

Physically-Accurate Synthetic Images for Machine Vision Design

J. M. Parker

Dept. of Mechanical Engineering,
University of Kentucky,
Lexington, KY 40506-0108

Kok-Meng Lee

George W. Woodruff School
of Mechanical Engineering,
Georgia Institute of Technology,
Atlanta, Georgia 30332-0405

In machine vision applications, accuracy of the image far outweighs image appearance. This paper presents physically-accurate image synthesis as a flexible, practical tool for examining a large number of hardware/software configuration combinations for a wide range of parts. Synthetic images can efficiently be used to study the effects of vision system design parameters on image accuracy, providing insight into the accuracy and efficiency of image-processing algorithms in determining part location and orientation for specific applications, as well as reducing the number of hardware prototype configurations to be built and evaluated.

We present results illustrating that physically accurate, rather than photo-realistic, synthesis methods are necessary to sufficiently simulate captured image gray-scale values. The usefulness of physically-accurate synthetic images in evaluating the effect of conditions in the manufacturing environment on captured images is also investigated. The prevalent factors investigated in this study are the effects of illumination, the sensor non-linearity and the finite-size pinhole on the captured image of retroreflective vision sensing and, therefore, on camera calibration was shown; if not fully understood, these effects can introduce apparent error in calibration results. While synthetic images cannot fully compensate for the real environment, they can be efficiently used to study the effects of ambient lighting and other important parameters, such as true part and environment reflectance, on image accuracy. We conclude with an evaluation of results and recommendations for improving the accuracy of the synthesis methodology.

1 Introduction

For machine vision system applications such as part presentation, the accuracy of image gray-scale pixel values far outweighs image appearance (Lee, 1991); in this paper, we present physically-accurate image synthesis as a rational basis for designing both hardware and software components of a vision system. This is a very complex task, since such systems consist of many parts and the most proficient systems are designed by considering the integrated hardware and software arrangement. Numerical simulation is a flexible, practical tool for investigating a large number of hardware/software configuration combinations for a wide range of parts.

Prior machine vision research includes the use of *photo-realistic* synthetic images as an aid in testing model-based vision algorithms (Wu et al., 1990). However, these images were generated with the simple image synthesis algorithms available with most commercial CAD systems, which assumed idealized or nonphysical reflectance models, limited light source models and unrealistic camera optics. While the images obtained with these packages were useful in gaining some insight into algorithm performance, their usefulness was limited (Chen and Mulgaonkar, 1991). These photo-realistic images were generated based upon work developed in the area of computer graphics, where appearance of the image to the viewer is generally the primary concern. Meyer et al. (1986) modeled the generation of a physically-accurate synthetic image. In their model, the environment description includes the scene's reflective and emitting properties, in addition to geometrical information and is processed via a simulation based upon the *physics* of illumination, instead of the idealized or nonphysical reflectance and illumination models used to produce photo-realistic images.

It was, in fact, attempts to improve the realism of photo-realistic

images that resulted in the development of highly accurate methods for calculating illumination, which supported the development of *physically-accurate* synthesis methods (Goral et al., 1984; Nishita and Nakamae, 1985; Sillion et al., 1991; Kajiya, 1986; Ward et al., 1988). The use of such physically-accurate images in an iterative method for improved image understanding was proposed by Gagalowicz (1990). Some researchers, while agreeing that understanding the illumination problem is important, felt that most physically-accurate synthesis methods were too computationally expensive to be useful in vision system design (Cowan, 1991). Rushmeier et al. (1992) have developed an efficient methodology for generating physically-accurate synthetic images that predict the gray-scale values of images captured by a computer vision system. Results from this research confirm that physically-accurate image synthesis methods, rather than those methods currently available with standard CAD packages, are necessary to sufficiently simulate captured images. Highly-accurate methods of calculating illumination establish the basis for generating the array of pixel radiances representing scene illumination. Most graphics methods, however, assume an infinitesimal pinhole to find a solution for the rendering equation (Kajiya, 1986), which calculates the average radiance incident on each image pixel for a wavelength band.

The benefits of this research are briefly summarized as follows: (1) It provides a rational basis for designing the hardware and software components of a machine vision system. Specifically, the models which account for the effects of finite-sized pinhole, sensor nonlinearity, and ambient illumination have been developed. (2) The models have been experimentally validated. Additionally, the results provide an opportunity to perform an in-depth study of the factors that can significantly degrade the performance of image-processing algorithms and aid in the determination of critical design parameters. (3) A third contribution is the foundation for a CAD-tool which utilizes physically-accurate synthetic images to accurately and inexpensively predict the performance of a proposed vision system design prior to implementation or the construction of a prototype, minimizing the need to build and test a

Contributed by the Manufacturing Engineering Division for publication in the JOURNAL OF MANUFACTURING SCIENCE AND ENGINEERING. Manuscript received May 1996; revised Feb. 1999. Associate Technical Editor: C-H. Meng.

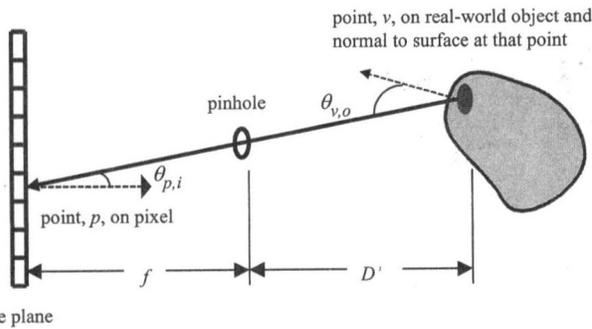


Image plane

Fig. 1 Geometry of illumination simulation for an infinitesimal pinhole camera

large number of hardware configurations. Such a tool would provide an effective means to compare algorithms and predict the optimal algorithm (and optimal performance) for a specific application earlier in the design phase, significantly reducing implementation time and improving industrial reliability.

