & \dfrac{\partial P_i}{\partial \ell_{i-1}} \\ \dfrac{\partial \ell_i}{\partial P_{i-1}} & \dfrac{\partial \ell_i}{\partial \ell_{i-1}} \end{bmatrix} \begin{bmatrix} \Delta P_{i-1} \\ \Delta\ell_{i-1} \end{bmatrix} = \begin{bmatrix} M_{PPi} & M_{P\ell i} \\ M_{\ell Pi} & M_{\ell\ell i} \end{bmatrix} \begin{bmatrix} \Delta P_{i-1} \\ \Delta\ell_{i-1} \end{bmatrix} = M_i \begin{bmatrix} \Delta P_{i-1} \\ \Delta\ell_{i-1} \end{bmatrix}, \tag{16}$$

where M_i denotes the sensitivity matrix of boundary surface r_i [16].

The change in the light source and its unit directional vector can be obtained by differentiating

$\ell_0 = [S\Phi \quad C\Phi C\Psi \quad C\Phi S\Psi \quad 0]^T$ to give

$$\begin{bmatrix} \Delta P_0 \\ \Delta \ell_0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & C\Phi \\ 0 & 0 & 0 & -C\Phi S\Psi & -S\Phi C\Psi \\ 0 & 0 & 0 & C\Phi C\Psi & -S\Phi S\Psi \end{bmatrix} \begin{bmatrix} \Delta P_{0x} \\ \Delta P_{0y} \\ \Delta P_{0z} \\ \Delta \Psi \\ \Delta \Phi \end{bmatrix} = M_0 \begin{bmatrix} \Delta P_{0x} \\ \Delta P_{0y} \\ \Delta P_{0z} \\ \Delta \Psi \\ \Delta \Phi \end{bmatrix}. \qquad (17)$$

Furthermore, the change in the sensor readings can be obtained by differentiating Eq. (12), i.e.

$$\begin{bmatrix} \Delta^n P_{nx} \\ \Delta^n P_{ny} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \frac{-\ell_{n-1x}}{\ell_{n-1z}} & \frac{-P_{n-1z} - t_{nz}}{\ell_{n-1z}} & 0 & \frac{\ell_{n-1x}(P_{n-1z} + t_{nz})}{\ell_{n-1z}^2} \\ 0 & 1 & \frac{-\ell_{n-1y}}{\ell_{n-1z}} & 0 & \frac{-P_{n-1z} - t_{nz}}{\ell_{n-1z}} & \frac{\ell_{n-1y}(P_{n-1z} + t_{nz})}{\ell_{n-1z}^2} \end{bmatrix} \begin{bmatrix} \Delta P_{n-1x} \\ \Delta P_{n-1y} \\ \Delta P_{n-1z} \\ \Delta \ell_{n-1x} \\ \Delta \ell_{n-1y} \\ \Delta \ell_{n-1z} \end{bmatrix}$$

$$= [M_{PPn} \quad M_{P\ell n}] \begin{bmatrix} \Delta P_{n-1} \\ \Delta \ell_{n-1} \end{bmatrix} = M_n \begin{bmatrix} \Delta P_{n-1} \\ \Delta \ell_{n-1} \end{bmatrix}. \qquad (18)$$

In Eq. (18), the sensitivity matrix M_n varies with t_{nz}, which indicates that the focal plane is not the orthogonal plane, and hence an astigmatism aberration is produced. Via successive applications of Eqs. (16)~(18), it can be shown that the change in the sensor readings can be expressed as

$$\begin{bmatrix} \Delta^n P_{nx} \\ \Delta^n P_{ny} \end{bmatrix} = M_n M_{n-1} \ldots\ldots M_{i+1} M_i M_{i-1} \ldots M_1 M_0 \begin{bmatrix} \Delta P_0 \\ \Delta \Psi \\ \Delta \Phi \end{bmatrix}$$

$$= \begin{bmatrix} M_{11} & M_{12} & M_{13} & M_{14} & M_{15} \\ M_{21} & M_{22} & M_{23} & M_{24} & M_{25} \end{bmatrix} [\Delta P_{0x} \quad \Delta P_{0y} \quad \Delta P_{0z} \quad \Delta \Psi \quad \Delta \Phi]^T. \qquad (19)$$

Equation (19) indicates that the overall sensitivity matrix of the optical system is given by the product of the sensitivity matrices of each individual boundary surface, i.e. $M_i \quad (i = 0,1,2,\ldots.n)$. A typical optical system comprises various elements and sensors, each of which has its own optical properties and resolution characteristics. However, provided that the sensor resolutions are known, Eq.(19) provides a convenient analytical tool with which to investigate the effects of a range of optical aberrations.

