Invited Paper

Practical Design Considerations for Modern Photographic Optics

Ken Rockwell, KenRockwell.com, Post Office Box 8778, La Jolla, CA 92038-8778

ABSTRACT

In the 1950s, cameras and lenses didn't communicate. Optical bench tests were sufficient to characterize lens performance completely. Today, optics are only one small factor in the much larger electro-optical systems which are today's cameras. Bench performance alone isn't enough. For a lens to be useful it must work in concert with modern cameras' focus, exposure and data systems. This paper explores these numerous issues and their influences on lens design. Outside-the-barrel ergonomic and systems issues are covered in SPIE paper 7428-19.

Keywords: SLR, TTL, Falloff, Distortion, Diaphragm, Digital Capture, Photography, Photographic, Autofocus.

1. INTRODUCTION

Many more factors in addition to optical performance are important as design parameters for modern photographic lenses.

This paper covers what a lens needs to do inside its own barrel to work well with today's cameras. Paper 7428-19 covers what needs to happen outside the barrel to make a lens a commercial success, and popular with photographers in the larger scheme of a system of several lenses as well as ergonomically.

As I go through history, note that the requirements added on in each era remain to this day as additional requirements for modern lenses.

This paper applies to all still photography. It's apparent concentration to 35mm and digital single-lens reflex (SLR) cameras is because these cameras have evolved to the highest level of sophistication.

2. 1839 - 1959

In the good old days before the 1960s, lenses simply bolted onto a camera, and you were done. There wasn't even light meter coupling.

The most advanced cameras were 35mm rangefinder cameras like the Leicas, Nikons and Contax that peaked in the 1950s, where at least the lens' focus was coupled to the camera's rangefinder by means of a cam and roller. Beyond this, the lens was on its own.

In those days, bench tests meant everything. Mechanical precision was tight, and the performance seen on a bench would almost exactly match performance as seen by the user in-camera and on-film. The biggest potential loss of performance was inaccurate rangefinder coupling for longer lenses.

Everything was shot on film. Unless one was making huge enlargements and was very careful about enlarger alignment and using the right enlarger lens at the right aperture to avoid diffraction (or a microscope), no one except the lens designers ever saw what a lens really could do.

Great lenses of this era are still good today, while the sloppier ones still passed muster with 99% of picture takers.

Decent acutance and freedom from distortion was enough.

Current Developments in Lens Design and Optical Engineering X, edited by Pantazis Z. Mouroulis, R. Barry Johnson,
Virendra N. Mahajan, Proc. of SPIE Vol. 7428, 742804 · © 2009 SPIE · CCC code: 0277-786X/09/$18
doi: 10.1117/12.828380

3. 1960: AUTOMATIC DIAPHRAGMS

The 1960s saw the popularization of the single-lens reflex (SLR) camera.

Automatic diaphragms were added that open and close as the picture is taken. This allows full-aperture viewing, focusing and metering.

The diaphragms needed to open and close at rates of 10 frames-per-second (FPS) or more, and close fast so there is as little delay as possible after the shutter is pressed.

They have to be fast, and they also have to be precise and accurate. In photography, one needs to be within not more than a third of a stop (±25% transmission) for correct exposure.

Poorer lenses still don't meet this acuracy requirement, and lead to exposure errors. Few cameras use closed-loop exposure measurement, so errors in the static positions of the diaphragm blades and the dynamic ability of the blades to get to the correct position go uncorrected by through-the-lens (TTL) light meters.

4. 1970: TTL COUPLED LIGHT METERS

The 1970s brought the popularity of TTL light (exposure) meters. This takes lens and filter transmission into account, and frees the photographer from having to compensate for them.

That's good, but since SLR cameras meter at full aperture, cameras and lenses need to communicate to the metering system by how much transmission will be reduced as the aperture is set automatically to the smaller openings used to make the photo.

This communication started as mechanical feelers reading the position of the aperture ring. Exposure was calculated electrically and read on a needle of a manual light meter in the camera.

Rheostats we usd at first as encoders for analog light meters. THese later became became digital encoders, but all of them have the same potential for mechanical error. Any errors in reading and interpreting the final shooting aperture leads to exposure errors, which again are not corrected my most camera systems because they run open-loop.