The remainder of this paper is organized as follows: the theory of physically-accurate synthetic image generation, which includes a discussion of the governing equations necessary to generate physically-accurate images is followed by a detailed comparison of the differences between photo-realistic and physically-accurate synthesis methods. Practical implementation issues of the synthesis methods are then discussed. The Experimental Investigation section begins with a description of the hardware test-bed, followed by a discussion of the computational model for radiative transfer, which provides the foundation for physically-accurate image synthesis. This is followed by the results of the specific experiments performed to better understand the effect of parameters, especially [source and ambient] illumination; the paper concludes with a discussion of the results and recommendations for future work.

2 Fundamental Equations

The geometry of the illumination incident at each image pixel is diagrammed in two dimensions in Fig. 1. Radiance incident on the sensor pixel is equal to the radiance leaving the real-world object that is visible to the sensor through the pinhole. The solid angle of the cone rays leading to the patch on the object is equal to that corresponding patch in the image. Thus, it can be shown that the irradiance (or the power per unit area) incident on the surface of a pixel is given by

$$I_i(\theta_{p,i}, \phi_{p,i}) = I_o(\theta_{p,i}, \phi_{p,i}) \cos \theta_{p,i} d\Omega \quad (1)$$

where

- θ and ϕ = spherical coordinates (polar angle, θ , and azimuthal angle, ϕ) which specify the location and direction in three-dimensional space,
- $I_i(\theta_{p,i}, \phi_{p,i})$ = irradiance falling on a point (x_p, y_p) on the pixel along the ray from the object,
- $I_o(\theta_{p,i}, \phi_{p,i})$ = radiance (or the power per unit area per unit solid angle) from the object in the direction toward the pinhole, and
- $d\Omega$ = solid angle subtended by the corresponding patch on the real-world object as seen from the center of the pinhole.

In machine vision, captured image values are proportional to $E_{\text{pixel}}(\lambda)$, the average energy/area incident on the sensor pixel, or

$$E_{\text{pixel}}(\lambda) = \int_{\text{pixel-area}} E_i(\theta_{p,i}, \phi_{p,i}) f(x_p, y_p, \lambda) dA_{\text{pixel}} \quad (2)$$

where $E_i(\theta_{p,i}, \phi_{p,i})$ is the energy/area incident on the sensor pixel at that point along that direction and $f(x_p, y_p, \lambda)$ is the spatial sensor sensitivity for a given wavelength band at that point on the pixel. The energy/area incident on the pixel for a specified period of time τ can be solved by integrating the irradiance over the pinhole area as follows:

$$E_i(\theta_{p,i}, \phi_{p,i}) = \int_{\text{pinhole}} I_o(\theta_{p,i}, \phi_{p,i}) \cos \theta_{p,i} \tau d\omega_{\text{pinhole}} \quad (3)$$

where τ represents exposure time to scene illumination; and $d\omega$ is the infinitesimal solid angle in the direction of the pinhole as seen from the patch of object. The irradiance, $I_o(\theta_{p,i}, \phi_{p,i})$, is equal to that leaving the visible real-world object and depends on the irradiance falling on the surface from all directions. This irradiance may be from light sources or from secondary reflections from other surfaces in the environment. For reflective surfaces, the object radiance is not known a priori and must be calculated; its value depends on the light energy incident on the object as illustrated in Fig. 2. The object radiance is given by the equation of radiative transfer (Siegel and Howell, 1981):

$$I_o(\theta_{v,o}, \phi_{v,o}) = \int_{\Omega} \rho_{bd}(\theta_{v,o}, \phi_{v,o}; \theta_{v,i}, \phi_{v,i}) I_i(\theta_{v,i}, \phi_{v,i}) \cos \theta_{v,i} d\omega_i \quad (4)$$

where

- $\rho_{bd}(\theta_{v,o}, \phi_{v,o}; \theta_{v,i}, \phi_{v,i})$ = bi-directional reflectance of the object at that point, and
- $I_i(\theta_{v,i}, \phi_{v,i})$ = incident radiance on the real-world object at point (x_v, y_v) on the object, and the integral is over the entire hemisphere of directions above that object at that point.

In Eq. (4), the incident radiance, $I_i(\theta_{v,i}, \phi_{v,i})$, may originate at light sources or be due to multiple interreflections within the environment. A light source visible in direction $(\theta_{v,i}, \phi_{v,i})$ contributes an incident radiance equivalent to the emitted radiance of the source in that direction. If something other than a light source is visible in direction $(\theta_{v,i}, \phi_{v,i})$, this equation is recursive and its complete solution accounts for all possible geometric illumination effects, e.g., shadowing, highlights, etc.

The nonlinearity of the gray-scale response of an imaging sensor is specified by an exponent "gamma γ " and thus, the sensor for a computer vision system is modeled here as a power law (Lee, 1994) as follows:

$$G_{\text{pixel}} = K \left(\int_0^{\infty} E_{\text{pixel}}(\lambda) s(\lambda) d\lambda \right)^{\gamma} + G_o \quad (5)$$