4. ABERRATION OF MONOCHROMATIC LIGHT

As discussed in the Introduction, optical aberrations are complex nonlinear functions of the constructional parameters of the optical system. Consequently, calculating the effects of these aberrations on the performance of the optical system using conventional mathematical techniques is a difficult, time-consuming task. However, by applying a sensitivity analysis approach, various aberration functions can be established to simplify the optical design task. To apply this approach, it is first necessary to partition Eq.(19) as follows:

$$\begin{bmatrix}\Delta^n P_{nx} \\ \Delta^n P_{ny}\end{bmatrix} = \begin{bmatrix}M_{11} \\ M_{21}\end{bmatrix}\Delta P_{0x} + \begin{bmatrix}M_{12} \\ M_{22}\end{bmatrix}\Delta P_{0y} + \begin{bmatrix}M_{13} \\ M_{23}\end{bmatrix}\Delta P_{0z} + \begin{bmatrix}M_{14} \\ M_{24}\end{bmatrix}\Delta\Psi + \begin{bmatrix}M_{15} \\ M_{25}\end{bmatrix}\Delta\Phi \cdot \tag{20}$$

The first three terms on the right-hand side of Eq.(20) describe the changes in the image coordinates, $[\Delta^n P_{nx} \quad \Delta^n P_{ny}]^T$, caused by changes in the object coordinates, $[\Delta P_{0x} \quad \Delta P_{0y} \quad \Delta P_{0z}]^T$. The matrices M_{11} and M_{22} in these terms describe the lateral magnifications of the system in the longitudinal and transverse directions, respectively, and are specified during the design stage, i.e. $M_{11,ideal}$ and $M_{22,ideal}$, respectively. Meanwhile, matrices M_{12}, M_{13}, M_{21} and M_{23} denote the various cross-sensitivities in the optical system. In general, distortion aberrations are produced within the optical system if either the lateral magnification matrices deviate from their specified values, or the cross-sensitivity matrices have non-zero values. The last two terms on the right-hand side of Eq.(20) describe the changes in the image coordinates $[\Delta^n P_{nx} \quad \Delta^n P_{ny}]^T$ caused by changes in the spherical angle coordinates of the object point, i.e. $\Delta\Psi$ and $\Delta\Phi$. Non-zero values of the matrices M_{14}, M_{15}, M_{24} and M_{25} in these two terms indicate that the rays are not concurrent in the image space, and hence spherical and coma aberrations will be induced.

In general, the numerical values of matrices M_{jk} $(j = 1 \sim 2, k = 1 \sim 5)$ depend on the construction parameters of the particular optical system. However, as a general principle, the smaller their magnitudes the better the quality of the image. Accordingly, these matrices can be used to construct an objective merit function with which to drive the optimization program used to search for the most suitable construction parameters.

5. CHROMATIC ABERRATION OF POLYCHROMATIC LIGHT

The refractive index of any medium other than a vacuum varies with the wavelength of the incident light. Consequently, when polychromatic light is refracted at the medium boundary, each constituent monochromatic light ray is separated and follows its own particular path to a unique image position. Therefore, chromatic aberrations are formed within the optical system.

