Edmund Optics™

best of EO® applicationNOTES

technical documentation

2000-2003

revised edition

2000-2003

Best of EO Application Notes

Table of Contents

Introduction

A Reading Assignment You'll Be Glad About!

Put on your thinking caps — because Edmund Optics is presenting its Best Application Notes ever.

These application notes cover a wide variety of topics, many in great detail. The notes focus on the things that Edmund Optics knows best: machine vision, the choice between custom vs. off-the-shelf optics, and telecommunications optics. All of the notes are written by in-house Edmund engineers, the very same engineers that are waiting to answer your Technical Support calls and e-mails.

And make sure you don't miss our Tech Tips scattered throughout! Want to know how to make an inexpensive light stop? Want to know how to tell the different axes of a polarizer? These little nuggets of wisdom have been passed along from our Applications Engineers to you.

So study up — because these notes will make your design work easier and more problem free. But don't take our word for it — get reading and find out for yourself!

John Stack
President and CTO

CONTACT INFO

PHONE
800-363-1992

FAX
856-573-6295

MAIL
Edmund Industrial Optics
Order Department PD002
101 East Gloucester Pike
Barrington, NJ 08007-1380

E-MAIL
sales@edmundoptics.com

WEBSITE
www.edmundoptics.com

How to Reduce the Cost of Configuring a Vision System

FIGURE 1: *Fundamental parameters of an imaging system include the resolution of the object, the field of view, and the depth of field that the user wishes to image. The working distance, from the object to the lens, is also important, as is the sensor size. The primary magnification is the field of view divided by the sensor size.*

When building a vision system, one must consider the application, resolution, illumination, depth of field, field of view, processing speed, and other elements. But all too often, systems are built that either fail to meet performance expectations or utilize components that are over-specified. Both pitfalls are expensive in the long run because an under-specified system that fails must be redesigned until it works; and an over-specified system contains components that are more expensive than needed. To avoid these pitfalls, pay attention to specifications. In this article we describe the parameters of a vision system so that you can specify a system that meets your needs. We also suggest some specific cost-saving strategies.

Because the purpose of a vision system is to extract necessary information from an image, the application determines the required image quality. A system with sufficient image quality for one application may not be sufficient for another. The opposite can also be true, with many applications using over-specified components that do little more than increase cost. But what is image quality? There are two complementary ways of looking at the issue: first, the image quality of a system is the result of the image quality of the components; second, image quality is specified not by a single number, but by several factors discussed below.

Equipment basics and the application

The imaging ability of a system is the result of the imaging ability of the components. Any vision system needs illumination, a lens, a camera, and either a monitor or a computer/capture board to analyze the images. Even the electronics cables and the user's eyes affect the entire system's image quality.

It does no good to specify a high-resolution camera for use with a low-resolution monitor. Ideally, one chooses components to fit the application and complement each other. By avoiding over-specifying the quality on some parts of the system, one ensures that none of the components is more expensive than necessary.

The needs of the application determine image quality — does the vision system have to capture images quickly? (That also affects the processing speed of the system.) Does it need to check the orientation or color or size of the workpiece, or just detect its presence? How large is the smallest necessary detail? How much contrast is necessary?

In order to talk about the needs of the application, we need a vocabulary for image quality.

Image quality

Image quality consists of a number of fundamental parameters (see Figure 1):

continued >

- Field Of View (FOV): The viewable area of the object under inspection. In other words, this is the portion of the object that fills the camera's sensor.
- Working Distance: The distance from the front of the lens to the object under inspection.
- Resolution: The minimum feature size of the object under inspection.
- Depth Of Field (DOF): The maximum object depth that can be maintained entirely in focus. The DOF is also the amount of object movement (in and out of focus) allowable while maintaining an acceptable focus.
- Sensor Size: The size of a camera sensor's active area, typically specified in the horizontal dimension. This parameter is important in determining the proper lens magnification required to obtain a desired field of view.

In addition to resolution and depth-of-field, as mentioned above, image quality is also a combination of three other properties: image contrast, perspective errors, and distortion.

Resolution, contrast, and MTF curves

By considering the relationship between resolution and contrast one can understand the tremendously useful modulation transfer function, or MTF.

Resolution is a measurement of the imaging system's ability to reproduce object detail. For example, imagine a pair of black squares on a white background. If the squares are imaged onto neighboring pixels, then they appear to be one large black rectangle in the image. In order to distinguish them, a certain amount of space is needed between them. Determining the minimum distance needed to see the two squares yields the limiting resolution of the system. This relationship between alternating black and white squares is often described as a line pair. Typically the resolution is defined by the frequency measured in line pairs per millimeter (lp/mm).

There are two different but related resolutions in play here: the resolution in object space (the size of elements in the object that can be resolved), and image space resolution (a combination of the lens resolution and camera resolution). The sensor's line pair resolution can be no more than half the number of pixels across on the sensor because a pair of pixels is the minimum required to discern a black and white area. The image and object space resolutions (described in lp/mm) are related by the primary magnification of the system.

The limiting resolution of the system can be determined experimentally by imaging a test target (see Figure 2). A bar target consists of line pairs with varying frequencies, whereas a star target consists of wedges with a continuum of frequencies. The orthogonal lines in a bar target are useful because they allow users to test the system for errors that show up differently in the x and y planes of an image (in other words, astigmatic errors). Bar targets, however, are limited by having a finite number of steps in frequency. Star targets do not have this drawback, however they can be more difficult to interpret.

Contrast describes how well the blacks can be distinguished from the whites. In real life black and white lines will blur to some degree into grays. Noise and blurring of edges will cause the contrast to go down. How effectively the differences between boundary areas on the image are reproduced relative to one another is often defined in terms of grayscale or signal-to-noise. For an image to appear well-defined, the black details need to appear black and the white details must

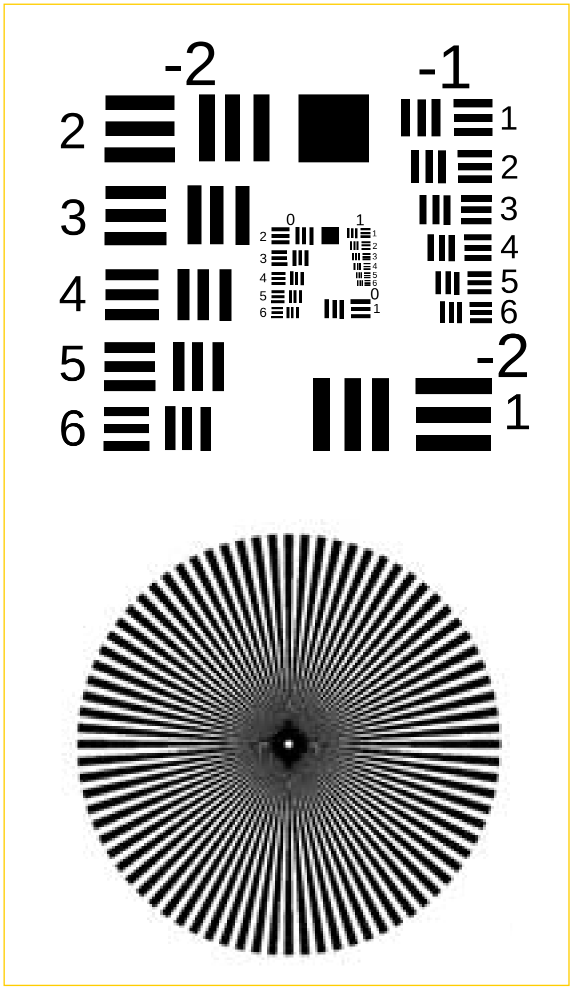


FIGURE 2: Two test targets: a bar target and a star target allow users to measure resolution and astigmatic errors.

$$\text{Image Space Resolution} = \frac{\text{(Object Space Resolution)}}{\text{(Primary Magnification)}}$$

continued >

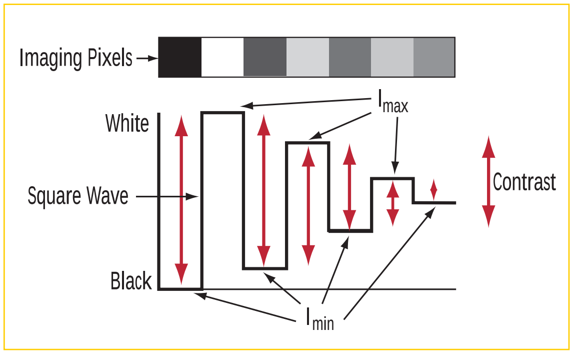


FIGURE 3: *Contrast is the difference in intensity between blacks and whites. For an image to appear well-defined, black details must appear black and white details must appear white. The greater the difference in intensity between a black and white line, the better the contrast. The human eye can see a contrast of as little as 1-2%. A typical limiting contrast of 10 to 20% is often used to define the resolution of a CCD imaging system.*

$$\% \text{ Contrast} = \frac{(I_{max} - I_{min})}{(I_{max} + I_{min})}$$

where I_{max} is the maximum intensity
and I_{min} is the minimum intensity

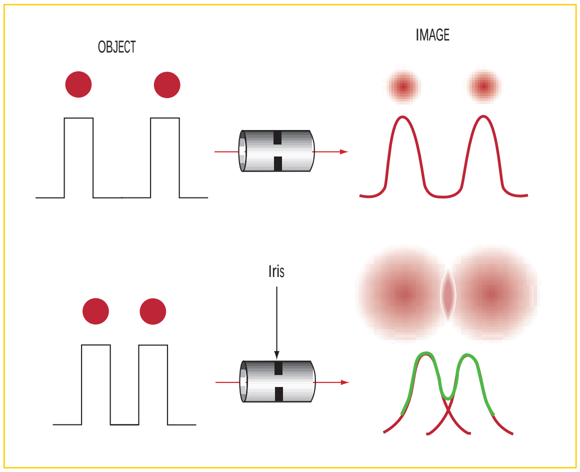


FIGURE 4: *Contrast is not constant! It depends on frequency. The dots at the top of the figure can be imaged through a lens. They blur slightly. If we moved the spots closer, their blurs overlap and contrast decreases. When the spots are close enough that the contrast becomes limiting, that spacing is our resolution.*

appear white (see Figure 3). The greater the difference in intensity between a light and dark line, the better the contrast. This is intuitively obvious, but it is more important than it may first appear. The contrast is the separation in intensity between blacks and whites.

Reproducing object contrast is as important as reproducing object detail, which is essentially resolution. The lens, sensor, and illumination all play key roles in determining the resulting image contrast. The lens contrast is typically defined in terms of the percentage of the object contrast that is reproduced. A sensor's ability to reproduce contrast is usually specified in terms of decibels in analog cameras and bits in digital cameras.

The resolution and contrast of an image can be defined individually, but they are also closely related. In fact, resolution is often meaningless unless defined at a specific contrast. Similarly, contrast depends on resolution frequency. Consider two dots placed close to each other and imaged through a lens (see Figure 4). Because of the nature of light, even a perfectly designed and manufactured lens cannot accurately reproduce an object's detail and contrast. Even when the lens is operating at the diffraction limit, the edges of the dots will be blurred in the image.

When they are far apart (in other words, at a low frequency), the dots are distinct, but as they approach each other, the blurs overlap until the dots can no longer be distinguished. The resolution depends on the imaging system's ability to detect the space between the dots. Therefore, the resolution of the system depends on the blur caused by diffraction and other optical errors, the dot spacing, and the system's ability to detect contrast.

Optical engineers usually specify a contrast level at a specific resolution. When a plot is made of contrasts at a range of frequencies, you have a Modulation Transfer Function (MTF) curve.

Suppose we imaged a target of black and white parallel lines. Consider the effect of progressively increasing the line spacing frequency of a target and how this might affect contrast. As one might expect the contrast will decrease as the frequency increases. The Modulation Transfer Function (MTF) is plotted by taking the contrast values produced by a series of different line pairs. The curve drawn from these points shows the modulation (in other words, the contrast) at all resolutions, not just at the limit resolution.

The high-resolution end of the curve is not always the most important part of an MTF! For many applications, a high contrast at a low frequency is more important than the limit of resolution. For such applications, a higher-resolution lens (for example, one designed to work with film rather than with CCDs) will not improve the overall system — although it will increase the cost. Instead, brighter illumination may be all that is needed.

Other strategies for reducing costs

There are some other specific strategies that designers can apply to reduce costs. The overriding theme is to eliminate unnecessary complexity: keep the system as simple as possible. Here are four strategies that can be used together or separately:

- eliminate colors
- fix apertures
- eliminate folds
- use off-the-shelf optics

Do you need white light or would monochromatic illumination work just as well? If you only use one color, then chromatic aberration

continued >

is no longer an issue. If the system does not need to be color-corrected over the entire spectrum, the lens design is simpler. Going monochromatic may also simplify the illumination system since monochromatic LEDs use less power and create less heat than white light incandescent bulbs.

If you can fix the system apertures, while maintaining the ability to focus the system, this also simplifies the optical design and can reduce the number of elements in the system. All other things being equal, fewer elements mean less cost.

Folds in the optical path introduce aberrations. If at all possible, lay out the path in a straight line. Also, you avoid the cost of the folding mirrors.

Finally, use off-the-shelf system elements when possible. Unless the vision system will be produced in quantities of several hundreds or thousands, off-the-shelf optics will be cheaper than custom-made optics. They will certainly be faster to obtain than custom optics.

Most design software packages have off-the-shelf lenses preloaded into them. The sooner off-the-shelf options are worked in, the better. Typical design software will give a starting point with custom lenses when one optimizes all surfaces. Then one can force the software to replace the custom lenses with the closest off-the-shelf matches and allow air spaces to compensate. The best time to do this is before starting on the mechanical designs.