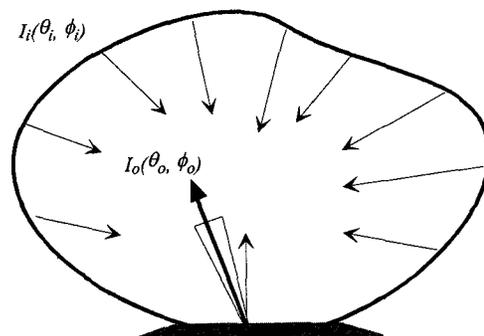


Fig. 2 Model of radiance leaving a surface

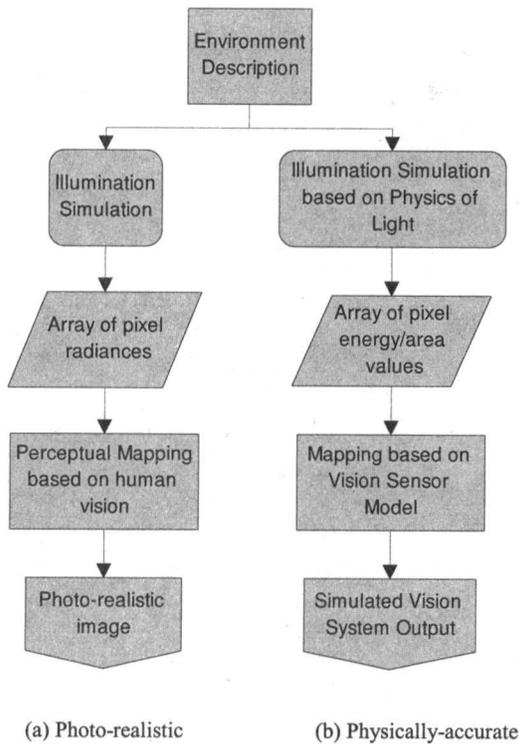


Fig. 3 Model of the synthetic imaging process

where G_{pixel} is the pixel gray-scale value; G_o is the dark-current (zero illumination) value; and $E_{\text{pixel}}(\lambda)$ is the pixel energy/area value.

Equations (2)–(5) provide an analytical model for generating the output signal for a computer vision system characterized by sensor sensitivity K and system response linearity γ . K and γ are empirically-determined constants and $s(\lambda)$ characterizes the sensor as a function of wavelength for a given sensor. A gamma of unity yields a linear response, whereas less than unity compresses the bright end and greater than unity compresses the dark end. In Eq. (5), the average energy incident over the pixel area is calculated from Eqs. (2) and (3), which require the solution to Eq. (4).

At this point, it is worthwhile to point out the primary differences between the photo-realistic and physically-accurate image generation processes, using the governing equations given above. Photo-realistic images are usually generated using limited light source (e.g., parallel or isotropic point light sources) and optics (i.e., infinitesimal pinhole) models, idealized or nonphysical (i.e., Lambertian or Phong) reflectance models, and spectral dependence is usually ignored. Physically-accurate synthesis methods, on the other hand, are able to incorporate finite aperture and lens effects [Eq. (3)], include spectral dependence [Eq. (2)] and sensor non-linearity [Eq. (5)], as well as model reflectance properties (e.g., specular, diffuse, anisotropic) and illumination distributions realistically [Eq. (4)].

3 Physically-Accurate Image Synthesis

Figure 3 compares the processes used to generate synthetic images for (a) photo-realistic and (b) physically-accurate synthetic images for vision system applications. In this context, the term “photo-realistic” refers to images that “look like they could be photographs.” Images generated for this purpose are often rendered via methods that use heuristics to make images appear realistic. They are not necessarily generated by simulating light propagation and the transfer function of the imaging device; therefore, though they appear realistic, they cannot be relied upon to yield gray-scale intensity values which are representative of the

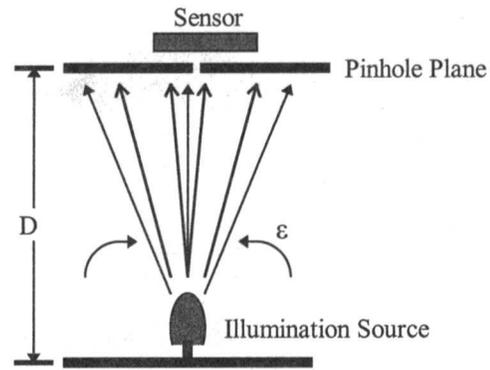


Fig. 4 Calibration experimental setup

actual scene. As shown in Fig. 3(b), physically-accurate images are based upon the physics of illumination and the behavior of the sensing device. An accurate mathematical model describes the physical scene and the vision system used to capture that scene. This model is used to simulate scene illumination, which is represented as an array of [pixel] radiances. This array of radiances is then converted to energy/area values, which are transformed by a mapping based on a model of the system sensor and how it converts incident light energy into gray-scale values.

The physically-accurate synthetic image is simulated in a two step process. In the first step, *Radiance*, a freely-distributed software package from the Lighting Systems Research Group of the Lawrence Berkeley Laboratory, is used to solve the radiative heat transfer equation, Eq. (4). In the second step, the sensor model for the computer vision system is modeled using the power law given in Eq. (5). Several implementation issues are discussed as the following subsections.

3.1 Finite-Size Pinhole Model. The specific Monte Carlo solution method used in *Radiance* has been validated by both physical measurements and comparison with other software packages written to calculate illumination (Grynberg, 1989). However, both the sensor pixels and pinhole, though small, have finite size that can not be accurately approximated by the infinitesimally small pixels and pinhole assumed by *Radiance*. This problem can be separated into two parts, and each solved separately. The first problem, that of finite pixel size, is minimized in *Radiance* by assuming an infinitesimal pinhole, shooting many rays through each pixel (i.e. sampling over the pixel area) and averaging the radiance values calculated for each ray. The software does not have an explicit mechanism for integrating over a finite aperture; therefore, the array of radiance for the finite pinhole is approximated by a weighted average of radiance calculated using a finite number of representative infinitesimal viewpoint positions within the finite pinhole. The weighting signifies the percentage of the finite pinhole represented by that infinitesimal viewpoint; in this

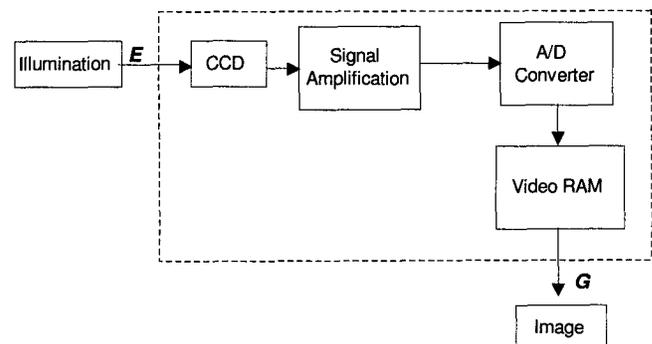


Fig. 5 Calibration of system response

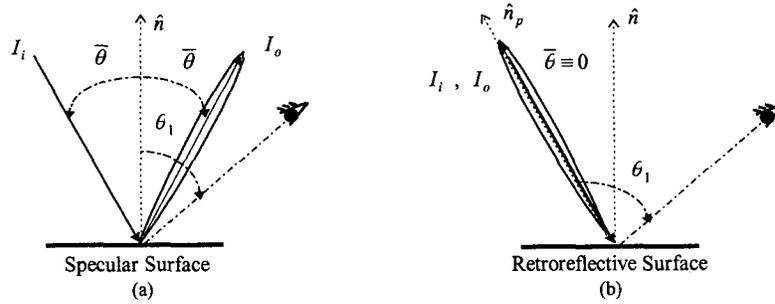


Fig. 6 Illustration of reflective surface modeling

investigation, viewpoint positions were chosen such that the representative areas and, therefore, weightings were equal. Mathematically, the weighting function is defined as