5. 1970: TTL COUPLED LIGHT METERS

The 1970s brought the popularity of TTL light (exposure) meters. This takes lens and filter transmission into account, and frees the photographer from having to compensate for them.

That's good, but since SLR cameras meter at full aperture, cameras and lenses need to communicate to the metering system by how much transmission will be reduced as the aperture is set automatically to the smaller openings used to make the photo.

This communication started as mechanical feelers reading the position of the aperture ring. Exposure was calculated electrically and read on a needle of a manual light meter in the camera.

Rheostats we usd at first as encoders for analog light meters. THese later became became digital encoders, but all of them have the same potential for mechanical error. Any errors in reading and interpreting the final shooting aperture leads to exposure errors, which again are not corrected my most camera systems because they run open-loop.

6. 1980: EXPOSURE AUTOMATION

The simplest automatic exposure systems of the late 1970s read the light coming through the lens, and adjusted the exposure time accordingly.

More complex systems in the early 1980s also varied the lens' aperture as needed. Requiring the camera to set the diaphragm introduces yet another source of error.

This added potential for error led some cameras of the early 1980s to use closed-loop light metering, like the Nikon FA. These systems meter open-loop as the photographer used the camera, but just before the actual instant of exposure, would sample-and-hold (S/H) one last light reading after the diaphragm had stopped down fully to its taking aperture.

This is great, but could slow the camera slightly, and introduced errors if the diaphragm moved slightly after the S/H took its reading. Diaphragm blades are flying around fast, and at small apertures, errors still crept in, even with closed-loop systems.

There are two common methods to control the diaphragm from the camera.

Some cameras, like all Nikons since 1959 and manual-focus Canons like the AE-1, use mechanical linkages.

Other cameras, like Canon's current EF system introduced in the late 1980s, use servo motors attached to the diaphragm in each lens, as controlled electronically from the camera.

7. 1985: AUTOFOCUS (AF) AND DATA

7.1 AF Sensing

Autofocus SLR cameras most often use phase-detection schemes to focus as quickly as possible.

The problem is that limited focus-sensor (CCD) resolution limits precision and limited camera processing power limits speed.

Even with unlimited speed and precision, accuracy is always limited by the mechaical accuracy of the positioons of the additional half-silvered and sub-mirrors added into the optical path between the lens and the AF sensors.

The AF sensors are typically in the bottom of the camera, seeing images passed through several relay mirrors, while the photographer is viewing the image on a screen at the top of the camera.

Any mechanical errors in the many mirrors leads to focus shifts because any change in path length to the sensors won't match the path length to the film.

AF systems typically work by processing an image from the lens from an annulus of about f/5.6. This lets these systems work with lenses of f/5.6 and slower.

If a lens has no spherical-aberration induced focus shift, there is no problem at other apertures.

If a lens exhibits focus shift, focus at larger apertures may exhibit an obvious shift either in front of or behind the intended subject. Focus at smaller apertures has never been a problem since the increased depth-of-feild hides it.

Modern AF systems read lens data electronically, and AF systems can be programmed by their manufacturers to compensate as needed.

As of about 2005, some Nikon and Canon AF cameras allow users to program their own slight offset corrections, however these offsets remain constant regardless of set aperture or zoom-lens focal length. Since spherical aberration and focus shift will vary with these settings, this one correction value can't possibly address all lens settings.

7.2 AF Drive

In order to allow lenses to be focused quickly via electric motors, frequently the mechanics are made with more play to allow less friction, thus faster focus with less battery power. This can lead to poor optical alignment in practice.

Clever lens designs often use internal or rear-element focus in lieu of moving the entire optical tube.

Some AF systems, like Nikon's, use mechanical coupling into the lens, with a motor in the camera.

Other systems, like Canon's and more recent Nikons, have only electronic coupling to the lens, with focus motors located in each lens.

SLRs autofocus in real-time, closed loop TTL. Rangefinder cameras lack the advantage of TTL focus, thus more errors creep in.

7.3 Data Communication

Not obvious until years later, Nikon's AF lenses, which although controlled mechanically, had ROMs to communicate static lens data to the camera. These ROMs always encoded the lens' focal length and maximum aperture so that the light meter could measure absolute, not just relative, subject luminance.