Off-the-shelf solutions and design

All of these parameters can lead to exacting specifications for a lens. This often leads integrators to want a custom lens, feeling that an off-the-shelf lens could not fit all of the parameters correctly. However this is often an easier task than it may seem. It's pretty obvious that an off the shelf solution will be more cost effective than a custom solution. In optics this is very true. It takes high volumes to efficiently manufacture lenses. There often is very costly and time consuming design necessary to come up with a lens that will work properly.

With the tremendous growth in machine vision, off-the-shelf video lenses are more common. We at Edmund carry a very wide selection of lenses designed specifically for machine vision applications. Usually a little bit of flexibility on one or more parameter allows for an easy selection of an off-the-shelf lens. A simple and cost effective way to work an off-the-shelf lens into a design is to choose a lens before the mechanical design of the entire system is finished. The majority of times I walk into a customer's facility where they are trying to solve a vision problem, the lens was the last thing to be considered. The delay in considering it often leads to very difficult mechanical constraints to deal with. The housings usually could have been changed early on, but by the time the lens is integrated it would be too difficult.

There are times custom makes sense. For example when you have very tight requirements for working distance and packaging. Also if specific fields of view are necessary a custom fixed lens that gets the correct field of view may still be cheaper than a zoom lens that happens to get the right field of view with a lot of unused adjustability. Also for very high volume the costs of design and set up can be amortized and repeating costs like the number of lenses and mechanical adjustments can often be minimized with a custom design.

The disadvantages of custom can also be avoided by using off-the-shelf elements in the design. Lead times for manufacturing of lenses is a significant factor to consider. Production of a custom lens can

continued >

easily take 8-10 weeks to manufacture. The price also associated with test plate and tooling can be high as well. By using off-the-shelf elements you avoid the set up fees and the lead times.

Most catalog lenses are already in popular design packages, which makes designing with them easier. This also allows for a quick prototype that helps in proving the concept of the design before an investment in design and manufacturing. Many applications can easily be solved with simple telephoto or reverse telephoto designs using two achromats.

Finally if the design is very demanding a ground up custom may be necessary. We see many applications each month where there is no other option than a custom design. In these cases we always make sure the customer knows that there will be a significant cost and lead time for a design. Though we can often produce a design in rather short time, the prototypes will always take sometime to produce if we could not use off-the-shelf components.

When we go to a custom design, the trade off we gain for the higher lead time and the higher cost of design is that we can often save money in the long run for high volume. One way to do this is, since we are designing the lens for a known application, we can remove some of the adjustments that are normally built into an off-the-shelf lens. If the illumination is going to be constant we can make the iris fixed to only one setting. We can design the lens to get the performance that is required without over designing it beyond the needs of the application.

It is also good to have the manufacturer be part of the design process. We do many custom designs because we also manufacture the lenses. We can often save time and money designing to tooling and test plates we already have. Also we understand our tolerancing and manufacturing better than anyone so we can design specifically to the manufacturing abilities.

Conclusion

The first step in an efficient vision system is to properly specify the necessary requirements and the degrees of freedom. The more degrees of freedom, the easier it will be to find an off the shelf solution. If a custom design is necessary look first to solutions using off-the-shelf components to make design, prototyping, and production faster and cheaper.

The basic cost-saving strategy for vision systems is to specify what you need, and no more. Apply as much intelligence as possible when specifying the system, and use some common sense tips to reduce costs during design. If you do, then your system will fill your needs efficiently.

TECH TIP ON CHOOSING MONITORS

Video systems require video compatible monitors rather than computer monitors. Monitor specifications, such as signal format, component level signals, and relative resolution, must match the input device. Based on the resolution of the human eye, the following equation may be used to yield the maximum monitor viewing distance:

$$\text{viewing distance} \leq \frac{4148}{\text{TVL}} \times \frac{\text{Screen}}{\text{Diagonal}} \quad \text{(in inches)}$$

Set Your Sights on Vision

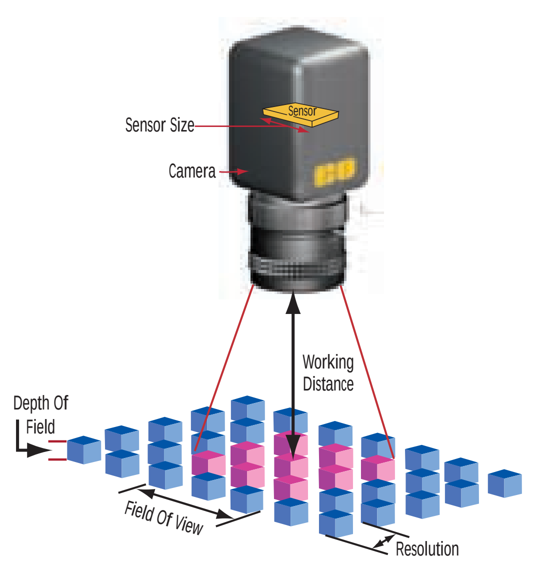


FIGURE 1: *Fundamental parameters of an imaging system.*

FUNDAMENTAL PARAMETERS OF AN IMAGING SYSTEM

Field Of View (FOV): The viewable area of the object under inspection. In other words, this is the portion of the object that fills the camera's sensor.

Working Distance: The distance from the front of the lens to the object under inspection.

Resolution: The minimum feature size of the object under inspection.

Depth Of Field (DOF): The maximum object depth that can be maintained entirely in focus. The DOF is also the amount of object movement (in and out of focus) allowable while maintaining an acceptable focus.

Sensor Size: The size of a camera sensor's active area, typically specified in the horizontal dimension. This parameter is important in determining the proper lens magnification required to obtain a desired field of view.

Semiconductor equipment such as wire bonders and surface profiling equipment requires integrated sensors that can monitor a process or locate material, and these sensors are often optical imaging systems. Despite this fact, many semiconductor equipment manufacturers which employ entire groups of mechanical, electrical and software engineers have only a single engineer in charge of optical systems.

And yet, the need for integrating optics into the machinery has never been greater. And because space is always at a premium in a fab cleanroom, system designers have little elbow room. Often, integrating a vision system means snaking the optical system through the equipment without interfering with the primary process, be it wire-bonding, die packaging, aligning wafers, or lining up registration marks before lithography or metrology — many of the common fabrication processes can benefit from using optics.

If you can't afford a large optical engineering department, you can apply a design strategy for implementing imaging systems within the tight space constraints of your equipment. The steps are straightforward:

1. Define the image quality you need
2. Determine whether the quality is feasible
3. Prototype
4. Place the lighting
5. Make it fit
6. Reduce production costs

Image quality

The primary purpose of any imaging system is to obtain sufficient image quality to extract necessary information. There is no single number that determines image quality. Before you can specify your vision system needs, spend some quality time with the object you want to view.

For the fundamental parameters of an imaging system, see Figure 1 and the sidebar, at left.

Another useful descriptor of the system, the primary magnification of the lens, is the ratio between the sensor size and the field of view. It is not typically used as a fundamental parameter.

In addition to resolution and depth-of-field (see sidebar, at left) image quality is also a combination of three other properties: image contrast, perspective errors, and distortion (see Figure 2).

The point of considering all these factors is to determine the minimum acceptable image quality. Defining the minimum image quality is crucial. Tightly packed optical systems all have one thing in common: they sacrifice lots of image quality to accommodate for mechanical constraints. In addition, truly understanding image quality requirements can mean huge savings in both time and money. Know your minimums!

Will it fit?

Once you have nailed down your basic parameters — what you absolutely must have — it is time to crunch some numbers to find a combination of focal lengths and object/image distances that will

continued >

work — or to determine that no feasible system can fit the requirements you've come up with.

The bad news is that this usually involves working through thin lens equations, which you probably saw last in a college physics textbook. In addition, those equations can lead to very misleading results. The good news is that you don't have to do it yourself. Get on the phone and start calling optical companies. Any modern optical company worth working with has optical design software that can quickly and easily provide a preliminary solution. The even-better news: Unless your problem is extremely complex, this service is usually free

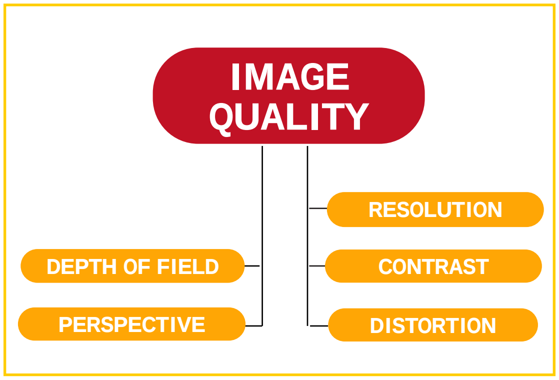


FIGURE 2: *A variety of factors contribute to the overall image quality, including resolution, image contrast, depth of field, perspective errors, and geometric errors.*

Prototype

Prototype using off-the-shelf components. You will find that off-the-shelf prototyping is fast, inexpensive, and it allows you to confirm image quality requirements. Even the best optical designer cannot perfectly predict the effects of illumination and object surface characteristics.

Another word of hard-won wisdom: set up your initial prototype in a straight line. The final optical system will, no doubt, have a number of bends and twists in it, but at this point, you need to understand the basic effects of lenses, apertures, CCD's and illumination. If you don't become familiar with these characteristics of your system at this point, debugging later can become a nightmare. Even if you have to make special metal to hold the straight-line system, it is well worth the time.

Finally, make sure your aperture sizes are realistic. Chances are, you have chosen lenses with diameters that won't fit in the mechanical space you've allotted for the optics. Use apertures to simulate the diameters that you can realistically expect.

Tightly packed optical systems sacrifice image quality to accommodate mechanical constraints.

Illuminate!

Most imaging systems that fail do so because the objects in the field of view are improperly illuminated. Without sufficient illumination, the system's contrast suffers, and the image quality therefore suffers as well. Contrary to popular belief, contrast is more important than resolution in many imaging systems.

Let's repeat that: contrast is more important than resolution in many imaging systems.

For an image to appear well-defined, the black details need to appear black and the white details must appear white. The greater the difference in intensity between them, the better the contrast. For the imaging system to have a chance of transmitting a good contrast image, however, the object has to be illuminated in a way that provides good contrast to begin with.

Know your angles: illumination is all about geometry. Consider the relationship between lighting geometry and surface features in these examples (see Figure 3):

Diffuse light from the front can be provided by fluorescent linear or ring lamps and minimizes shadows and specular reflections, but it also makes surface features less distinct.

Single-directional glancing incidence lighting, such as from fiberoptic light guides, goes to the opposite extreme: it shows surface defects and topology very well, but also causes extreme shadows and bright spots.

Directional illumination, provided by one or more fiberoptic light guide offers more moderate properties: strong relatively even lighting

continued >

but with some shadows and glare.
Ring lights, provided by fiberoptic or LED ring light guides, reduce shadows and provide relatively even illumination, but can sometimes be difficult to mount and can sometimes create a circular glare problem from highly reflective surfaces.
Polarized lighting, provided by a regular light source with a filter attached, provides even illumination but offers less intensity through the polarizer.
Diffuse axial lighting can be offered by LED axial illuminators or fiberoptic-driven axial adapters, and offers shadow-free even illumination with little glare, but requires an internal beam splitter which reduces the intensity.
Structured light, which can be provided by a line-generating laser diode or a fiberoptic line lightguide, is very useful for extracting surface features, but the disadvantage of using a laser is that some colors may absorb the intense light and heat up.

Most imaging systems fail because the object is improperly illuminated.

Optical engineers love LEDs. The use of monochromatic LED's solves a lot of imaging problems and simplifies optical designs: the main benefit is that if you use only one color of light, then chromatic aberration simply isn't a factor. As with most things, however, there is a price to pay: LED illumination can be uneven and not provide enough energy where you need it. To fit your purpose, LED generated light may need to be reshaped, diffused or directed by a lens.

Debugging illumination can be tricky. Two tools you should not be without are a flat mirror and chrome ball bearing. These two surfaces accurately show the location and intensity of your illumination sources regardless of object surface characteristics.

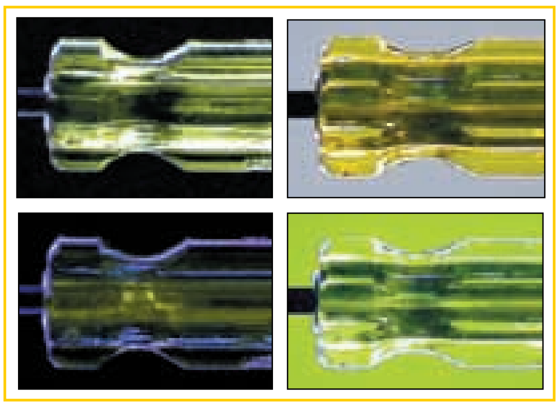


FIGURE 3: *What you can see depends on the light you shed on the object. Different types of illumination can solve or create problems for your imaging system.*

Making it fit

Now that you've ironed out the basic optical path and illumination, you get to make the system fit in the space allotted to it. When you start adding folds and combining optical paths, you start earning your keep. This looks easy on paper, but it can be a tolerancing and debugging hell. While we can't make this easy, we can mention some details to think about:

Mirror thickness has a direct effect on image quality. While it may be tempting to specify ultra thin mirrors and beamsplitters, doing so makes it impossible for optical manufacturers to guarantee surface flatness and thus image quality. Just holding the mirror can deform its shape. If you need surface flatness of ¼ wave or less, a good rule of thumb is to use a 6:1 ratio between surface size and thickness. If you have to specify thinner optics, take a great deal of care when mounting the parts to avoid deforming them due to strain in the mechanical fixtures or from bonding. One last point on mirrors: Mounting them from the front can alleviate the need for tight thickness tolerances.