$$w_i = \frac{a_i}{A}$$

and

$$\sum_{i=1}^N w_i = 1$$

where a_i is the area represented by the (infinitesimal) viewpoint sample; A is the total pinhole area; and N is the number of (viewpoint) samples used to approximate the finite pinhole.

3.2 Sensor Model. To transform the array of irradiance determined by the computational model into predicted gray-scale values corresponding to a particular system, K and γ , are calibrated experimentally for the system which accounts for the system non-linearity. Figure 4 suggests a general technique for calibrating K and γ , where a known illumination source of a single wavelength band is sampled. Therefore, the spectral sensitivity function $s(\lambda)$ is constant in Eq. (5) and its contribution can be absorbed into the value for K . Note that the uncertainty in the value of the power output for the LED does not affect the value of γ ; however, it would "scale" K , shifting the histogram of gray-scale values left or right. For a specified source of intensity P_s (power per unit solid angle) with an included angle, ϵ , the small planar patch of area on the sensor at a distance D from the source subtends a solid angle of $\pi \tan^2(\epsilon/2)$. Over an exposure time of τ seconds, the intensity of the source on the sensor is

$$E = \tau P_s \pi \tan^2(\epsilon/2) \quad (6)$$

A series of images obtained over a wide range of exposure times ensures that a full range of experimental gray-scale output values (from the dark current value to saturation) is obtained.

Figure 5 shows a schematic of the experimental setup where the

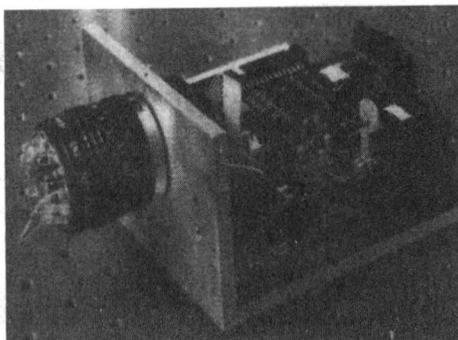


Fig. 7 FIVS with illumination system

imaging sensor, signal-conditioning amplifier, and A/D converter are treated as an integrated unit. The relationship between the average gray scale value and the average energy of the source over the exposed area on the sensor can be determined from Eq. (5), which can be rewritten in the following form to facilitate the determination of K and γ :

$$\ln(G - G_o) = \gamma \ln E + \ln K. \quad (7)$$

3.3 Reflectance Model. Perfectly specular scatters result in a reflection of 100 percent of the irradiation into an infinitesimal solid angle (i.e. a single ray) in the mirror direction. In practice, however, the specular lobe is finite in size and has a width proportional to the roughness of the scattering surface. Therefore, the radiance leaving a specular surface is modelled as a Gaussian lobe in the mirror direction as shown in Fig. 6(a). Retroreflective surfaces are modeled by using a texture map to perturb the surface normal so that it faces the direction of incident illumination. Figure 6(b) shows how the specular bi-directional reflectance definition can be modified to model a retroreflective surface by perturbing the surface normal. A detailed process of modeling the retroreflective surface for use with the software *Radiance* can be found in (Parker, 1996).

4 Experimental Investigation

While opportunities to increase image accuracy by studying and understanding the parameters which influence image accuracy exist in many areas, the most important parameters include illumination distribution and ambient lighting, part reflectance, and geometric alignment. This study concentrated on the effects of illumination distribution and ambient lighting on image accuracy. Four sets of experiments were conducted to examine the effects of design parameters on image accuracy. First, the sensor/system sensitivity and linearity constants required to establish the sensor model [in Eq. (5)] were empirically determined using the procedure described in Section 3.2; the precision of the sensor model was then examined. Secondly, the ability of generating synthetic images of retroreflective sensing was demonstrated. The captured images were used a rational basis for evaluating these synthetic images. Thirdly, the accuracy of image edges were examined experimentally. Finally, the effect of ambient lighting on image

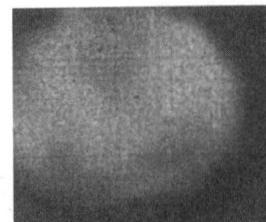


Fig. 8 Captured image of retroreflective field

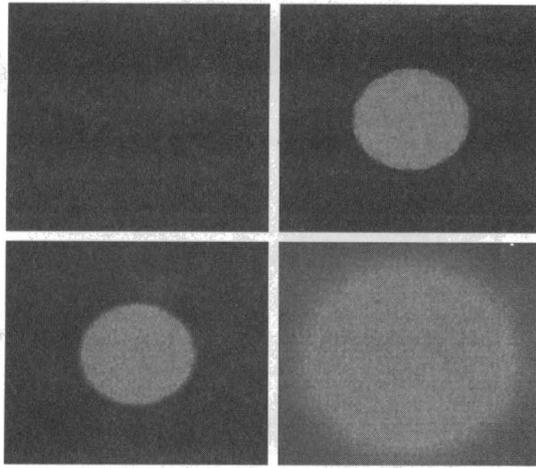


Fig. 9 Synthetic images of retroreflective field—Clockwise from upper-left-hand corner: (a) generated with a CAD-package reflectance model (b) Radiance-generated image, using infinitesimal pinhole and cone-angle illumination model (c) image with finite pinhole model and Gaussian illumination distribution (d) finite pinhole, cone-angle illumination model

accuracy was investigated using synthetic images and validated experimentally.