The changes in the unit directional vector, $\Delta\ell_i$, caused by changes in the refractive index, ΔN_i, when a polychromatic light ray is refracted at the ith boundary, can be obtained by differentiating Eq.(6) to obtain

$$\Delta\ell_i = \begin{bmatrix}\Delta\ell_{ix} \\ \Delta\ell_{iy} \\ \Delta\ell_{iz}\end{bmatrix} = \begin{bmatrix} \dfrac{n_{ix}N_i(1-C\theta_i^2)}{\sqrt{1-N_i^2+(N_iC\theta_i)^2}} + (\ell_{i-1x} + n_{ix}C\theta_i) \\ \dfrac{n_{iy}N_i(1-C\theta_i^2)}{\sqrt{1-N_i^2+(N_iC\theta_i)^2}} + (\ell_{i-1y} + n_{iy}C\theta_i) \\ \dfrac{n_{iz}N_i(1-C\theta_i^2)}{\sqrt{1-N_i^2+(N_iC\theta_i)^2}} + (\ell_{i-1z} + n_{iz}C\theta_i) \end{bmatrix} \Delta N_i = M_{Ni}\Delta N_i \tag{21}$$

Clearly, if the ith boundary is a reflective boundary, then $\Delta N_i = 0$. According to Eq. (21), the changes in the sensor readings, $[\Delta^n P_{nx} \quad \Delta^n P_{ny}]^T$, can be obtained by successively applying the sensitivity equations, i.e.

$$
\begin{aligned}
&[\Delta^n P_{nx} \quad \Delta^n P_{ny}]^T = M_n [\Delta P_{n-1x} \quad \Delta P_{n-1y} \quad \Delta P_{n-1z} \quad \Delta\ell_{n-1x} \quad \Delta\ell_{n-1y} \quad \Delta\ell_{n-1z}]^T \\
&= M_{PPn}[\Delta P_{n-1x} \quad \Delta P_{n-1y} \quad \Delta P_{n-1z}]^T + M_{P\ell n}[\Delta\ell_{n-1x} \quad \Delta\ell_{n-1y} \quad \Delta\ell_{n-1z}]^T \\
&= M_{PPn}[M_{PPn-1}\Delta P_{n-2} + M_{P\ell n-1}\Delta\ell_{n-2}] + M_{P\ell n}M_{Nn-1}\Delta N_{n-1} \\
&= M_{PPn}[M_{PPn-1}[M_{PPn-2}\Delta P_{n-3} + M_{P\ell n-2}\Delta\ell_{n-3}] + M_{P\ell n-1}M_{Nn-2}\Delta N_{n-2}] + M_{P\ell n}M_{Nn-1}\Delta N_{n-1} \\
&= M_{PPn}M_{PPn-1}[M_{PPn-2}\Delta P_{n-3} + M_{P\ell n-2}\Delta\ell_{n-3}] + M_{PPn}M_{P\ell n-1}M_{Nn-2}\Delta N_{n-2} + M_{P\ell n}M_{Nn-1}\Delta N_{n-1} \\
&= \cdots\cdots\cdots\cdots \\
&= M_{PPn}M_{PPn-1}M_{PPn-2}\cdots M_{PP3}M_{P\ell 2}M_{N1}\Delta N_1 + M_{PPn}M_{PPn-1}M_{PPn-2}\cdots M_{PP4}M_{P\ell 3}M_{N2}\Delta N_2 \\
&+ \cdots\cdots + M_{PPn}M_{PPn-1}M_{PPn-2}M_{Nn-3}\Delta N_{n-3} + M_{PPn}M_{PPn-1}M_{Nn-2}\Delta N_{n-2} + M_{P\ell n}M_{Nn-1}\Delta N_{n-1} \\
&= \begin{bmatrix} \eta_x \\ \eta_y \end{bmatrix} \qquad (22)
\end{aligned}
$$

Eq. (22) provides exact formulae with which to calculate the chromatic aberration associated with the passage of polychromatic light through an optical system. The value of each instance of ΔN_i $(i = 1 \sim n-1)$ varies as a function only of the wavelength of the incident light. It is noted that the leading matrices in Eq. (22) indicate the weight by which each instance of ΔN_i contributes to the overall constitutive equation and are the sensitivity matrix of the corresponding ΔN_i.