Your system may be using infrared LEDs, which is great, but good IR mirrors take a little getting used to. They are often gold mirrors, which are soft and easily damaged. Have a talk about these issues with both your supplier and the production people before your design hits the manufacturing floor.

A wise designer allows for adjustments. Long optical paths can be very sensitive to centering, boresight and angular tolerances. Folding the optical path multiplies this problem by a factor of three. What works well in the lab may fail on the production floor. If possible, design gimble adjustments for all of your folds. An x-y adjustment in the CCD plane can help to adjust for boresight errors. If you can't do

continued >

Tricks of the trade: When debugging your illumination prototype, keep a flat mirror and chrome ball bearing on hand. These two surfaces accurately show the location and intensity of your illumination sources regardless of object surface characteristics.

that, do a stringent geometric tolerance analysis to minimize the effects listed above.

Don't ignore back reflections. When you use a beamplitter to combine the illumination and imaging optics into the same path, only 20 to 40% of the illumination is used. The rest of the light passes out of the system. But when the stray light hits a piece of metal in the machine and reflects back into the optical system, you get problematic back reflections. Even black surfaces can reflect light! Prevent this problem by baffling the excess light. Threaded barrels can really make a difference. Or, you can make your own light stop (see Tech Tip below).

Production

Finally, your design is nearing production. Although you started with off-the-shelf components, most systems end up with some custom components. Talk to your supplier: what should you expect in terms of price and delivery of custom components?

Some common customizations, which won't break your budget, include edging lenses (to a smaller diameter) and resizing mirrors and beamsplitters.

If you are considering ordering custom lenses, consider the quantity you need. If you need at least 500 pieces of a single lens or doublet, then custom lenses may make sense. If you need fewer pieces, off-the-shelf lenses will probably be more economical.

Determining whether to order custom or off-the-shelf compound CCD lenses is more complicated. If you need more than 250 pieces, a custom lens can really make sense. One of the advantages of a custom lens is that you can eliminate adjustable aperture stops and helical focus, which reduces costs dramatically. In addition, most off-the-shelf lenses turn out to be larger than if they were designed for a specific application. So if you are working in a tight space (which this article assumes), a custom lens can be a real advantage. If you have done your homework, a good optical company can make ordering a compound CCD lens painless and cost effective.

In addition to adding gimbles for adjusting mirrors, make the lenses adjustable too. The more leeway you can provide in terms of focus and alignment, the easier life on the production floor will be. At the very least, you must provide for focus adjustments!

Conclusion

While designing an optical system into the tight constraints of semiconductor equipment is rarely easy, we've provided both a strategy and some tactics that will make it possible.

TECH TIP ON MAKING A LIGHT STOP

Many optical applications call for a non-reflective light stop to be inserted into the system. A quick solution to this problem is to run the same piece of paper through the copy machine multiple times with the document cover open and nothing on the glass. This produces a black piece of paper. After several passes through, the buildup of toner will provide an excellent, inexpensive, non-reflective light stop.

Micro-Optics and Fiber Optic Systems

FIGURE 1: Micro-optics of various sizes and shapes are used in fiber-coupling and collimating applications.

Micro-optics are critical to fiber optic systems as they help connect the fiber to the components responsible for manipulating light. For this reason, an understanding of micro-optics is important to not only optical engineers but to all designers working with fiber optic systems so that they can specify the right micro-optics for their application.

Because network designers, manufacturers, and optical engineers think very differently, we need to define some terms before we can talk constructively across these disciplines. For example, what is throughput? Is it the bandwidth of data, the number of widgets that a production line makes in a day, or the intensity of light that passes through a lens? For the purposes of this article, we will use the terms of optical engineers – throughput is the intensity of light through a lens. Light intensity is an important factor because a lack of intensity creates a noisy or weak signal, which is not useful for carrying data.

Flavors of micro-optics

Micro-optics are lenses, mirrors, prisms, windows, and other elements, used to manipulate light, that have dimensions between 0.5 and 3 mm (see Figure 1). Among the host of lens types are PCX (with one planar and one convex side), DCX (with two convex sides), ball, drum, and gradient index (GRIN) lenses. The latter are popular because they can be made to guide light toward their axis, which can be very useful for guiding light into a fiber core. There is more to the micro-optics for fiber than GRIN lenses, however (see "Battle of the Lenses", next page).

Nearly all of these manipulative elements perform one of two basic functions: they either collimate light or couple light from one device to the next. Collimating optics catch and reshape the spreading beam that emerges from a laser diode. Coupling optics have more varied jobs: they are employed where the beam magnification needs to change, which is typical any time light moves between a fiber and another component (for example, multiplexers/demultiplexers, circulators, gratings, or switches).

Micro-optics are made of either glass or plastic. The most typical glass is BK7, and elements made from this material have standard characteristics. By using materials with higher refractive indexes, a lens with the same radius can have a shorter focal length and a higher numerical aperture. High-index materials include LASF9 and cubic zirconium.

Plastics, such as PMMA, SMMA, or polycarbonate, have considerably lower indices of refraction than BK7, but are used because they can be made easily by molding and are less expensive than glass. Furthermore, the molding process allows manufacturers to incorporate mechanical structures or aspheric surfaces into the elements.

Plastics, however, are difficult to coat, and coatings are essential for fiber applications. Standard vapor-deposition coating methods cannot be used on plastics because the materials cannot withstand the

II

continued >

temperatures. Plastics can be coated by dipping techniques, but such coatings are not as complex as those possible with vapor deposition. Coatings maximize throughput, reduce reflection, and filter stray light.

More complex coatings can filter for polarization as well; this function cannot be done on plastic. For collimators, however, where narrow passbands are not needed, plastic microlenses work well.

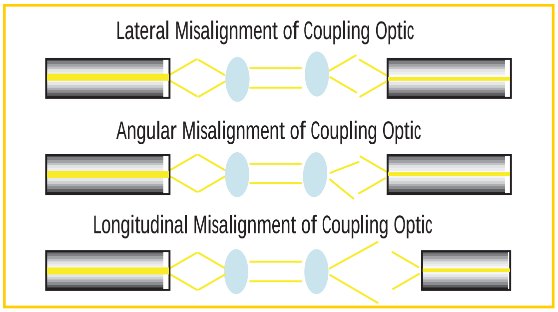


FIGURE 2: *Different types of misalignment in fiber couplers are caused by separation, offset, and tilt.*

Using micro-optics

Micro-optics increase throughput by fighting back-reflection and alignment errors. Reflections not only reduce efficiency but can cause feedback in the laser. For example, when coupling two fibers with plane faces or collimating light from a laser, one can reduce feedback by using discrete elements or coatings, or both. Antireflection coatings cut down on the amount of reflection at each surface.

Fiber coupling is subject to three types of misalignment (see Figure 2): separation, offset, and tilt. In separation, the fibers may not be close enough together: if there is an unplanned-for distance along the z-axis between them, light from one fiber core will spread out and lose much of its intensity. When offset, the fiber cores may be displaced laterally along the x-axis, so that light from one core hits the cladding layer of the second fiber, also reducing the light throughput. Finally, one fiber may be tilted (rotated around the x- and z-axis) so that the light will hit the cladding of the second fiber when launched.

The mechanical effects and tolerancing of the way the fibers are held certainly prompt alignment errors. Optical tolerances apply to mounting devices as well as to the optics; they share the total error budget. If the mounts are made of molded plastic, it is hard to hold tight tolerances to the mounts, and the optics have to be much more precise in order to stay within the budget.

The gross error in molded plastic housings is sizeable. If you have ever pulled the cover off of patch panels, you can see that the efficiency of typical connectors needs improvement. To compensate for mechanical errors, optical tolerances are driven very hard. A better solution would be to improve the accuracy of molded plastic connectors, or at least to evaluate where more improvement can be made to meet the requirement for an application.

With optical fiber, axial positioning tolerances make a big difference. If you have a 50µm fiber core, and the beam entering the fiber is decentered by 10 to 30µm, the system may still work, but it will lose throughput.

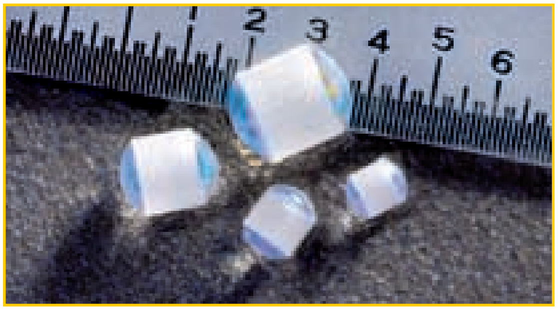


FIGURE 3: *Drum lenses can be used for many fiber applications, with fiber-coupling efficiencies comparable to GRIN lenses. Drum lenses are easily available and inexpensive.*

Battle of the lenses

Everyone who works with optical fiber seems to love gradient index (GRIN) lenses. They do have some excellent characteristics: flat faces and no spherical aberration on axis. They are expensive, however, and at times PCX or ball lenses can work just as well at a lower cost.

For a 50µm core fiber, one can use other types of lenses and get the same efficiency. Consider the following options:

- A drum lens, which is an edged-down ball lens (see Figure 3). Drum lenses are readily available and comparable to the per formance of GRIN lenses.
- A ball lens has the same effect as a drum lens, is compact, and the focal length is the diameter. Within the past year, ball lenses that are coated uniformly on all sides have been commercialized. The lens can be dropped into a system without worrying about an axis or uncoated region, and because of its shape, mounting is straightforward.
- PCX lenses, which are effectively half of a ball lens, work fine for

continued >

efficiency in a lot of moderate bandwidth systems, such as OC1 or less. They are available off-the-shelf, and their biggest drawback is spherical aberration. For a 50µm core fiber, however, one can achieve 90% to 95% efficiency with a PCX lens at lower bandwidths. By keeping track of numerical aperture effects within the system, systems designers can maintain that efficiency.

In the 1980s, an analysis of three types of lenses by A. Nicia showed them comparable in fiber-coupling efficiency.[1] In experiments, ball lenses were shown to reach the theoretical expectations. The efficiency of these devices really depends on the packaging.

Tolerancing

What about tolerancing in the micro-optics themselves? Glass tolerancing is well understood and built into popular optical design programs. Plastics are not as well characterized. After being molded, plastics shrink tremendously – for example, to get a 7mm diameter element, a 10mm mold may be needed. This limits the feasible tolerances of plastic parts and certainly includes a different set of issues than for glass. Plastic's material qualities also limit surface accuracy and centering (see table at left).

TYPICAL TOLERANCES FOR MICRO-OPTICS

Tolerance	Glass	Plastic
CT to ET ratio	2:1	4:1
Surface Quality	20-10	40-20
Dimensional	0.03mm	0.015mm
Centering	20min	20min
EFL	2%	0.50%
Power	2 fringes	2 fringes
Irregularity	1/2 fringes	1/2 fringes

The issues apply to elements other than lenses. Windows have many of the same tolerancing and coating issues. Microprisms are manufactured and coated differently: special tools must be created to grind, polish, and coat the right-angle prisms used for switching. To some extent, the tolerances depend on the tooling.

As the size of the optic shrinks from macroscopic to micro-sizes, two conflicting tolerancing issues occur. First, the error budget is smaller because the system and its components are smaller. But small size also works for micro-optics. For example, consider a wedge: the total indicator runout for this part is the product of the angle times the diameter. Because the diameter is so small, the wedge is less sensitive to errors in the angle. Tolerancing issues do not translate directly from macro-optics.

Consider surface roughness. Imagine that a lens is specified to have a quarter-wave roughness. This specification means that at some points on the surface, the roughness may be as large as a quarter wave. As the aperture gets smaller, the smaller area is more likely to be within spec because it is less likely to contain one of the roughest spots. The probability of having a deviation within the field is smaller, for a smaller field. For a micro-optic, one might achieve the same quality by specifying only a half-wave surface roughness.

The scratch-dig surface requirements work the same way. Specifying a quality of 20-10 should not change the price. As the diameter decreases, so does the difficulty of holding quality over that smaller area. Obtaining a surface quality of 10-5 is more difficult.

Conclusion

For your application, consider what kind of optics and coatings you need to get the performance you want. As the sidebar suggests, PCX lenses are inexpensive and readily available in many diameters and can solve many of the same problems as GRIN lenses.

When you specify the elements, pay some attention to the tolerancing: if you can inject some intelligence into the specifications to make them fit your application, you may bypass some expensive manufacturing problems that are not strictly necessary. Although the size of micro-optics suggests that they be more precise than macro-optics, the reality is that the tooling is more difficult and tighter tolerances may not be necessary.

REFERENCES

[1] A. Nicia, Lens coupling in fiber-optic devices: efficiency limits, Appl. Opt. 20 (18), p. 3136-3145 (15 Sep 1981).

Contrast Enhancement Through Filtering

FIGURE 1: *An application where multi-colored pills are sent down a conveyor belt and sorted via an imaging system.*

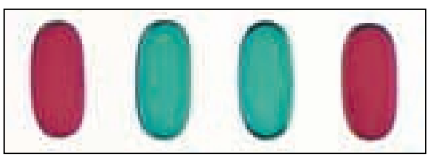


FIGURE 2: *A close-up of the colored pills (above) and the grayscale values of those pills using various filters (below).*

Application requirements

During packaging, pharmaceutical pills of different colors need to be sorted. An automated imaging system, which distinguishes between the different colored pills, is essential in increasing production efficiency. In such a system, pills are inspected for specific characteristics as they travel down a trough-like conveyor belt prior to sorting. A minimum of 60% contrast is needed for the software to be able to differentiate between the different pills.