4.1 Determination of Vision System Parameters. The vision system chosen for system emulation and experimental validation is FIVS (Flexible Integrated Vision System, shown in Fig. 7), which was developed at Georgia Tech (Lee and Blenis, 1994). To determine the values of K and γ for the system, a single AlGaAs LED, which has a dominant wavelength of 650 nm, a typical radiant intensity of 47.1 mW/sr and a 7 deg cone angle, was chosen as the illumination source and positioned as shown in Fig. 4. Images were captured for exposure times ranging from 0 to 32 ms, yielding a dark current gray-scale value of 22 and a saturation value of 181. Using the method of least squares (as described in Section 3.2) and the typical radiant intensity and cone angle values given above to estimate the source power/area, a K value of 3.38×10^4 and a γ of 0.85 were obtained.

4.2 Synthetic Images of Retroreflective Sensing. The configuration (Fig. 7) utilizes a structured, co-located illumination system consisting of twelve AlGaAs LED's (with a 7 deg cone angle) placed circumferentially (in a 25 mm diameter) around the lens (focal length = 10.59 mm). This co-located illumination system (Lee, 1994) as shown in Fig. 7 is an effective prototype for the validation of the proposed methodology for several reasons. The use of retroreflective materials in this application enables the capture of reliable gray-scale images with a high object-to-background contrast without a detailed prior knowledge of object geometry and surface reflectance. Since an ideal retroreflector returns the incident ray in the same direction, retroreflective vision sensing also facilitates the use of low-power, low cost light sources for part presentation. The chosen CCD sensor has a strong spectral

response at the dominant wavelength provided by the illumination sources and is reasonably insensitive to the effects of ambient lighting.

Typical experimentally obtained properties of commonly used retroreflective materials are available in the following references (Federal Test Standard, 1977; Rennilson, 1980). In this investigation, 3M[®] Scotchlite 3870 retroreflective sheeting was used, the characteristic properties of which can be found in Lee (1994). To qualitatively compare various methods of generating synthetic images with retroreflective vision sensing, two experiments were conducted.

I Synthetic Image of Retroreflective Background

Synthetic images of a retroreflective background were generated and compared to a captured image of the retroreflective field (Fig. 8) obtained with the FIVS located 1 meter above a flat retroreflective sheeting. The comparison of synthetic images is given in Fig. 9.

As seen in Fig. 9(a), a CAD-generated image assuming an ideal diffuse surface results in an image that is nearly black for the illumination level and exposure time range of captured image conditions. The CAD-generated image confirms the need for an accurate reflectance model, as used in Figs. 9(b)–(d). Figure 9(b) illustrates *Radiance's* ability to model a retroreflective background. However, the illuminated area is too small and too sharply defined. Incorporation of the finite aperture [Fig. 9(d)] results in an image with a more acceptable transition between the illuminated and non-illuminated areas, but the illuminated area is still too small. The importance of accurate source emission distribution modeling is shown in Fig. 9(c).

II Synthetic Image of Retroreflective Calibration Board

A calibration board containing a 5×5 grid of retroreflective circular dots of $\frac{3}{8}$ inch (15.625 mm) diameter, which were spaced by a $\frac{1}{4}$ inch (19.05 mm) center-to-center distance, was the subject of captured and simulated images. The retroreflective calibration pattern was created by printing a partially diffuse black pattern on retroreflective sheeting. The captured and synthetic images of the calibration board are compared in Fig. 10. The comparison illustrates the limitations of infinitesimal pinhole modeling (Fig. 10(b)); averaging over the finite aperture results in an image (Fig. 10(c)) that more nearly reflects the gray-scale values of the captured image.

4.3 Prediction of Edge Transition. In machine vision, the detection of edges is often necessary. Thus, an experiment was conducted to compare synthetic and captured edge images. Representative captured and synthetic images of the transition between a black diffuse surface and a retroreflective background are shown in Fig. 11. Ten successive images of the transition between the black matte and retroreflective surfaces were captured. An analytical edge model (Lee, 1994) facilitated image analysis and provided an additional comparison. In the model, the ratio of the radius of the pinhole projection from point P on the target plane to the distance, D' , between the pinhole plane and the target plane, is derived from the geometry to be

$$\frac{R}{D'} = \frac{1}{2F} \left(1 + \frac{f}{D'} \right) \quad (8)$$

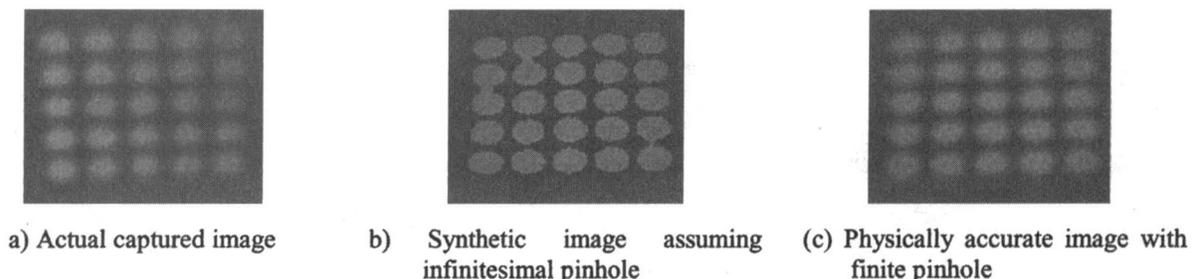


Fig. 10 Comparison of captured and synthetic images

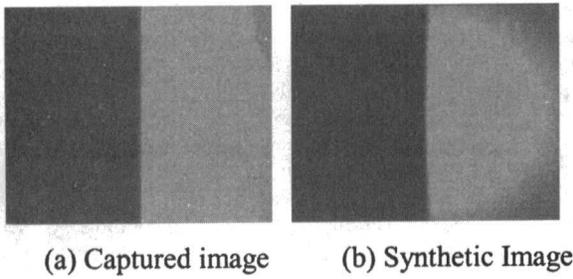


Fig. 11 Edge images

where

- R = the radius of the pinhole projection of point P on the pinhole plane onto the target plane,
- D' = the distance between the pinhole and target planes,
- $F = f/d_p$ for a pinhole of diameter d_p (as simulated in this experiment), and f is the focal length.