6. OPTICAL MERIT FUNCTION

Modern software-based approaches for optical systems analysis and design generally use a single number, referred to as the merit function, Γ, to indicate the optical quality of the system. As discussed in the Introduction, Γ usually has the form of the sum of the squares of many individual image defects. Although these image defects may be of many different types, they are typically related to the image quality. One common merit function traces a large number of rays from the same object point and then calculates the root-mean-square (rms) error of the resulting focused spot sizes. Adopting an alternative approach, Eq. (20) provides four potential defect items, namely M_{14}, M_{15}, M_{24} and M_{25}, with which to evaluate the point imaging aberrations caused by a lack of concurrency of the rays in the image space, i.e.

$$
\Gamma_1 = \sum_{i=1}^{3}(M_{14}^2 + M_{15}^2 + M_{24}^2 + M_{25}^2)\,. \qquad (23)
$$

As indicated in Eq. (23), the values of the image defects are evaluated at three different locations in the field of view.

The distortion aberration caused by a non-uniform magnification at the image plane also provides a meaningful index with which to assess the imaging quality of an optical system. Generally speaking, this distortion increases with an increasing image size, and can be quantified using the following image defects:

$$
\Gamma_2 = \sum_{i=1}^{3}[(M_{11} - M_{11,ideal})^2 + (M_{22} - M_{22,ideal})^2], \qquad (24)
$$

$$\Gamma_3 = \sum_{i=1}^{3}(M_{12}^2 + M_{13}^2 + M_{21}^2 + M_{23}^2), \tag{25}$$

where $M_{11,ideal}$ and $M_{22,ideal}$ are the desired magnification values in the x- and y- coordinate directions, respectively. In essence, Γ_2 expresses the distortion directly in terms of the deviation of the actual magnifications in these two directions from the required values. In general, it is desirable that no cross-sensitivity exists between the x- and the y- coordinate directions. In other words, any increase in the image height in the x-coordinate direction (or image depth in the z-coordinate direction) should ideally have no effect on the image height in the y-coordinate direction. Hence, the merit function Γ_3 was the mutual magnification interference to avoid serious image distortion.

As discussed in Section 5, the refractive index of an optical medium varies with the wavelength of the light incident upon it. As a result, chromatic aberrations are produced when a polychromatic light passes through an optical system. Accordingly, a further merit function can be introduced based on the variance of the image position caused by differences in the refractive index of the optical medium, i.e.

$$\Gamma_4 = \sum_{i=1}^{3}(\eta_x^2 + \eta_y^2). \tag{26}$$

In general, the quality of the focused image is highly sensitive to variations in the object-to-image distance. In practice, any deviation from the specified focus length, t_0, is liable to result in a significant degradation in the image quality. Therefore, a further merit function can be defined based on the deviation of the actual object-to-image distance from the desired value, i.e.

$$\Gamma_5 = (t - t_0)^2 \tag{27}$$

An overall merit function can then be constructed simply by aggregating the five individual merit functions given in Eqs. (23) ~ (27), i.e.

$$\Gamma = \upsilon_1\Gamma_1 + \upsilon_2\Gamma_2 + \upsilon_3\Gamma_3 + \upsilon_4\Gamma_4 + \upsilon_5\Gamma_5, \tag{28}$$

where υ_i $(\upsilon_i \geq 0, \upsilon_1 + \upsilon_2 + \upsilon_3 + \upsilon_4 + \upsilon_5 = 1, i = 1 \sim 5)$ is a weighting factor which allows the relative contribution of each merit function to be adjusted in accordance with the basic structure of the optical system.

The numerical values of the merit functions presented above depend on the constructional parameters of the particular optical system. However, as discussed previously, the smaller the value of the merit function, the better the optical performance. The f/2.8 Petzval lens system shown in Fig. 3 is considered for illustration purposes. The prescription data for this lens are summarized in Table 1 (Laikin, 1995 [8]). In the Petzval lens, the incident light is parallel to the optical axes of the four lenses and is focused at the origin point of the image coordinate system. Figure 4 illustrates the variation of the focusing quality of the Petzval lens, i.e. Γ_2, with the position of the image plane, i.e. T10 (see Fig. 3). The results show that the optimal focusing performance is obtained when the image plane is located at a distance of 49.6316 mm from lens A. Figures 5 and 6 illustrate the variations in the merit function Γ_2 with the positions of lens A and lens B, respectively, with T10 maintained at a constant value of 49.6316 mm. The results clearly reveal that the optimal focusing quality is obtained when the optical components in the Petzval lens are deployed in accordance with the original design specification.