System Requirements Given by Customer:

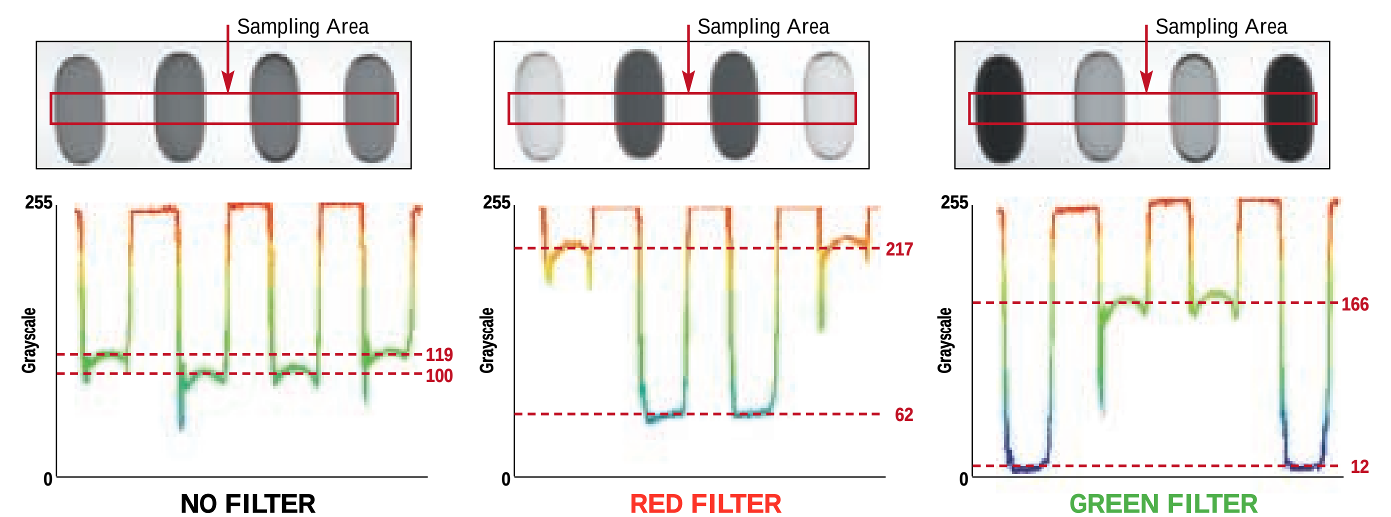
Working Distance:	~350-450mm
Field of View:	~70mm
Minimum Contrast:	60%

Component selection

The 35mm MVO™ Double Gauss imaging lens, used with a ½" CCD format camera, yields an appropriate field of view and working distance for this application. The Sony XC-ST50 high resolution monochrome CCD camera offers a suitable amount of resolution and dynamic range (grayscales). A fiber optic area backlight is placed underneath the slotted trough to diffusely illuminate the pills. A capture board is used to digitize the camera signal for further image processing. In order to meet the minimum contrast level, filtering is required. The process of the filter selection is shown below. 14

Effects of filtering

Monochrome cameras cannot inherently discriminate between different colors. In this example, both the red and green pills appear nearly identical when imaged with the Sony XC-ST50 (see Figure 2). Filtering can



continued >

$$\% \text{ Contrast} = \frac{(I_{max} - I_{min})}{(I_{max} + I_{min})}$$

where I_{max} is the maximum intensity and I_{min} is the minimum intensity

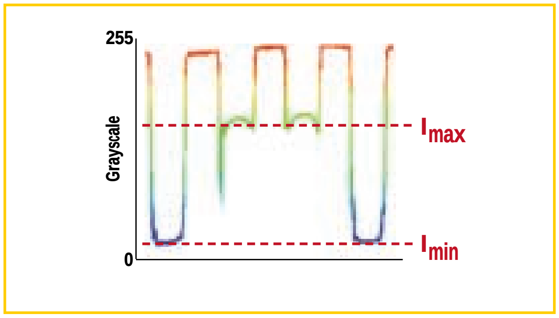


FIGURE 3: An example on calculating contrast using the equation given above.

be used to improve the contrast between pills of different colors and enables the system to differentiate between them. The images, along with their associated grayscale profile curves, are illustrated in Figure 2. All curves are generated only for the sampling area indicated.

Calculating contrast

A visual interpretation of the images and grayscale profile curves can be quite subjective. However, a contrast value can be calculated from the curves to determine which filter offers the highest contrast (see Figure 3).

NO FILTER: $\text{CONTRAST} = \frac{119 - 100}{119 + 100} = 8.7\%$

RED FILTER: $\text{CONTRAST} = \frac{217 - 62}{217 + 62} = 55.6\%$

GREEN FILTER: $\text{CONTRAST} = \frac{166 - 12}{166 + 12} = 86.5\%$

Conclusion

In order to differentiate between the colored pills, the software needs a minimum of 60% contrast. A grayscale profile can be generated from the sample area in order to calculate the contrast. The original monochrome image only has a 8.7% contrast difference between the red and green pills. The contrast can be increased beyond the minimum requirement by ~25% by attaching a green filter to the front of the lens. This allows the user's customized software to operate on a go/no-go principle and accurately sort the pills.

TECH TIP ON USING FILTERS

For best results, point the coated, or "mirror-like" surface towards your light source. This will minimize any thermal effects resulting from the absorption of the heat by the glass on the other side. Placing the filter in the opposite direction will still work, though it will cut down your throughput and you will not get the maximum desired effect. Also, having the "mirror-like" side facing away from the source will cause an interference pattern when the source is a coherent beam of light. The coated surface is easily determined by looking at the edge of the substrate, from the direction of the center of the filter at a slight angle so looking at the inside edge. If you can see the actual edge (thickness) of the glass, then the coating is on the other side. From the coated side, the edge is not visible. This is more difficult to check on coatings that transmit in the visible, but the edge can still be detected by viewing the filter at a steep angle.

Also, be aware of the tilt of the filter. For filters in general, as the angle of incidence (the angle your source light hits the filter) increases, a filter's transmission curve will shift to lower wavelengths. The effect of large angles from the center of the optical system is the same as tilting a filter from a perpendicular position to an optical system. As the angle of tilt gets larger, the curve will start to change shape, this typically means the transmission will steadily drop and the slopes in the curve will start to change. Most filters are designed for a 0º angle of incidence, but some filters (such as Hot Mirrors) can be designed for other angles of incidence. Keep this in mind when specifying a filter.

Correcting Perspective Errors with Telecentricity

FIGURE 1: *An application where computer key connectors need to be analyzed in off-line inspection.*

SYSTEM PARAMETER EQUATIONS

Equation 1.0:

$$\text{CCD Res. } (\mu m) = 2 \times \text{CCD Pixel Size (Horiz, } \mu m)$$

Equation 2.0:

$$\text{Primary Mag. (PMAG)} = \frac{\text{Sensor Size (Horiz, mm)}}{\text{FOV (Horiz, mm)}}$$

Equation 3.0:

$$\text{Object Res. } (\mu m) = \frac{\text{CCD Resolution } (\mu m)}{\text{PMAG}}$$

combining all of these expressions with the given values yields:

Equation 4.0:

$$\text{Object Res. } (\mu m) = \frac{2 \times \text{Pixel Size } (\mu m) \times \text{FOV (mm)}}{\text{Sensor Size (mm)}}$$

Equation 5.0 (this example only):

$$\textbf{Pixel Size } (\mu m) \textbf{ / Sensor Size (mm) = 1}$$

Application requirements

In this example, a system is required to inspect the prototype of a hardware computer key connector to verify the placement of its pins. This is a laboratory setup requiring no automation. A precise measurement between each pin is determined using measurement software.

System Requirements Given by Customer:

Expected Vertical Pin Separation (center-to-center):	~2.5mm
Number of Pins Viewable Simultaneously:	~7
Object Resolution to Meet Measurement Accuracy:	36µm

System parameter calculations

In order to accommodate the simultaneous inspection of multiple pins, the minimum Field of View (FOV) should be about 18mm.

By using some basic equations (see left), we can specify the parameters of our system and pick a suitable CCD camera.

Our system requirements dictate an 18mm field of view and a 36µm object resolution. Using these values, Eqn. 4.0 can be reduced to a ratio (Eqn. 5.0).

This ratio can be used to compare the resolution of different cameras for a specific field of view (while factoring in the sensor size). We can calculate this ratio for some of our high resolution digital and analog CCD monochrome cameras:

Redlake MEGAPLUS ES 1.0	9.0/9.07= 0.99
Sony XC-ST30	6.4/4.8 = 1.3
Sony XC-ST50	8.4/6.4 = 1.3

The Redlake MEGAPLUS ES 1.0 camera is the best match to the desired ratio. Using Eqn. 2.0, we calculate PMAG = 0.5X for the imaging lens. And from Eqn. 1.0, the camera's resolution is 18µm. If we assume that the lens is not the limiting factor in the resolution of the system, the corresponding object resolution is 36µm (Eqn. 3.0).

Note: Although the Sony XC-ST30 camera has a higher resolution (13µm) it only yields about 50µm object resolution because it has a smaller sensor.

Component selection

Since the camera has been selected, a 0.5X PMAG imaging lens needs to be decided on. Conventional lens designs suffer from perspective errors which are noticeable when imaging objects with significant height/depth, as in this example. Telecentric lenses optically correct this problem as illustrated in the images on the next page.

continued >

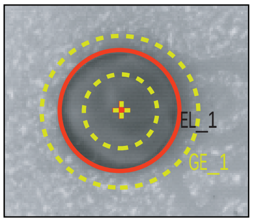


Image 2: Telecentric

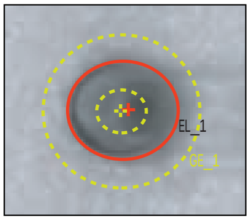


Image 4: Conventional

TELECENTRIC LENS

Image 1: High DOF (F16)

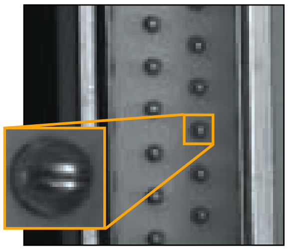
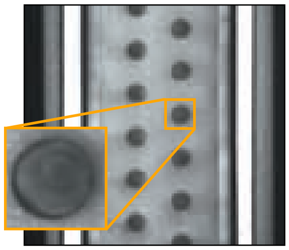


Image 2: Low DOF (F2.8)



CONVENTIONAL LENS

Image 3: High DOF (F16)

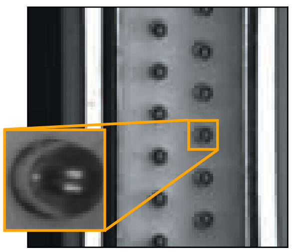
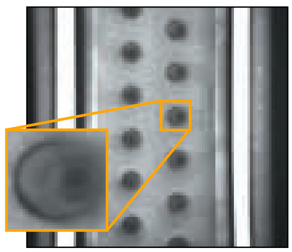


Image 4: Low DOF (F2.8)



Measurement software

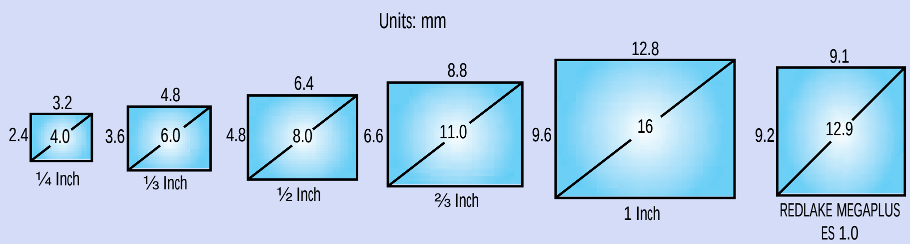
Edge detection analysis at low depth of field (F2.8) was used to determine the center of the pins. The telecentric design (Image 2) maintains a symmetrical blurring within the pin diameter. The result is an accurate circular fit to the pin by the measurement software. On the other hand, the conventional design results in a perspective blur which yields an elliptical fit. This introduces error into the prediction of the pin center and also other measurements. For example, the conventional system measures 3.51mm center-to-center separation between two diagonally adjacent pins. The telecentric system measures a 3.21mm separation. The actual pin separation is 3.16mm.

Conclusion

The Redlake MEGAPLUS camera, model ES 1.0 digital camera offers the best combination of high resolution and sensor size to meet the measurement accuracy requirement for this application. The telecentric lens was selected because it corrects for the perspective errors by maintaining constant magnification over the depth of field. Since the center of the object does not shift as it blurs, the telecentric lens offers a huge advantage when measuring center-to-center separation. In this example, the telecentric lens, in combination with the Redlake MEGAPLUS camera, improves the overall measurement accuracy by 86%.

TECH TIP ON CCD SENSOR SIZE

The size of the sensor's active area is important in determining the system's field of view. Given a fixed primary magnification (determined by the lens), larger sensors yield greater fields-of-view. The nomenclature of these standards date back to the Vidicon vacuum tubes used for television, so it is important to note that the actual dimensions of the chips differ. All of these standards maintain a 4:3 (horizontal:vertical) aspect ratio.



Another issue is the ability of the lens to support certain CCD chip sizes. If the chip is too large for the lens design, the resulting image may appear to fade away and degrade towards the edges because of vignetting (extinction of rays which pass through the outer edge of the lens). This is commonly referred to as the "tunnel" effect, since the edges of the field become dark. Smaller chip sizes do not yield such problems.

Manipulating Distortion Out of Your Image

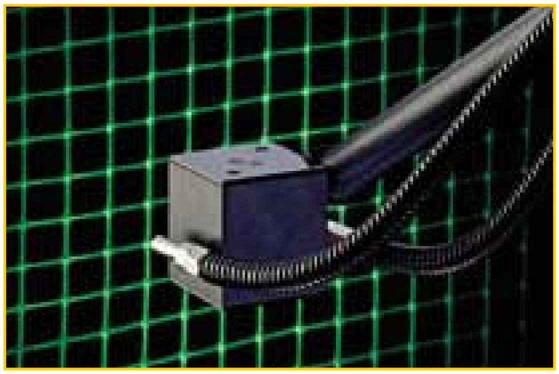


FIGURE 1: An application where plastic mesh fencing is measured to ensure all dimensions meed standards.