The intensity of the point P on the detector is proportional to the overlapping area between the pinhole projection and the bright area on the target plane. The highest and lowest intensities, I_{\max} and I_{\min} , correspond to the cases when the pinhole is completely covered by the bright and dark areas, respectively. Therefore,

$$\frac{I - I_{\min}}{I_{\max} - I_{\min}} = 1 - \frac{1}{\pi} [\cos^{-1} S - S \sqrt{(1 - S^2)}] \quad (9)$$

$$S = \frac{D'}{R} (\tan \beta - \tan \beta_o) \quad (10)$$

where $|S| \leq 1$

- I = pixel gray-scale intensity (Maximum and minimum values denotes maximum and minimum gray-scale intensities near the edge).
- S = the normalized displacement from the edge to the center of the pinhole projection, and
- β = the entrance angle (the angle between the surface normal of the target plane and the incident light).

The edge model was used to refine the “estimated true edge position” of the captured edges, and this mean position estimate used as the edge position in the synthetic images. Two major differences exist between imaging conditions for the synthetic and

captured images (Fig. 11). First, the real image was captured in ambient conditions with a lens, rather than the pinhole camera and controlled lighting conditions simulated. It was determined experimentally using Tsai’s (1987) camera calibration that the effect of the lens distortion is small. However, the addition of a lens results in higher (i.e., brighter) gray-scale values in the captured image. Second, the effect of additional ambient illumination (in Fig. 11(a)) is clearly shown in the absence of the perfect semicircular illuminated area predicted in the synthetic image (Fig. 11(b)). The vertical edge in the synthetic image was placed to match that of the captured image as nearly as possible, corresponding to a pixel position of 96.7, $\beta_o = 0.0015$ radians, $D' = 1$ m, $f = 10.59$ mm and $d_p = 0.1$ mm. An exploded view of the edge transition predicted using Eq. (5) with non-linearity values of 0.85 and 0.90 is shown in Fig. 12.

As shown in Fig. 12, the captured edge has a similar transition to the simulated edges, with a nonlinearity factor between 0.85 (the experimental value) and 0.90. The edge model or Eq. (9), which does not account for the illumination distribution, is inadequate to predict the edge transition accurately, even after it is corrected for γ . These results show the importance of the γ (sensor nonlinearity) calibration for the system and confirm that the empirical value of 0.85 is reasonable. They further illustrate the importance of an accurate illumination simulation that incorporates finite pixel and pinhole approximations in synthetic edge construction.

4.4 Effect of Ambient Illumination. While the synthetic images illustrate that correct reflectance modeling is important, the effects of ambient lighting are also critical in accurately simulating captured images for applications where accuracy is of particular concern. We explore the effect of the ambient lighting on calibrating the focal length of the FIVS using Tsai’s (1987) camera calibration algorithm. In this context, ambient lighting refers to external lighting in addition to the co-located LED illumination (Fig. 7). The external lighting consists of two 400 W metal halide lamps located at $(x = \pm 0.06$ m, $y = 0.06$ m, $z = 0.82$ m) relative to the coordinate frame of the calibration board. The emission distribution of two lamps was modelled as Gaussian (with a beam width of 7 degrees) and the RGB emission values (in $W/m^2 \cdot sr$) were chosen to reflect a power of 400 watts and to emulate the typical distribution of metal halide lamps.

The simulated orientation of the calibration board (Subsection 4.2 part II) was at an optimal position for yielding the calibrated focal length (Parker and Lee, 1997), with rotations about the x -, y -, and z -axes of 46, 6 and -2 degrees, respectively, and a focal length of 12.5 mm. The resulting synthetic image is shown in Fig.

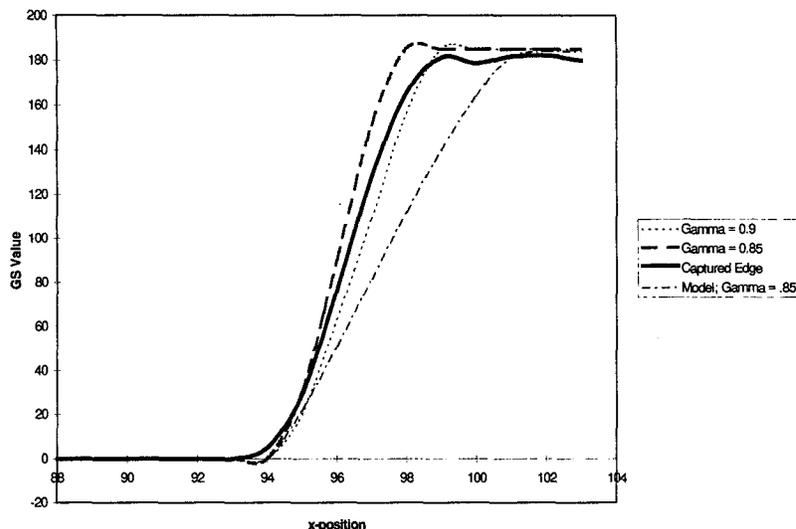


Fig. 12 Effect of gamma on synthetic edge modeling

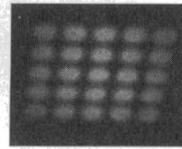


Fig. 13 Typical synthetic calibration image at a particular orientation

Table 1 Effect of variability in lighting conditions on calibrated focal length

Lighting Condition	Focal length	
	Captured Average	Synthetic Image
Controlled	12.59	12.450
Ambient	12.85	12.546

13. As shown in Table 1, the image simulated under ambient conditions yielded a longer apparent focal length.

The effect of ambient lighting on camera calibration was then experimentally validated using captured images. Ten consecutive controlled and ambient lighting images captured of the board were used to determine the effects of ambient lighting on Tsai's camera calibration. As in the simulated case, images captured under ambient conditions also yielded a longer calibrated focal length. The results suggest that illumination has a significant influence on camera calibration results. The difference in apparent focal length is not as great in the simulated case (0.8 percent versus an average increase of 2 percent in the captured images). This is believed to be due, at least in part, to the presence of less ambient illumination in the synthetic image, most noticeably, the unmodeled sources such as sunlight and its inter-reflection present in the captured scene.