Even the earliest lenses from Nikon and Canon encoded the focal length settings of zoom lenses, and this data from 1980s lenses is read and recorded today by all their current digital SLRs.

Nikon added distance encoding in 1992, calling these lenses "D" lenses. By encoding the subject distance, the exposure system had additional data with which to predict flash exposure regardless of subject reflectance, and could make better guesses about the subject to enable Nikon's very clever "Matrix" meter to calculate exposure even more consistently.

(The Nikon Matrix meter (and other evaluative meters), beyond the scope of this paper, don't just measure light. They attempt to interpret the subject based on absolute luminance, and calculate a pleasing exposure. For instance, if the camera measures luminance which would suggest a white object in full sunlight, it knows to give additional exposure so the white object doesn't appear gray instead.)

8. 2000: DIGITAL CAPTURE

With film, it required skill and instruments to see what the lens really put on the film.

With digital capture, anyone can zoom in all they want on their computers to see any flaws, whether photographer error or lens design.

There is a significant portion of the market, as always, more concerned about lens sharpness than photography itself.

Now these folks easily can see, compare and share their impressions, even if misguided.

Regardless of how misguided, the Internet makes it easy to share any observations, and the worse they are, the faster they travel.

Good photographs are never an imitation of reality. Good photographs are designed by photographers to communicate one message as clearly and strongly as possible. This is done by designing the layout of an image in such a way that our eyes are naturally attracted to their basic patterns and relationships, and then once there, have details which lead there eyes in and around the image. Just like designing lenses, it takes quite a few years of higher education (art school) to learn this, so I won’t get into it here.

8.1 Modulation Transfer Function (MTF) Performance

In the old days, designers balanced contrast, acutance against ultimate resolving power. In the 1950s, the Germans tended to strive for ultimate resolving power at the expense of lower-spatial-frequency contrast, while the Japanese shot for high contrast at the expense of ultimate resolving power. To see this, compare some Nikon to some Leica or Schneider lenses of the 1950s.

Today, since users can see things so easily digitally, but digital sensors are blind above the Nyquist frequency, it's best to trade slightly lower peak performance for more consistent performance across the field. Casual users looking for lens limitations can see variations across the field more easily than they can see absolute performance.

A lens which exhibits a constant 80% performance (an arbitrary figure) from center to corner will look wonderful under peering eyes.

A design which exhibits 97% performance in the center, but falls to 80% performance in the corners, will look worse because casual observers will see the degradation that otherwise would have no reference inside each image.

Likewise, it's better to have saggital and meridional MTFs track each other, even if at a lower value, than to have them diverge. Divergent saggital and meridional MTFs lead to smearing, which looks worse as an obvious lens design limitation than simply becoming benignly softer.

Users today may return lenses to dealers as "defective" when they see this. This is common in ultra-wide angle zooms.

8.2 Distortion Performance

Many tools exist to correct distortion electronically.

In Adobe Photoshop, recent versions (since CS2 or about 2006) provide simple corrections for first-order distortion.

It may be better to have more first-order distortion, because it is easily corrected, than to have less overall distortion, but with a more complex, and therefore more difficult to correct, signature.

8.3 Out-of-Focus Performance (Bokeh)

In actual photography, most of the image is not in perfect focus. Residual spherical aberration has a very visible effect on out-of-focus areas.

In a perfect lens, blur circles are circles, convolved with Airy disks. We'll ignore the Airy disks here and concentrate on the blur circles, because these effects become obvious when the blur circles become large enough to see.

Perfect blur circles have sharp edges, thus every point of light in the subject is represented by a corresponding blur circle in the image. If we have a subject with bright points of light, the blur circles become obvious.

Residual spherical aberration alters the distribution of light within each circle.

In photography, we use out-of-focus areas with deliberate intent to make that part of the image, usually a background, "go away" and become as flat and smooth as possible.

Perfect blur circles have sharp edges. These sharp edges contain energy at higher spatial frequencies, which adds perceived detail in a part of an image a photographer wants to be a softly blurred background.

What photographers really want are blur circles with Gaussian, not rectangular, distribution functions.

Photographers use "bokeh," a word from Japanese, to describe the look of the out-of-focus areas of an image.