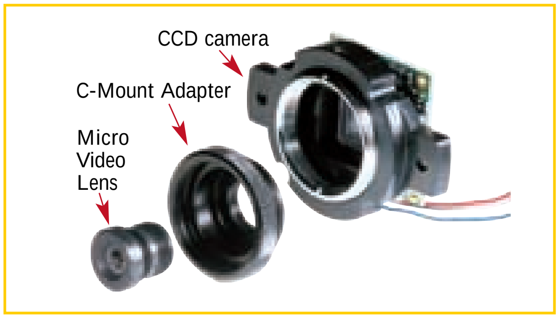


FIGURE 2: Assembly of CCD camera used in the above inspection example.

$$\text{Distortion (\%)} = \frac{\text{Actual Distance (AD) - Predicted Distance (PD)}}{\text{Predicted Distance (PD)}} \times 100$$

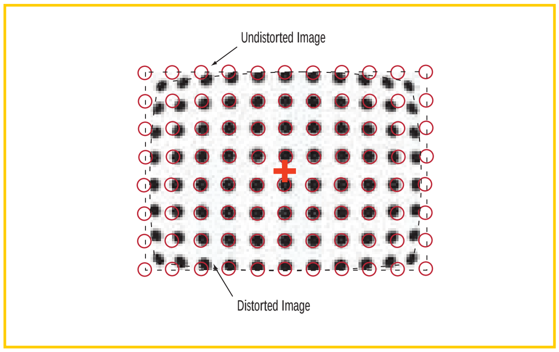


FIGURE 3: Calculating the Actual Distance (AD) and Predicted Distance (PD) for barrel distortion.

Application requirements

Precise measurements of plastic mesh fencing are needed during production runs to ensure that all dimensions fall within the specified tolerances. In this situation, the space reserved for the imaging system is extremely limited. The housing for the CCD and lens is integrated into the mounting of the machinery.

System Requirements Given by Customer:

Working Distance:	~50mm
Horizontal Field of View:	~50mm
Measurement Tolerance:	±0.3mm
Component Housing:	<40mm cube

Component selection

Although their minimum working distance is longer than desired, the MVO™ Micro Video lenses are compact, making them ideal for this application. The 4.3mm focal length MVO™ Micro Video lens has a 60° angular field of view under normal conditions. By introducing 0.25mm of space between the lens and the camera, the horizontal field of view is reduced to 50mm at a 50mm working distance. A high resolution monochrome board camera offers the appropriate resolution and size. The illumination is provided by a fiber optic illuminator with a dual branch flexible light guide. Due to the macro configuration and wide angle of the lens, distortion has been introduced into the image. This distortion must be taken into account in order to make accurate measurements.

Calculating distortion

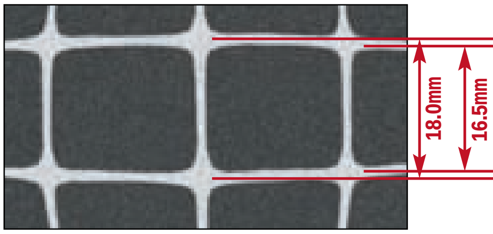
Distortion is a geometric optical error (aberration) in which information about the object is misplaced in the image, but not actually lost. Using measurement software and a dot target of known size (shown left), we can measure the distortion at different distances from the center of the image. Note: Distortion is not linearly correlated to the distance from the center of the field.

Factoring distortion out

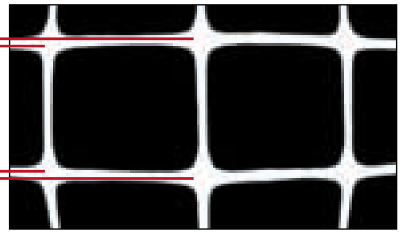
Once the amount of distortion is calculated, it can be factored out in order to yield an undistorted image. In this example, –16% (barrel) distortion is measured at the edges of the field. The distortion has a negative value because the edge of the field is closer to the center of the image than it should be. Since we are using a ⅓" format CCD camera (6mm diagonal sensor size), the corner of the sensor is 3mm from the center. Based on the amount of distortion, this point would actually be located at a distance of 3.57mm in an undistorted image. Since distortion must be measured for each point on the image, repeated calculations are required. Once this is done, distortion can either be processed out of the image (as shown on next page) or taken

continued >

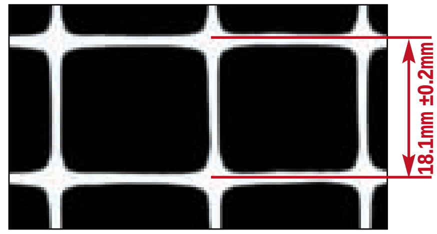
into account during measurement. We can also calculate the corresponding positions on the object by dividing the image distances by the primary magnification (PMAG=0.096). The edge of the field of view, the part of the mesh measured to be 31.3mm (=3/0.096) from the center mark, is actually 37.2mm (=3.57/0.096) away.



DISTORTED IMAGE: This is an initial distorted captured image in which the contrast is not ideal.



BINARY IMAGE: A binary image (black and white, no grays) can be generated through image processing. Note: It is not necessary to convert the image into binary to subtract distortion.



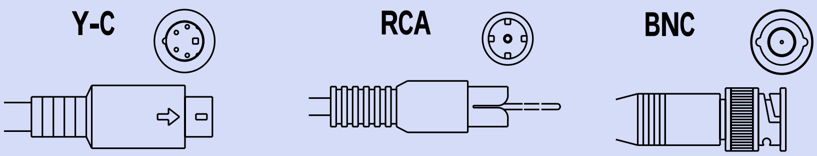
CORRECTED IMAGE: Having measured the distortion accurately, it can be removed through image manipulations. The resulting image is a precise representation of the original object.

Conclusion

The MVO™ Micro Video lens and board level camera offer an ideal solution for this space-limited application. There is a high degree of distortion within the lens because of its large angular field of view. Once measured, this distortion can be factored out of the image in order to obtain more accurate measurements. In this example, we are interested in measuring the height of the mesh (center row). Without taking distortion into consideration, the height fluctuates from 16.5 to 18.0mm. Once the distortion is taken into account, we realize that the range of the height of the mesh is actually 17.9-18.3mm, well within the ±0.3mm tolerance.

TECH TIP ON SIGNAL FORMATS

There are four basic signal types used in CCD cameras: Composite (NTSC, EIA), Y-C (S-video), RGB and Digital (RS-422). NTSC (RS-170A/Color) and EIA (RS-170/Monochrome) signals are the most common and will accommodate most applications. Y-C and RGB separate the image into components and therefore provide superior image quality for video recording and image analysis. Digital cameras provide a level of performance that make them unique. Used in conjunction with image capture boards, digital cameras do not suffer from the visual constraints imposed by video formats. The result is greater flexibility in image acquisition and quality. In any electronic system, the signal format should be constant. Any accessories added to the common camera-monitor system are added directly after the camera.



Each video signal format corresponds to specific cable connectors, as shown above. Composite signals can use either BNC or RCA type connectors. Y-C uses four pin-DIN type and RGB uses four BNC connectors.

Off-the-Shelf Optics Offer Speed and Economy

FIGURE 1: Stock lenses have a wide variety of popular diameters, each with a wide variety of focal lengths, to provide customers with many choices to fit their application.

FIGURE 2: Many elements are blocked onto one tool and ground and polished at the same time.

Off-the-shelf optics are usually much less expensive and easier to use than custom optics, though not in every case. This article takes a look at when off-the-shelf optics should be used, and how to use them. Off-the-shelf optics are continually produced in large quantities, and kept in stock by manufacturers and distributors. These stock optics are typically designed in a wide variety of sizes and focal lengths from which to choose (see Figure 1).

Customers frequently struggle with the decision of when to use catalog optics and when to buy lenses custom made for their application. As a manufacturer, we have often been asked to quote prices for custom optics in volumes at which they are not as economical as off-the-shelf elements. Because many different customers use the same lens, off-the-shelf optics allow an economy of scale, even when one customer needs only a few lenses. As a general rule of thumb, custom lenses make economic sense only when one needs thousands of lenses. But, as with any rule of thumb, there are always exceptions.

Off-the-shelf advantages

For many reasons, off-the-shelf components are more economical than custom components. The first and most obvious is that economy of scale can be gained by using off-the-shelf. To understand why volume is important, one must understand how most lenses are made.

The vast majority of lenses are produced the same way today as they were made during World War II. This involves blocking many lenses onto one tool and grinding and polishing with pitch (see Figure 2). Several tools can grind or polish at the same time on a single machine. Making one lens takes as long as making several hundred. And because material often is a small portion of the cost of manufacturing common glass lenses, making one lens costs about the same as making 50.

Deterministic grinding and polishing machines can be used to manufacture lenses one at a time. These machines have their own associated expenses. This method is usually used for low volume manufacturing. Tooling is more of a consideration with deterministic polishing. This is still fairly new technology and not as prevalent as the conventional pitch polishing.

Once the lenses are polished, they must be tested. Test plates are typically used to test a lens. If the lens has a radius that is not currently being used in the optical shop, then the costs of manufacturing the lens increases because a specific test plate must be made for each radius used. Special tooling may also be necessary for custom lenses, further increasing cost.

In addition to volume, off-the-shelf optics are (by definition) available more quickly than custom lenses. A common request from customers who do not understand that lenses are made in batches is to receive just the first few lenses that are made, assuming this will save

continued >

time. The first few custom lenses in a batch might be available a day earlier, due to the time taken for testing. Manufacturing a simple lens can take on the order of one to three months. If test plates must be manufactured as well, then expect to add another month or two. For quick lead times, off-the-shelf optics cannot be beaten.

Lead time is an obvious consideration for prototypes, but also needs to be understood for production. Sometimes a customer using a custom lens suddenly has a dramatic increase in business and needs to have twice as many lenses as they forecasted. If this sudden increase in demand cannot be filled, it could shut down their assembly line. Although some manufacturers try to keep a safety stock of lenses, such stocks are more difficult to maintain for custom lenses than for off-the-shelf lenses.

Off-the-shelf disadvantages

Off-the-shelf optics, however, also have disadvantages that should be considered. Customers often want to buy off-the-shelf optics to insert in their own optical designs. Ideally, off-the-shelf optics should be incorporated in the initial design. Altering a finished design can be costly.

Changing the lens inevitably means that the mounting must be changed to accommodate any changes in focus. Even lenses with identical focal lengths can mount differently because a change in radius alters where the lens is mounted. Also, the optical design must be redone. Engineering for these kinds of changes will have associated costs that may outweigh the savings of an off-the-shelf lens. If FDA or similar approvals are required, the validation involved in a design change also can lead to severe costs.

Another consideration is the effect on tolerances. If one uses more elements to correct aberrations without using custom optics, then the stack-up of tolerances can decrease performance. Also, some designs require a specific tolerance for a specific element, which may not be standard to off-the-shelf optics. Sometimes a very specific focal length is required, or a specific lens form, such as a meniscus lens, to correct aberrations — these may not be available off the shelf. Special coatings are a popular reason for a custom lens. Sometimes designs require very low reflectance at a specific wavelength or an antireflection coating in the UV or near-IR wavelengths. Sometimes there is no off-the-shelf solution and a custom lens is unavoidable.

Options are available to customize off-the-shelf elements, including edging down a lens or a custom coating. Edging down, or changing the diameter of a stock lens, can be done quickly and often inexpensively. This is useful for mounting in an existing housing or accommodating space limitations. Another simple customization is applying a special coating on an uncoated stock lens. The cost for custom coating a batch of lenses can be quite low, and the lead time very short.

The sooner off-the-shelf options are worked in, the better.

Designing in-stock lenses

Designing with off-the-shelf optics can be made easier in several ways. The first is to design using these elements. Most design software packages have off-the-shelf lenses preloaded into them. Software such as Zemax, Code V, Oslo, OLIVE, and others all include complete catalogs of off-the-shelf lenses.

The sooner off-the-shelf options are worked in, the better. Typical design software will give a starting point with custom lenses when one optimizes all surfaces. Then one can force the software to replace the custom lenses with the closest off-the-shelf matches and allow air

continued >

spaces to compensate. The best time to do this is before starting on the mechanical design.

Many design tricks can be used to produce superior performance with off-the-shelf components. Consider a laser objective, which is used to focus a laser beam to a small spot. The lens system can be designed in two ways. A custom solution would use a "best form lens" — a single lens with two different curvatures selected to reduce spherical aberrations.

An off-the-shelf solution would use two identical PCX (plano-convex) lenses in place of the single, more expensive custom lens (see Figure 3). Two 30-mm focal length PCX lenses close together can, for example, replace a best-form lens with a 15.5-mm focal length. The two PCX lenses yield a smaller spot (around six times smaller), because four surfaces bend the light instead of only two. Two PCX lenses also can be cheaper than a single best form, because only one radius is being manufactured. The radii of the lenses in the off-the-shelf design are also longer, which often will be cheaper as well because it is easier to manufacture.

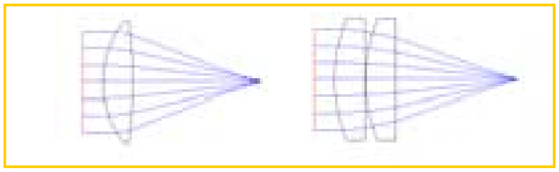


FIGURE 3: *Two stock lenses can replace a single custom lens and provide good performance for less money. Design software is Zemax.*

Volume makes the difference

Volume is the key to deciding when to use off-the-shelf or custom optics. Low volume will always favor off-the-shelf elements, but as volume goes up the advantages diminish and other factors take over.