5 Illustrative Design Tool

The primary motivation of this investigation is to demonstrate the use of synthetic image generation as a tool for machine vision system design through identifying the critical illumination design parameters on camera calibration.

A 2⁴ full factorial experiment was designed to generate the 16 synthetic images used to study the effects of the chosen factors on Tsai's algorithm (Table 2). The variations investigated represented reasonable permutations possible in the FIVS system. As modeled, a perfectly aligned LED ring is parallel to the *xy*-plane; therefore, an "arbitrary" misalignment was modeled as a rotation of the ring about the *y*-axis, followed by a rotation about *z*. Small variations in the source intensity and illumination distribution were modeled by perturbing nominal intensity and (Gaussian) beam width values, respectively. Similarly, variations in the direction of the source optical axis were characterized by a small nominal angle of deviation from the original optical axis of the LED and a perturbation parameter. For each factor, two sets of eight samples corresponding to the high and low level are compared.

For the analysis, the designed focal length was 12.5 mm, viewing distance was 0.79 m (31.1 in), and, exposure time was 130 ms.

Table 2 Factor and response description for illumination and calibration investigations

Factor and Response Description	Low Level	High Level
Factor A: Illumination Offset (LED ring)	$\alpha = \beta = 0^\circ$	$\alpha = 1.2^\circ, \beta = -85^\circ$
Factor B: LED Intensity [$W / m^2 \cdot sr$]	2324	2324 nominal, $a=0.15$
Factor C: LED Beam Width	7.6°	7.6° nominal, $a=0.10$
Factor D: LED Optical Axis Offset	0°	1° nominal, $a=0.90$
Response: Calibrated Focal Length ($n=1$)		

Table 3 Effect of design factors on calibrated focal length

Factor Level				Calibrated Focal Length	
A	B	C	D	Mm	(Abs.) %Error
L	L	L	L	12.450	0.40
L	L	L	H	12.152	2.78
L	L	H	L	12.373	1.02
L	L	H	H	12.528	0.22
L	H	L	L	12.588	0.70
L	H	L	H	12.575	0.60
L	H	H	L	12.558	0.46
L	H	H	H	12.605	0.84
H	L	L	L	11.037	11.70
H	L	L	H	12.763	2.10
H	L	H	L	11.150	10.80
H	L	H	H	12.371	1.03
H	H	L	L	11.612	7.10
H	H	L	H	12.367	1.06
H	H	H	L	11.624	7.01
H	H	H	H	12.164	2.69

The results showing the effects of the factors on the focal length calibration are given in Tables 3 and 4.

As shown in Tables 3 and 4, the design factors A, D, and interaction AD (LED ring orientation, LED optical axis variation and the interaction of these two effects, respectively) have a significant effect on calibrated focal length. For unreplicated factorials (i.e., the sample size at each level of the design is 1), an effect, t_i , is determined to be significant if its absolute value exceeds the random variable, $t_{\alpha/2, m/3}$, where α represents the significance level (or Type I error) and m is the number of responses (Lenth, 1989). For this investigation, α was chosen to be 0.05 and the number of observations for a 2⁴ full factorial experiment is 16; therefore, the critical value, $t_{0.025, 5}$, was 2.571. A misalignment in LED ring orientation led to the highest error in calibrated focal length and, therefore, had a very significant effect on calibrated focal length. The interaction between LED ring orientation and variation in optical axis was the next most significant cause of error in calibrated focal length, having a slightly higher effect (due to the strong influence of LED ring misalignment) than variations in LED optical axis alone. Variations in LED intensity and beam width were not found to have a significant effect on calibrated focal length at the investigated values.

6 Conclusions

This paper discusses a method for generating physically accurate synthetic images for manufacturing applications where accurate location and orientation of the object outweigh its appearance. The synthesis method is able to incorporate finite aperture and lens effects, includes spectral dependence and sensor non-linearity, as well as models reflectance properties (e.g., specular, retroreflective) and illumination distributions realistically.

The physically-accurate synthetic image is simulated in a two step process. In the first step, *Radiance*, a freely-distributed soft-

Table 4 Significance of design parameters on calibrated focal length

Effect ID	t_i	
	f (mm)	Abs. %Error
A	-5.91	5.02
B	1.58	-1.32
C	-0.21	-0.33
D	5.15	-3.83
AB	-0.47	-0.82
AC	-0.96	0.21
AD	5.42	-4.34
BC	-0.26	0.75
BD	-1.84	1.06
CD	-0.26	-0.16
ABC	0.48	0.21
ABD	-2.28	1.42
ACD	-1.54	0.58
BCD	-0.13	0.76
ABCD	0.85	-0.24

ware package from the Lighting Systems Research Group of the Lawrence Berkeley Laboratory, is used to solve the radiative heat transfer equation. In the second step, differences between photo-realistic and physically-accurate image generation models are highlighted, and methods to overcome limitations for generating physically accurate images are discussed: Physically-accurate pixel values are related to the energy/area falling on the sensor pixel and must be transformed by a mapping based on the vision sensor model. Modeling the relationship between pixel gray-scale and energy/area values as a power law, and using the empirically-determined sensitivity constants for the FIVS system results in synthetic images which more correctly predict captured image values. Furthermore, the results show that both the sensor pixels and pinhole, though small, have finite size that can not be accurately approximated by infinitesimally small pixels and pinholes; as demonstrated, the finite pinhole sampling method produces significantly improved images. Additionally, the significance of ambient lighting effects on the captured image and, therefore, on the camera calibration was shown; if misunderstood, these effects can introduce apparent error in calibration results, which include the focal length, lens distortion and the transformation to determine part location and orientation. While synthetic images cannot fully compensate for the real environment, they can be efficiently used to study the effects of ambient lighting and other important parameters, such as true part and environment reflectance, geometric alignment, etc. on image accuracy.