Most lenses have neutral bokeh: blur circles are circles.

A lens with great bokeh has softer edges to its blur circles, eliminating higher spatial frequency distractions in the out of focus areas. This renders backgrounds soft and undistracting.

Bad bokeh is when a lens has blur circles in which there are more intensity at the edges of the circle, with darkened centers. These blur circles emphasize their own edges, and can render distracting detail in supposedly out of focus areas.

Bokeh changes sign behind or in front of the in-focus subject. Neutral is neutral, but if a lens has good bokeh for out-of-focus objects behind the subject, it will have bad bokeh for anything out-of-focus in front of the subject.

Photographers strive not to have out-of-focus objects in front of the subject. It looks weird and distracts the eye from the intended subject. Backgrounds are often supposed to be out of focus, which helps the viewer's eyes and mind concentrate on the intended subject.

Lenses should be designed for better bokeh in the backgrounds (the region behind the subject), at the expense of bokeh for objects in front of the subject.

8.4 Diaphragm Design

In some situations, the diaphragm opening has a visible effect on the image.

There are two visible effects: the shape of blur circles of out-of focus areas, and sometimes shafts of light are introduced which appear to radiate from brilliant points of light.

Diaphragms should have odd numbers of blades, like 5, 7 or 9 blades. 6- or 8- bladed diaphragms are less desirable.

The shape of out-of-focus areas' blur circles matches the shape of the open aperture. Out of focus points of light become hexagons with 6 diaphragm blades. These distract viewers, since they occupy the brain's processing time that should be used to pay attention the photographer's intended subject. Likewise, 8-bladed diaphragms beget octagonal shapes, and key our brains into thinking "Stop!" subconsciously.

5-bladed diaphragms are poor here, since we recognize pentagons, but consciously, they aren't that distracting since they aren't suggesting Stop,! snowflakes or honeycombs. There aren't many 5-sided things in nature to distract our subconscious attention. Likewise, septagons and nonagons have few, if any, natural parallels.

Worse, even numbers of blades create blur circle polygons with parallel sides. This potentially adds to poorer bokeh and more distracting out of focus backgrounds. The spatial energy in the images' out of focus areas can have the energy from straight sides of these blur circles combine at fewer angles. Odd-sided polygons have no parallel sides, so spatial energy is more evenly distributed by angle.

Most importantly, bright specular highlights, like the sun, bright lights in nighttime images or reflections of the sun from car windshields, often turn into little sunstars with lines radiating from them. Look at any car magazine, and you'll see these stars in the photos of cars made outdoors.

Light is diffracted from the knife-edge of the blade, and reflected off its edge. Straight blades make sharp lines from specular sources, while curved blades draw broader strokes.

Photographers use these little star patterns when they design photographs.

6-bladed diaphragms beget simple 6-pointed stars, and 8-bladed diaphragms beget 8-pointed stars.

Most photographers prefer the more subtle 10-, 14- or 18-pointed stars that are created from 5-, 7- or 9-bladed diaphragms. We get twice as many radiating lines from odd-numbered diaphragms because the radiating lines from even-numbered diaphragms superimpose upon each other.

Photographers are visual artists. Subtle effects like sunstars and bokeh are significant — much more significant and visible than sharpness. As a photographer myself, I have I have always considered a lens' number of diaphragm blades as a key part of my purchase decisions. Some brands, like Nikon, use odd numbers (7 and 9), while other brands, like Canon, usually use even numbers (6 and 8). This is another reason I've preferred using Nikon for over 25 years. The effects the diaphragm has on the image cannot be changed later in Adobe Photoshop.

8.5 Peripheral Falloff of Illumination Performance

Digital photography sensors are far more sensitive to light at the same signal-to-noise ratio (SNR) than film. Therefore they easily operate at photographic speeds (ISO ratings) ten to a hundred times more than film with the same SNR.

Even with these high speeds, longer focal length zoom lenses are often shot at or near their full apertures in daylight, especially at the longest ends of their zoom ranges.

Therefore care should be paid to the falloff characteristics, since often photos made with long zooms at long settings include a blank sky, which will make obvious any falloff.

Especially annoying is falloff which abruptly changes as one nears the corners of the image. Far better is falloff which changes gradually as one leaves the center of the image.