Custom is almost always out of the question for prototypes and proof of concept. When one needs a single or a small run of prototypes, off-the-shelf elements should be used whenever possible. In these quantities, custom lenses are astronomically expensive and require long lead times. Deterministic grinding and polishing can make this more cost effective, but also will lead to high prices. Off-the-shelf lenses provide a major benefit by being available quickly, because speed is critical for most prototypes. Also, if the prototype shows that the design must be changed, the customer must repeat the one-time expenses associated with custom lenses.

When a low volume of between 100 and 1,000 pieces is needed, economy of scale still causes off-the-shelf to be the more economical option. At this volume, off-the-shelf optics can save the customer from having to commit to a supply for a year or more. Stock lenses offer the considerable advantage of allowing the customer to buy on demand and have stock available in case volume increases. However, if a custom solution is necessary it can be done at a reasonable cost at this volume.

For moderate volumes of 1,000 to 100,000 pieces, both custom and off-the-shelf elements are viable options. Lenses are generally not stocked in these volumes unless a need for them is forecasted for a specific customer. Increases in volume are still easier to accommodate with an off-the-shelf option, because there is less risk in overstocking a stock lens than in overstocking a custom lens. The savings that might occur by using custom lenses start to be important at these volumes.

For high volumes above 100,000 units, custom lenses are almost always used. If elements can be eliminated, custom is almost exclusively used. This volume provides the customer with economies of scale for the custom lenses. The cost per piece to manufacture 200,000 pieces is not significantly less than the cost per piece of 100,000 pieces.

Saving time with stock lenses

Here's an example of the design process that demonstrates many of

continued >

the trade-offs between stock and custom lenses. The initial request to Edmund Industrial Optics was to redesign an imaging lens system for use in iris identification. This large optical system was reduced to a package about the size of a baseball.

The redesign used custom lenses, a stock PCX (plano-convex) lens and two custom best-form lenses. The system required monochromatic near-IR illumination. The working distance of the lens system needed to be short and ideally would use as few elements as possible. Due to a large production volume, a custom design was chosen to improve image quality.

A prototype run of about 50 pieces was manufactured. This was very expensive, but necessary for proof of design. OptoTech lens grinders and polishers made the prototypes more quickly and less expensively than traditional pitch polishing. Tooling for four separate radii was necessary, increasing both the cost and lead time.

After testing the prototypes, the customer changed the design specification radically. The second redesign used off-the-shelf optics to reduce lead time. In the first prototype run, the centering tolerance on the two PCXs had driven up the cost of the metal housing. A cemented achromat doublet greatly eased the tolerances. The second prototype was manufactured using items that were in stock. C-Mount tubes were used as the mounting platform.

The total redesign took about one week from specification to final prototype. The costs of the achromat were less that the cost of the two best-form lenses, and the cost of the housing was reduced. The customer approved the off-the-shelf solution, even for volumes up to 10,000.

TECH TIP ON CLEANING LENSES

Dust is the most common contaminant and can usually be removed using pressurized gas. If more cleaning is necessary, hold the lens in lens tissue and apply a few drops of reagent-grade acetone or lens cleaning solution. Slowly turn the lens while applying pressure in the center and working outward, to pull dirt off the lens instead of redistributing it on the surface. Fingerprints on a coated lens should be cleaned as soon as possible to avoid staining or damaging the optic. Larger dirt particles, however, should be removed with a dust-free blower before attempting to clean the optic with lens tissue. Larger particles trapped under the cloth will scratch the surface you are attempting to clean. If the lens is still dirty after using acetone — for instance, if the oil was just redistributed and not cleaned off the optic — then a mild soap solution can be used to gently wash the lens. Repeat the procedure with acetone to eliminate streaks and soap residue.

Micro-optics may also be cleaned using acetone but, due to their extremely small size, they require special handling and care. Delicate tweezers may be used to securely hold a micro-lens by its edge, or a vacuum pick-up tool may be used.

Also, choosing the proper cleaning supplies and using the proper techniques are as important as cleaning the component itself. Using improper cleaning practices can damage polished surfaces or specialized coatings that have been used on a substrate or lens. Always check with the manufacturer of the component to determine proper care and cleaning procedures.

Optics and Machine Vision

FIGURE 1: The fundamental parameters of an imaging system include the resolution of the object, the field of view, and the depth of field the user wishes to image. The sensor size and the working distance from the object to the lens are also important. The primary magnification is the field of view divided by the sensor size.

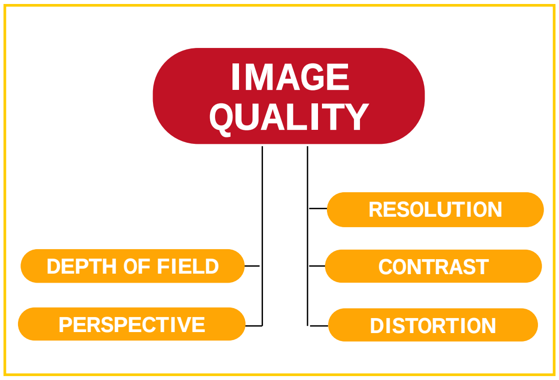


FIGURE 2: A variety of factors contribute to the overall image quality, including resolution, image contrast, depth of field, perspective errors, and geometric errors.

Image quality in a machine vision system is determined primarily by the quality of the system's components, such as lenses and frame grabbers. And, image quality can be measured and specified fairly easily. You can build a machine vision system by trial and error — by picking lenses, a CCD, and electronics at random and hoping they will work together and provide an image quality sufficient for your application. Many prototype systems are, in fact, built this way and many require considerable troubleshooting to get them working.

There is a better, faster way to build imaging systems that often can yield a cheaper system than you get by guesswork. By starting with an understanding of image quality, you can choose components that fit the application and complement one another. And none of the components will be more expensive than necessary.

The first step is to understand how image quality is specified. Next, by considering the relationship between resolution and contrast you will understand the tremendously useful modulation transfer function. Third, you take into account other factors related to image quality, including the relationship between f-number and resolution; the diffraction limit; and aberrations, depth of field, distortion, and perspective error.

Fundamentals

The fundamental parameters of imaging systems (see Figure 1) include:

- Field of View (FOV). The viewable area of the object under inspection, i.e., the portion of the object that fills the camera's sensor.
- Working Distance. The distance from the front of the lens to the object under inspection.
- Resolution. The minimum feature size of the object under inspection.
- Depth of Field (DOF). The maximum object depth that can be maintained entirely in focus. The DOF is also the amount of object movement (into and out of focus) allowable while maintaining an acceptable focus.
- Sensor Size. The size of a camera sensor's active area, typically specified in the horizontal dimension. This parameter is important in determining the proper lens magnification required to obtain a desired field of view.

Another useful descriptor of the system, the primary magnification of the lens, is the ratio between the sensor size and the field of view. It is not a fundamental parameter:

magnification = sensor size (mm) / field of view (mm)

In addition to resolution and depth of field, image quality is also a combination of three other properties: image contrast, perspective errors, and distortion (see Figure 2).

The primary purpose of any imaging system is to obtain enough image quality to allow the extraction of necessary information. A system that works for one application might not for another.

continued >

Furthermore, the use of over-specified components might do little more than increase system cost. With this in mind, let's take a closer look at the components that determine image quality.

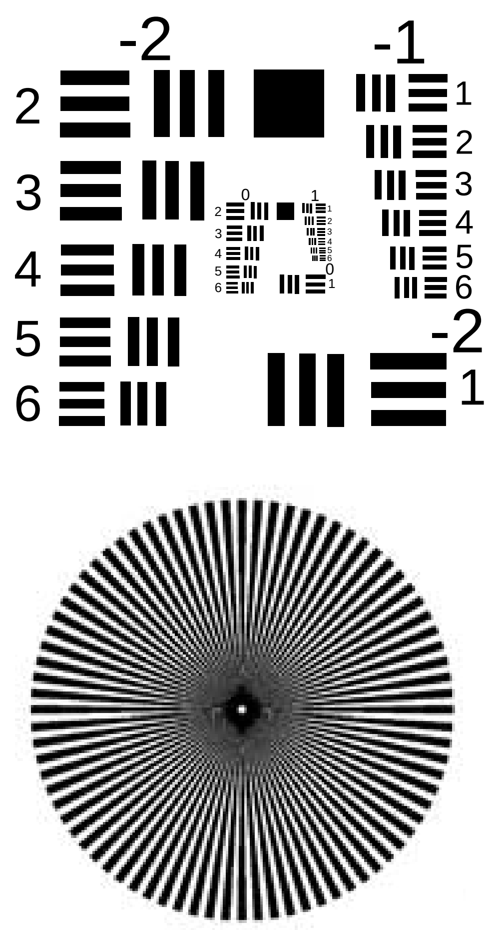


FIGURE 3: Two test targets: a bar target and a star target allow users to measure resolution and astigmatic errors. The orthogonal lines in the bar and the radial pattern in the star allow users to test the system for astigmatic errors. The star target's wedges have continuous frequencies that can be calculated by radial distance, unlike the finite number of steps in frequency offered by the bar target.

Resolution

Resolution is a measurement of the imaging system's ability to reproduce object detail. Imagine, for example, a pair of black squares on a white background. If the squares are imaged onto neighboring pixels, they appear to be one large black rectangle. To distinguish one from the other, a certain amount of space must exist between them. Determining the minimum necessary space yields the limiting resolution of the system. This relationship between alternating black and white squares is often described as a line pair. The resolution is typically defined by the frequency measured in line pairs per millimeter (lp/mm).

Two different but related resolutions are in play here: the resolution in object space (the size of elements in the object that can be resolved) and image space resolution (a combination of lens resolution and camera resolution). The sensor's line pair resolution can be no more than half the number of pixels on the sensor because a minimum of two pixels is required to discern a black and white area. The image and object space resolutions (described in lp/mm) are related by the primary magnification:

$$\text{image space resolution} = (\text{object space resolution}) / (\text{primary magnification})$$

The limiting resolution of the system can often be found by imaging a test target (see Figure 3). A bar target consists of line pairs with varying frequencies; a star target consists of wedges with a continuum of frequencies. The orthogonal lines in a bar target are useful because they allow an operator to test the system for astigmatic errors, which are errors that show up differently in the X and Y planes of an image. Bar targets, however, are limited by having a finite number of steps in frequency. Star targets do not have this drawback but can be more difficult to interpret.

Contrast

Although the resolution and the contrast of an image can be defined individually, they are closely related. We have already examined resolution as an independent parameter that describes object detail. Let's now consider contrast independently before relating the two concepts.

Contrast, which describes how effectively the differences between boundary areas on the image are reproduced relative to one another, can often be defined in terms of grayscale or signal-to-noise. For an image to appear well defined, the black details must appear black and the white details, white (see Figure 4). The greater the difference in intensity between a light and a dark line, the better the contrast. This is intuitively obvious, but more important than might first appear. The contrast is the separation in intensity between blacks and whites:

$$\%\ \text{contrast} = (I_{max} - I_{min}) / (I_{max} + I_{min})$$

Reproducing object contrast is as important as reproducing object detail, which is essentially resolution. The lens, sensor, and illumination all play key roles in determining the resulting image contrast. The lens contrast is typically defined in terms of the percentage of the object contrast that is reproduced. A sensor's ability to reproduce contrast is usually specified in terms of decibels in analog cameras and bits in digital cameras.

continued >

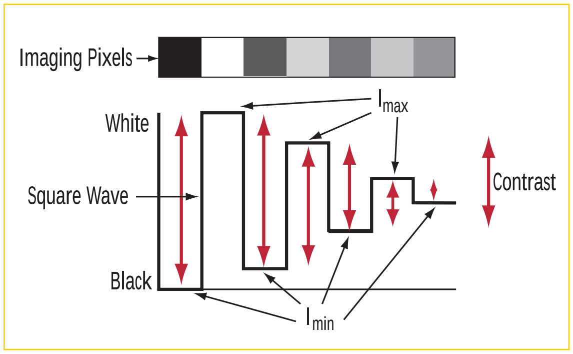


FIGURE 4: Contrast is the difference in intensity between blacks and whites. For an image to appear well-defined, black details must appear black and white details must appear white. The greater the difference in intensity between a black and white line, the better the contrast. The human eye can see a contrast of as little as 1-2%. A typical limiting contrast of 10 to 20% is often used to define the resolution of a CCD imaging system.

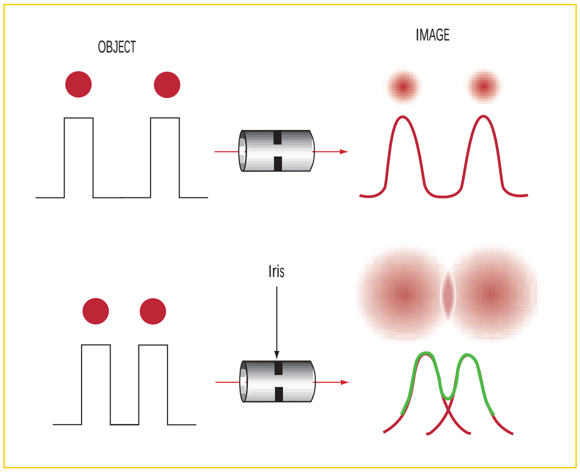


FIGURE 5: Contrast is not constant; it depends on frequency. The dots at the top of the figure can be imaged through a lens. They blur slightly. If we moved the spots closer, their blurs overlap and contrast decreases. When the spots are close enough that the contrast becomes limiting, that spacing is our resolution.