Future work includes investigating improved reflectance and illumination modeling; this provides significant opportunities for improvement: It is unlikely that the simple reflectance modeling used to represent the board background or the retroreflective dots is sufficient to adequately predict captured-image performance. Also, the light sources chosen for the FIVS system are unlikely to be identical and may vary significantly from the typical pattern used in the simulations shown above (e.g., the distribution pattern may not be symmetrical). Distribution of ambient lighting sources may vary similarly. After addressing these issues and sufficiently

validating the methodology, simulated images will be used to study a variety of hardware design parameters and their effects on image accuracy.

Acknowledgments

This work was partially supported by a NSF Graduate Fellowship and the Woodruff Faculty Fellow Award. Comments and suggestions from Dr. Holly E. Rushmeier (currently with the IBM Watson Research Center) are greatly appreciated.

References

Chen, C.-H., and Mulgaonkar, P. G., 1991, "CAD-Based Feature Utility Measures for Automatic Vision Programming," *IEEE Workshop on Directions in Automated CAD-Based Vision*, Maui Hawaii, IEEE Computer Society Press.

Cowan, C., 1991, "Automatic Camera and Light-Source Placement Using CAD Models," *IEEE Workshop on Design in Automatic CAD-Based Vision*, Maui Hawaii, IEEE Computer Society Press.

Federal Test Method Standard 370, 1977, "Instrumental Photometric Measurement of Retroreflective Materials and Retroreflective Devices," March 1.

Gagalowicz, A., 1990, "Collaboration Between Computer Graphics and Computer Vision," *Scientific Visualization and Graphic Simulation*, D. Thalmann, ed., pp. 233-248.

Goral, C. M., Torrance, K. E., Greenberg, D. P., and Battaile, B., 1984, "Modeling the Interaction of Light Between Diffuse Surfaces," *Computer Graphics*, Vol. 18, No. 3, pp. 213-222.

Grynberg, A., 1989, "Validation of Radiance," Lawrence Berkeley Laboratory, Applied Sciences Division, Lighting Systems Research Group, Report #LBID 1575.

Kajiya, J. T., 1986, "The Rendering Equation," *Computer Graphics*, Vol. 20, No. 4, pp. 143-150.

Lee, K.-M., 1991, "Flexible Part-Feeding System for Machine Loading and Assembly, Part I—A State-of-the-Art Survey," "Part II—A Cost Effective Solution," *International Journal of Production Economics*, Vol. 25, pp. 141-168.

Lee, K. M., 1994, "Design Concept of an Integrated Vision System for Cost-Effective Part Presentation," *ASME JOURNAL OF ENGINEERING FOR INDUSTRY*, Vol. 116, pp. 421-428.

Lee, K.-M., and Blenis, R., 1994, "Design Concept and Prototype Development of a Flexible Integrated Vision System," *Journal of Robotic Systems*, Vol. 11, No. 5, pp. 387-398.

Length, R. V., 1989, "Quick and Easy Analysis of Unreplicated Factorials," *Technometrics*, Vol. 31, No. 4, pp. 469-473.

Meyer, G. W., Rushmeier, H. E., Cohen, M. F., Greenberg, D. P., and Torrance, K. E., 1986, "An Experimental Evaluation of Computer Graphics Imagery," *ACM Transactions on Graphics*, Vol. 5, No. 1, pp. 30-50.

Parker, J. M., 1996, *An Analytical and Experimental Investigation of Physically-accurate Synthetic Images for Machine Vision Design*, Ph.D. Thesis, Woodruff School of Mechanical Engineering, Georgia Institute of Technology.

Parker, J. M., and Lee, K. M., 1997, "An Investigation of Vision System Design Factors which Affect the Performance of Camera Calibration," *Proceedings of the IEEE/ASME International Conference on Advanced Intelligent Mechatronics*.

Nishita, T., and Nakamae, E., 1985, "Continuous Tone Representation of Three Dimensional Objects Taking Account of Shadows and Interreflections," *Computer Graphics*, Vol. 19, No. 4, pp. 23-30.

Rennilson, J. J., 1980, "Retroreflective Measurements" A Review," *Applied Optics*, Vol. 19, No. 8.

Rushmeier, H. E., Parker, J. M., and Lee, K. M., 1992, "Synthetic Images for Cost-Effective Machine Vision/Illumination Design," *SME Applied Machine Vision Conference Proceedings*.

Siegel, R., and Howell, J. R., 1981, *Thermal Radiation Heat Transfer*, Hemisphere, Washington, DC.

Sillion, F. X., Arvo, J. R., Westin, S. H., and Greenberg, D. P., 1991, "A Global Illumination for General Reflectance Distributions," *Computer Graphics*, Vol. 25, No. 4, pp. 187-198.

Tsai, R. Y., 1987, "A Versatile Camera Calibration Technique for High Accuracy 3-D Machine Vision Metrology Using Off-the-Shelf TV Camera and Lenses," *IEEE Robotics and Automation*, Vol. RA-3, No. 4, pp. 323-344.

Ward, G. J., Rubinstein, F. M., and Clear, R. D., 1988, "A Ray Tracing Solution for Diffuse Interreflections," *Computer Graphics*, Vol. 22, No. 4, pp. 85-92.

Wu, C. K., Cheatham, J. B., Lin, Y. H., and Cleghorn, T. F., 1990, "Computer Graphics Modelling for Simulating and Testing Robot Vision Systems," *International Journal of Modelling and Simulation*, Vol. 10, pp. 67-70.