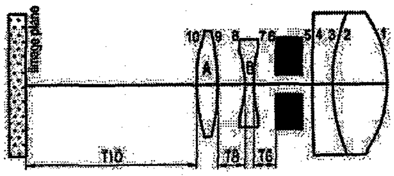

Figure 3: Petzval Lens

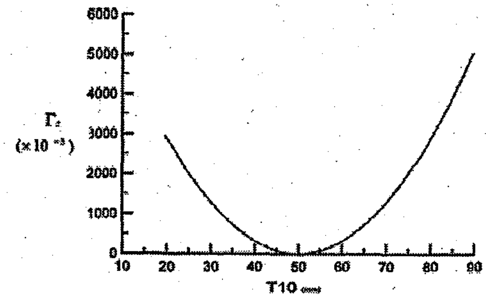

Figure4: Variation of focusing quality of Petzval lens as function of image plane position

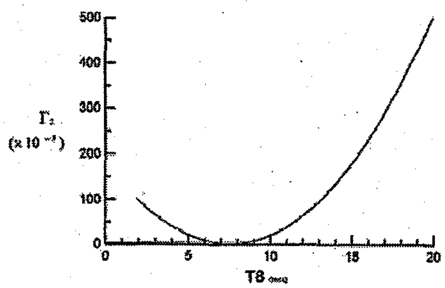

Figure 5: Variation of focusing quality of Petzval lens as function of lens A position

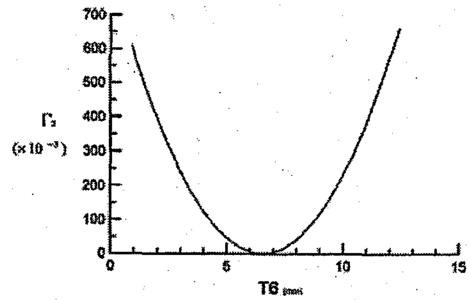

Figure 6: Variation of focusing quality of Petzval lens as function of lens B position

Table 1: The prescription data of Petzval lens

data / boundary i	radius (mm)	thickness (Ti) (mm)	refraction index ξ_i
1	38.2219	15.8496	1.61521
2	56.0857	0	1
3	56.0857	5.969	1.72311
4	590.682	3.0226	1
5	0	7.62	1
6	0	6.4008	1
7	41.7957	2.5146	1.52583
8	29.3446	7.9248	1
9	63.5635	6.096	1.61521
10	56.8655	49.6316	1

7. CONCLUSION

This paper has developed an analytical method based upon the use of a 4 x 4 homogeneous transformation matrix and skew ray tracing to model and evaluate the performance of optical systems. In the proposed approach, the directions of the reflected/refracted rays at each boundary surface in the optical system are determined in accordance with Snell's law. The effects of aberrations on the optical performance of the system are taken into account for both monochromatic and polychromatic light via the application of suitable sensitivity matrices at each optical boundary. Finally, a simple analytical merit function based on five defect items, each of which is a merit function in its own right, has been proposed and verified using an f/2.8 Petzval four-lens system for illustration purposes. The systematic analytical approach presented in this study provides a comprehensive and robust means of modeling optical systems and evaluating their performance.

ACKNOWLEDGMENTS

The author gratefully acknowledges the financial support provided to this study by the National Science Council of Taiwan under Grant No. NSC93-2212-E-269-014.

REFERENCES

1. Conrady, A. E., 1929, Applied Optics and Optical Design, Part I, Oxford Press, reprinted by Dover Publications, Inc., New York, 1957.

2. Herzberger, M., Modern Geometrical Optics, Interscience Publishers, New York, 1958.

3. Lin Psang Dain and Te-Tan Liao, "Skew-Ray Tracing and Sensitivity Analysis of Geometrical Optics," Transactions of the ASME Journal of Manufacturing Science and Engineering, Vol. 122, 2000, pp. 338-349.