Linking resolution and contrast

Resolution and contrast are closely linked. In fact, resolution is often meaningless unless defined at a specific contrast. Similarly, contrast depends on resolution frequency. Consider two dots placed close to each other and imaged through a lens (see Figure 5). Because of the nature of light, even a perfectly designed and manufactured lens cannot accurately reproduce an object's detail and contrast. At best, if the lens is operating at the diffraction limit (which we will discuss later), the edges of the dots will be blurred in the image.

When the dots are far apart (i.e., at a low frequency), they are distinct; as they approach one another, the blurs overlap until the dots can no longer be distinguished. The resolution depends on the imaging system's ability to detect the space separating one dot from another. System resolution therefore depends on the blur caused by diffraction and other optical errors, the dot spacing, and the system's ability to detect contrast.

Because contrast and resolution are so closely related, it is often beneficial to specify a contrast level at a specific resolution. The result of a range of frequencies being measured is the modulation transfer function (MTF) curve.

Modulation Transfer Function

Suppose we imaged a target of black and white parallel lines. Consider the effect of progressively increasing the line spacing frequency of a target and how this might affect contrast. As we might expect, the contrast will decrease as the frequency increases. The MTF is plotted by taking the contrast values produced by a series of different line pairs. The curve drawn from these points shows the modulation (i.e., the contrast) at all resolutions, not just the limit resolution.

For many images, having a high contrast at a lower frequency is more important than the limit resolution. Many high-speed systems fail because the designers don't understand this.

There is another way to think about MTF. Instead of plotting the contrast in the frequency domain, suppose we look at the intensity in the spatial domain. We said earlier that no optical system can reproduce an object's detail and contrast, and we discussed how an image of a dot has blurred edges. If, instead of a dot, a single point of light is imaged through a lens, it also spreads out. This can be measured by the point-spread function (PSF) of a lens, which is a function of the intensity vs. the linear distance across the image of the spot.

The MTF is the Fourier transform of the point spread function. Because the PSF is a single cross section of the spot, we look at two orthogonal PSFs to get a more complete picture of the image. This leads to two (a sagittal and a tangential) MTF curves for each image point. The two curves are sometimes averaged to make a clearer graph.

It is necessary to understand the PSF because of its relevance to real MTF measurements. In fact, it is a point source and not a target that is most often used to determine MTF values. Because MTF curves represent only a single point on the image, it is necessary to show multiple field points or curves to accurately define the full image. For example, sample points taken on the optical axis, at 0.7 the full field and at the full field, will yield a very accurate representation of image MTF. (The 0.7 full field is used because it represents half the area of the full field.)

continued >

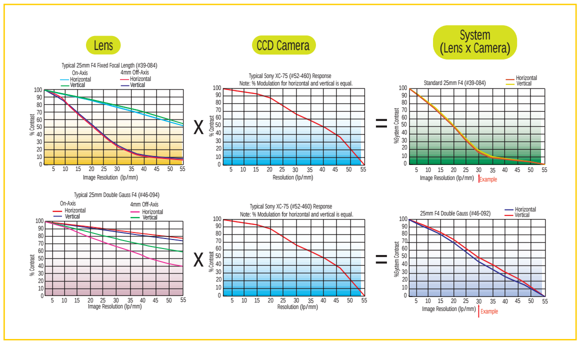


FIGURE 6: *Each system component has its own MTF (lens, camera, cables, capture board, and monitor). Multiplying each MTF yields an overall system MTF.*

Using MTF to choose a lens

Each component of an imaging system has an MTF curve associated with it — even the non-optical components such as capture boards and cables. The MTF for each device describes the relationship between the contrast and resolution (measured in frequency) for that component. By understanding your needs and choosing components with the curves you require, you can integrate a system without paying for components with unnecessary performance.

Consider, for example, a machine vision system set up to look at an object on an assembly line. Assuming that the information needed from the image is not a small detail, the integrator can concentrate on maximizing contrast (increased signal to noise) at low resolutions. This will allow the imaging system to capture the necessary data while allowing the user to run the assembly line faster than with a high-resolution but low-contrast imaging system.

This is also a chance to save money. When we pick a lens for the system, our goal is to maximize contrast. Assume we look at the MTFs of two lenses, one designed for 35mm film cameras, and the other designed to work with CCDs (see Figure 6). The CCD lens is less expensive than the 35mm lens, and doesn't offer usable contrast at high resolutions. However, if we consider the MTFs of both lenses, we can see that at the low frequencies of interest to us for this application, the CCD lens outperforms the more costly 35mm lens by providing higher contrast.

Assuming that the other components are also chosen with an eye to enhancing contrast at low resolutions, the final system MTF — which is a combination of the component MTFs — will provide the desired performance for a high-speed assembly line application. Remember: the choice depends on the application. If, as in the case outlined above, high contrast at low frequency is important, then pay more attention to the left side of the MTF curve.

Using MTF

In traditional system integration, a rough estimate of system resolution is often made by assuming it is limited by the component with the lowest resolution. Although this approach is useful for quick estimations, it is flawed because every component in the system contributes error to the image, yielding poorer image quality than the component with the lowest resolution. A more accurate system resolution can be calculated by combining the MTF of each component.

In addition to the lens, every component in an imaging system also has an MTF associated with it: the cameras, capture boards, cables, monitor, and user's eyes all have MTFs. When looking at the MTF curves, the more you depend on the application and the detector. If the limiting resolution is important, then you want a curve with the 10%–20% contrasts as far to the right as possible. If, as is often the case, high contrast at low frequencies is important, then you would pay more attention to the left side of the MTF curve.

Lenses and apertures

MTF describes contrast and resolution, but what about other image quality factors such as depth of field and geometrical errors? These are consequences of dealing with lenses.

The diffraction of light limits the performance of a lens. The diffraction limit of a lens is affected by the size of the aperture. The aperture is inversely proportional to the f-number, which describes the light-gathering ability of an imaging lens. As the lens aperture

continued >

decreases, the f-number increases. The diffraction limit dictates that the smallest spot that can be imaged through a lens is proportional to the f-number.

Often, however, the limiting factor for a lens is not diffraction; instead, optical errors and manufacturing tolerances limit performance. When this is the case, lens performance can often be improved by increasing the f-number.

Using telecentric lenses to overcome perspective errors

Perspective error, also known as parallax, is part of our everyday experience in gauging distance: we expect closer objects to appear larger than objects that are the same size but farther away. Perspective also exists in conventional imaging systems in which the magnification of the object changes with its distance from the lens. While this is useful for estimating the distance of objects with known sizes, it gets in the way of measuring objects of an unknown size. Perspective is most troublesome in measurement applications involving objects with depth or objects moving relative to the lens.

Telecentric lenses are designed to minimize perspective error. They optically correct for perspective, and objects remain the same perceived size, independent of their location within a depth of field and field of view defined by the lens (see Figure 7).

While telecentric lenses do not inherently have more depth of field than conventional designs, the images tend to blur symmetrically. Because the center of the blur corresponds to the center of the object, however, no error is introduced in measuring the center-to-center separation of objects. This is true even if the object is not in focus. The field of view of a telecentric lens is limited by the front diameter of the lens. Because the magnification is constant for a telecentric lens, different lenses are necessary for different fields of view.

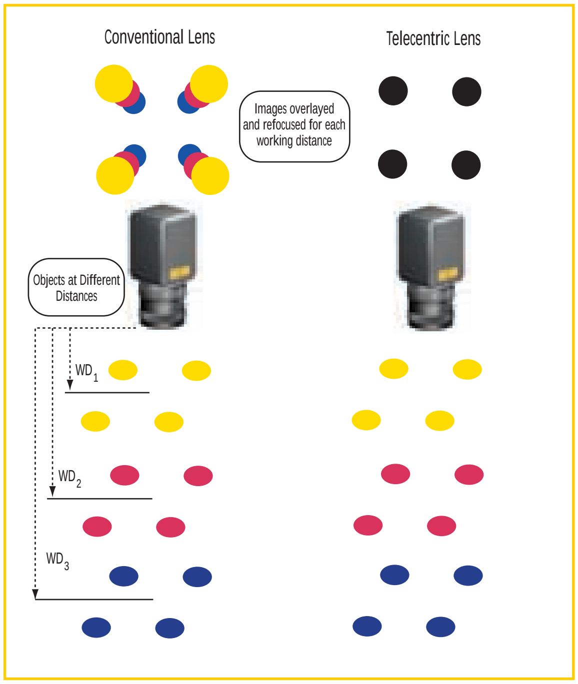


FIGURE 7: *Like railroad tracks that appear to converge at the horizon, perspective error makes the square of dots at the longest working distance appear to be closer together than the square of dots closest to the camera (left). A telecentric lens corrects this error within a range of working distances and over a certain field of view (right).*

Depth of field

As discussed earlier, the depth of field (DOF) is one of the fundamental parameters of image quality. The DOF of a lens describes its ability to maintain a desired amount of image quality as the object is moved closer to and farther from the best focus position. As the object moves closer or farther than the working distance, both the contrast and resolution suffer. The DOF therefore makes sense only when defined at both a specific contrast and resolution. Lenses used at higher f-numbers have larger depths of field.

Although the DOF can be calculated at the diffraction limit, the technique isn't useful if the lens is limited by other factors, which is often the case. (This also means that although two lenses may have the same f-number and thus the same diffraction limit, they do not necessarily offer the same DOF.) Instead, the DOF can be measured at a specific contrast and resolution for an application.

Depth of field is tested using a target with a regularly marked surface that slopes at a 45° angle. Testers can either eyeball the image to see where the image blurs or can calculate the contrast from looking at something called the line-spread function. If the lens includes an iris, its f-number can be raised by closing the iris. In this case, a user who needs more depth of field may be able to gain it by raising the f-number, but at the expense of resolution.

Distortion

Distortion also limits the image quality. There are a host of optical aberrations that cause the lens to change magnification at different

continued >

points in the image. The magnification changes with distance from the center of the field. One important point to remember about distortion is that no information is lost — it is merely misplaced.

All lenses have some distortion, which is worst at the edges of the field. The difference between the actual (distorted image) and predicted (non-distorted object) position can be expressed in terms of a percentage from the center of the field. Distortion can often be fairly well corrected, although it is more difficult to correct for this aberration in short focal length lenses, such as wide-angle or fisheye lenses.

Distortion is troublesome for measurement applications, but it can be corrected. Because no information is lost, once the distortion has been measured (using a distortion target), it can be factored into the calculation of measurements. Furthermore, the images can be corrected by software.

The amount of distortion that is acceptable depends, again, on the application. If the distortion at the edge of the sensor is less than the size of a pixel, it will not have any effect on the image. If the distortion is less than ~2%, the human eye will not perceive it.

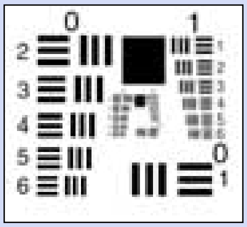
Conclusion

If you are building a machine vision system, you need to understand the characteristics of image quality that we have discussed. Once you understand the tradeoffs associated with the optical system, you can build an efficient system that works for your application. With this information, you can specify a lens that fits the needs of the measurement, without compromising performance or paying for features you don't need. You can also optimize the overall system for your application before prototyping and thus cut development time dramatically. We've seen some of our customers reduce their time-to-market by half. In the process, you might also reduce overall cost.

TECH TIP ON CHOOSING A CAMERA

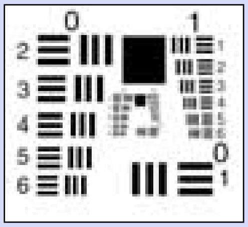
When choosing a camera for an industrial application, many system specifiers instinctively select color because they feel a monochrome image is inferior. That, however, is incorrect. Monochrome cameras have higher resolution, better signal-to-noise ratio, increased light sensitivity, and greater contrast than similarly priced color cameras. Although color imaging may be preferable, the eye perceives spatial differences more clearly in gradients of black and white. In addition, industrial applications requiring a computer interface typically operate with a black and white camera, since a color image requires more processing time and does not yield significantly more information about the object.

When a high resolution color image is necessary, it is beneficial to use a 3-chip (also called 3-CCD or RGB) camera. By utilizing three CCD sensors, these cameras offer the best of both worlds — yielding greater spatial resolution and dynamic range than single chip color cameras. The image is directed to each sensor using a prism and is then filtered to provide independent red, green and blue signals. The RGB output from a three chip camera is considered to be superior to the standard NTSC and Y-C formats, because the color information is on three separate signals.



Using a Single Chip Color Camera

Using a Black and White Camera



WDM Optics Push Optical Coating Technology to New Limits

Applying optical coatings is labor and time intensive — after a technician sets specifications in the coatings chamber, the coating process follows a series of nine separate steps, each critical to producing a quality coating. And the more complex the coatings are, the more effort and expertise coating production will require. The acceleration of optical technology has challenged coating vendors to create increasingly elaborate coatings. These vendors must understand the capabilities of the chamber and the coatings, utilize coating monitoring technology, and consider the costs involved in custom vs. off-the-shelf choices.

Market forces are pushing the performance of optics to their limits. Optical components must be developed to provide the best possible combination of manufacturability, performance, and price. One vital step to success in creating WDM optics lies in a discipline that is often overlooked or misunderstood — coating engineering.

Coating requirements for components such as wavelength-division multiplexers, which require complex and difficult-to-achieve wavelength accuracy and durability, are driving the development of coating technology.