4. K. H. Kim, K. F. Eman and S. M. Wu, "Analysis alignment errors in a laser-based in-process cylindricity measurement system," Trans. of the ASME J. of Eng. for Industry, Vol. 109, 1987, pp. 321-329.

5. J. Ni, and S. M. Wu, "An on-line measurement technique for machine volumetrics error compensation," Trans. of the ASME J. of Eng. for Industry, Vol. 115, 1993, pp. 85-92.

6. K. Lau, R. J. Hocken and W. C. Haight, "Automatic laser tracking interferometer system for robot metrology," Precision Engineering, Vol. 8, No. 1, 1986, pp. 3-8.

7. E. Hecht, Optics, 3rd Edition, Addison-Wesley Publishing Company, 1998.

8. M. Laikin, Lens Design, Second Edition, New York, Marcel Dekker, Inc., 1995.

9. Te-Tan Liao and Psang Dain Lin, "Analysis of Optical Elements with Flat Boundary Surfaces," Applied Optics. Vol. 42, No. 7, 2003, pp. 1191-1202.

10. Psang Dain Lin and Chi-Kuen Sung, "Matrix-based Paraxial Skew Ray-tracing in 3D Systems with Non-coplanar Optical Axis," Optik, Vol. 117, 2006, pp. 329-340.

11. Chuang-Yu Tsai and Psang Dain Lin, "Prism Design Based on Changes in Image Orientation," Applied Optics, Vol. 45, No. 17, 2006, pp. 3951-3959.

12. M. Born and E. Wolf, Principles of Optics, Pergamon, New York, 1985.

13. W. T. Welford, Aberrations of the Symmetrical Optical System, Academic Press, New York, 1974.

14. R. Kingslake, Lens Design Fundamentals, Academic, New York, 1977.

15. C. J. Progler and D. M. Byrne, "Merit Functions for Lithographic Lens Design," J. Vac. Sci. Technol. B, Vol. 14, No. 6, Nov/Dec, 1996, pp.3714-3718.

16. R. P. Paul, Robot Manipulators-Mathematics, Programming and Control, MIT press, Cambridge, Mass., 1982.

17. Psang Dain Lin and Te-Tan Liao, "A New Model of Binocular Stereo Coordinate Measurement System Based on Skew Ray Tracing," Transactions of the ASME Journal of Dynamic System, Measurement, and Control, Vol. 126, March, 2004, pp.102-114.

NOMENCLATURE

$(xyz)_0$	World coordinate frame.
$(xyz)_i$	Coordinate frame imbedded in ith boundary surface.
r_i	The ith boundary surface with unit normal n_i.
P_i	Incident point on ith boundary surface.
Q_{i+1}	Arbitrary intermediate point on light ray following refraction at ith boundary surface.
ℓ_i	Unit directional vector of light ray following refraction at ith boundary surface.
s_i	$s_i = 1$ or $s_i = -1$ such that cosine of incident angle $C\theta_i > 0$
A_{kj}	Configuration of frame $(xyz)_j$ with respect to frame $(xyz)_k$.
θ_i, $\underline{\theta}_i$	Incident angle, refraction angle.
N_i	$N_i = \xi_{medium,i-1}/\xi_{medium,i}$, where $\xi_{medium,i}$ is relative refractive index of medium i with respect to vacuum.
ΔP_i	Differential change in incident point P_i.
$\Delta\ell_i$	Differential change in unit directional vector ℓ_i of refracted ray.
Φ, Ψ	Spherical angles of arbitrary ray emitted from light source at P_0
M_{Ni}	Sensitivity matrix defined as $[\Delta P_i \quad \Delta\ell_i]^T = M_{Ni}\Delta N_i$ when light ray encounters ith boundary surface.
M_i	Sensitivity matrix defined as $[\Delta P_i \quad \Delta\ell_i]^T = M_i[\Delta P_{i-1} \quad \Delta\ell_{i-1}]^T$ when light ray encounters ith boundary surface.
Γ_i	The ith merit function
υ_i	Weighting factor of Γ_i