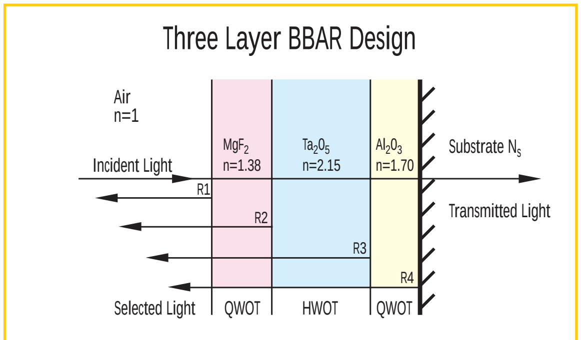


FIGURE 1: *Three-layer broadband antireflection coating uses two quarter-wave optical thickness layers and one half-wave, or absentee, layer.*

How coatings work

Coatings use constructive and destructive interference in thin films to create a specific spectral response (which can be a mirror, a partially reflecting mirror, an antireflection coating, or a filter) over the spectral region of interest. The coatings are thin dielectric films deposited on glass. Dielectric materials are non-absorbing (in other words, exhibit very high transmission) from the UV through the visible and, of particular interest for WDM applications, well into the IR.

To understand interference, consider light as a sine wave. When the lightwave encounters an interface, the reflected portion of this wave changes phase. The total phase change is a result of the combination of phase changes at interface reflections in combination with phase changes due to the optical path length the light travels. The phase change is related to the thickness of the interface layer. Typically, dielectric layers are deposited on the surface of the component in alternating high and low refractive indexes of quarter-wave optical thickness (QWOT).

The QWOT is prevalent throughout optical coating designs because it produces the maximum change in phase for any single dielectric layer. Layers of half-wave optical thickness (HWOT), also known as absentee layers, do not alter the performance at the design wavelength but may be used to modify transmission away from the design wavelength. The resultant added wavefront from all the reflections presents either an additive or subtractive effect. An additive effect is that of a high reflector, and a subtractive effect result as in an antireflection

continued >

coating (see Figure 1).

A coating chamber

It is easier to describe how a coating is made if one understands the parts of a coating chamber. A typical optical coating chamber is 24 to 40 inches in interior diameter and contains an array of components. The coating chamber includes several subsystems (see Figure 2). In this article we focus on the process and limitations of the most common method for making WDM components — vapor deposition.

The first subsystem holds and rotates the components being coated. It is either a planetary dual rotation or calotte single rotation mechanical structure. Planetary tooling is preferred if precision and uniformity are critical; the calotte is used if tight tolerances are not specified, and provides more parts per coating run. The planetary spins the components. Each tool includes a set of standard diameter holes that hold custom inserts, which in turn hold the components being coated. These inserts are made, if not already available, for each type of component being coated.

Moving down in the chamber, the next subsystem is the element heaters. These are placed along the perimeter of the chamber to aid in heating the chamber and specifically the substrate or components being coated. The chamber is typically heated to between 250°C and 300°C.

Next is the focus point of the chamber: an electron-beam gun vaporizes a target, held in a crucible, to create the vapor that fills the chamber and deposits onto the components (as well as all the other surfaces in the chamber). A complex system of crucibles and shutters allows the correct material to be vaporized for the correct amount of time. These crucibles are loaded into a rotating wheel. The coating machine or the operator moves the correct material in front of the gun at the correct time to deposit the next layer. The shutter stops vaporization after the correct material thickness is deposited.

In some systems an ion gun is used to add energy to the material as it is vaporized for better control of the process. This ion-assisted deposition (IAD) method increases the density, or packing factor, of the coating. This in turn decreases the voids in the coating and opportunity for moisture to comingle with the layer. Moisture changes the effective index of a thin film and causes the coatings properties to shift. Moisture in the coating limits the accuracy possible in a coating.

The layers are required to be a specific thickness, on the order of 1/10 of a wavelength of light. Two primary measuring methods are quartz crystal frequency monitoring and optical monitoring.

Crystal monitoring is based on the film being deposited on the crystal the same as the components of interest. As the thickness builds up, the characteristics of the crystal change accordingly. This change can be monitored and directly related to the thickness of the film. The second method uses the same concept, with the exception that it uses an optical detection basis. All of these systems work in concert to deposit very accurate layers of dielectric films to produce the result of the coating design. The chambers can create complex coating structures in excess of 100 layers.

Nine-step process

Coating a single surface takes nine separate steps, and a two-sided component takes 16 steps. Each step is labor- and time-intensive (see Figure 3). A typical broadband antireflective (BBAR) coating can take more than three hours of machine cycle time.

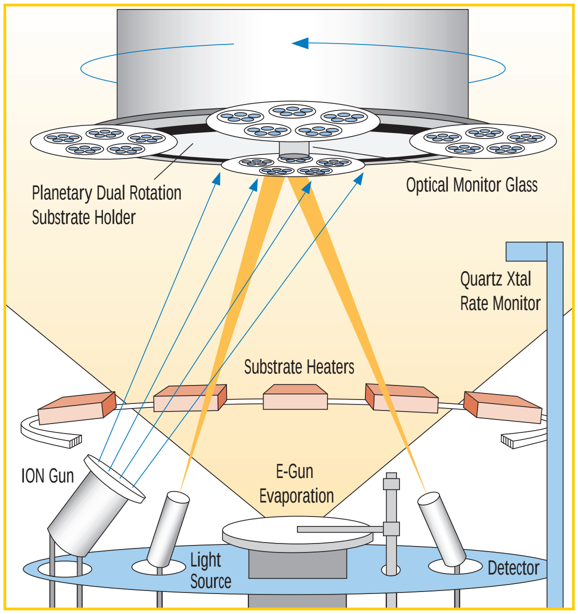


FIGURE 2: Coating chamber subsystems, from top to bottom, may include planetary tooling that holds the components, a quartz crystal deposition rate monitor, substrate heaters, ion-beam gun, e-beam system for vaporizing material from targets and depositing the material onto the components, and an optical deposition rate detector (including the light source and detector).

FIGURE 3: A technician sets specifications on a coating chamber. A labor- and time-intensive process, coating deposition involves nine separate steps.

continued >

It takes the same coating time to coat an entire chamber full of parts as it does to coat a single part. The nine-step process involves:

Prepare the tooling inserts for the coating run. If these inserts do not exist for the specific parts, they must be machined. The machining process can take up to several days depending on complexity of the components to be coated, and the number that can fit into an insert. Clean and load the components into the tooling. Depending on the size of the part, and the number of them to coat, this process can take from seconds per part to minutes.

Prepare the coating chamber for the run. The chamber needs to put through a series of checks to make sure all systems are functioning and all necessary surfaces in the chamber are covered. Load the planetary tools into the coating chamber.

Evacuate the chamber down to 2 x 10^{-5} Torr, and heat the chamber to between 250°C and 300°C. The vacuum removes airborne containments and moisture from the chamber as well as allowing more mobility to the material being vaporized.

Deposit the coating onto the component. Depending on the complexity of the coating, this process can take from half an hour to days. In complex filtering technologies, such as those in telecommunications, multiple hours to days is the standard.

Cool and vent the chamber back to room temperature and pressure. Remove the components from the chamber and test the witness sample. The witness sample is a window that is coated along with the components. This window is the piece that will go into the spectrometer to determine the spectral response of the coating. This window is necessary because the spectrometer cannot test a part with a curved surface. In addition to spectral testing, most coatings are checked for adhesion and abrasion resistance. Depending on the application, coatings may also be required to pass other environmental tests such as high humidity, high/low temperature cycling, salt spray and resistance to various solvents. Inspect and package the components.

Challenges to repeatability

Designing and making a coating is not an exact science. The design of a coating is highly dependent on the deposition chamber in which it will be made. The designer and operator must know and understand the nature of the calibration of the machine, as well as any issues with the performance of the individual subsystems being used. All the factors contribute to the accuracy and repeatability of the coating from run to run. In coatings that require many multiple layers, the risk goes up for effective monitoring of the process.

Monitoring is particularly problematic for the telecommunications industry. The complexity of these coatings drives the performance of typical optical coating chambers and special machines have been made to accommodate the production of wavelength-division multiplexers.

Other telecommunications coatings, expected to function for more than 20 years, can be made in typical coating chambers equipped for ion-assisted deposition. This is particularly helpful for coatings that require tight position accuracy of the center wavelength. Notch or edge filter coatings also benefit greatly from this enhanced technology.

Cost factors

In many cases, tolerances are the key to the simplicity or complexity of manufacture. The engineer who specifies the coating can reduce his

continued >

company's costs and improve the coating yield by asking for realistic performance. If the standard offerings from coating vendors will not meet the customer's need, the customer will do well to keep his requirements as close to the standard versions as his application will allow. Better yet, call the coating company and discuss your requirement with a coating designer. Working with the coating vendor during the design stage can save money, time, and headaches during production.

As with any other product, however, off-the-shelf coatings are less expensive than custom coatings. Any standard coating eliminates costly development and should be available at a shorter lead time. Using tried and tested processes also reduces the probability of failure in the coating chamber.

Coating failures do happen, and no one wants to see several weeks — or months — worth of precision-manufactured glass tossed out because it has a bad coating. Designs made on the computer always claim that a coating is manufacturable, but the execution in the coating chamber can be a different story.

WDM applications demand precision wavelength positioning.

Monitoring coating deposition ensures correct wavelength positioning

WDM applications demand precision wavelength positioning. Picking out one channel and rejecting the others requires that the center wavelength of coating pass or stop band be highly accurate. The accuracy of wavelength positioning depends on how carefully the deposition is monitored.

For typical edge filters, the accuracy can be ±1.0%, and deposition is monitored with a quartz crystal. For DWDM filters, however, accuracy can be as fine as ±0.002%, and deposition is optically monitored with state-of-the-art equipment.

Conventional optical monitoring charts the reflectance (or transmittance) at a preselected wavelength during the deposition of the layer. As the layer approaches quarter-wave optical thickness (QWOT), the percentage reflection or transmission reaches a turning point on the chart (because the QWOT produces the maximum change in R or T). The turning point may be used as a layer termination trigger or as a calibration level from which to calculate the eventual termination value.

Optical monitoring can be highly effective in producing coatings of tight tolerance provided the core design is a regular quarter wave stack. This is a result of a highly effective error compensation feature inherent in the turning point detection and layer termination method: if, during the coating process, the turning point is over — or under — shot, it is compensated by terminating the deposition at the turning point of the next layer. This compensating effect minimizes the cumulative error in the multilayer stack and can result in very accurate filter wavelength positioning, as demonstrated by "successful" production of DWDM narrowband transmission filters.

Quartz crystal monitoring measures the physical thickness of the depositing material, and, therefore, does not involve any turning point methodology. While this technique does not provide any error compensation, it is a very useful when monitoring layers that are significantly thinner than one quarter wave (layers less than QWOT have no turning point and therefore present difficulties for optical monitoring). This quartz crystal methodology is favored in the production of designs such as broadband antireflection coatings, where layers as thin as 10 nm are common. Precision multilayer coatings such as edge

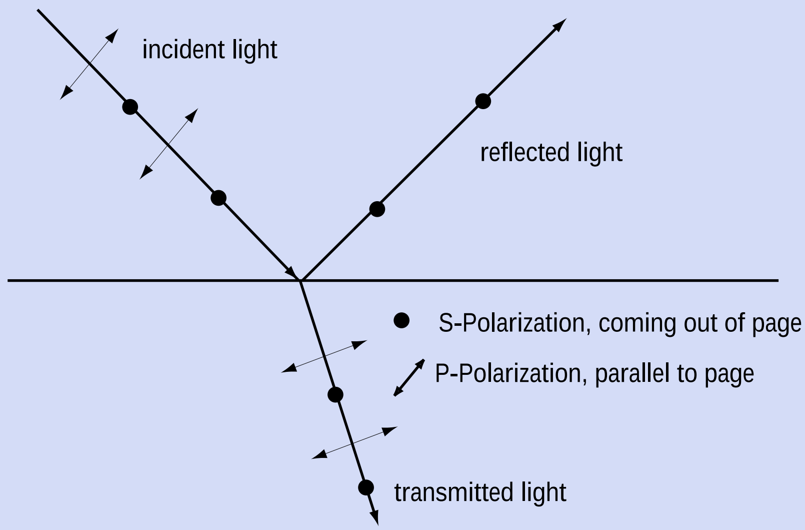
continued >

filters can be produced using quartz crystal monitoring; however, without the compensatory effects of optical monitoring, accurate material characterization and very tight process control are necessary to achieve specification.

Most coatings do not exhibit any significant polarization effects at angles of incidence less than 20°. At higher angles, the S and P states behave quite differently. This is a consequence of their effective angular refractive index, given by: N_s = Ncos(q) and N_p = Ncos(q). At 45°, the variation in S and P performance can be very significant. A 50:50 beamsplitter (random polarization) may transmit 75% P and only 25% S. Polarization insensitive coatings can be produced at high angles for single-wavelength operation. Achieving nonpolarization over a broad waveband, however, presents a difficult challenge to both thin-film designers and engineers. In telecommunications, polarization control is critical, and many coatings include a minimum polarization-dependant loss (PDL) specification. In some telecommunication applications, PDL is significant at angles as small as 15°.

TECH TIP ON S & P POLARIZATION

S & P polarization refers to the plane in which the electric field of a light wave is oscillating. S-Polarization is the plane of polarization perpendicular to the page in the figure below. P-polarization is the plane of polarization parallel to the page in the figure below.



The axis of a linear polarizer determines the plane of polarization that the polarizer passes. There are two ways of finding the axis of a polarizer. A simple method is to start with a known polarizer with a marked axis. Place both the known and unknown polarizer together and transmit light through them. Rotate the unknown polarizer until no light passes through the pair of polarizers. In this orientation, the unknown polarizer's axis is 90° from the axis of the known polarizer.

If a known polarizer with a marked axis can not be found, the axis can be found by taking advantage of the Brewster effect. When light reflects at glancing incidence off of a nonmetallic surface, the S-polarization is reflected more than the P-polarization (see figure above). A quick way to do this is to look at the glare off of a tiled floor or another nonmetallic surface. Rotate the polarizer until the glare is minimized. In this position, the polarizer is oriented so that the axis is vertical.

Keys to Cost Effective Optical Design & Tolerancing

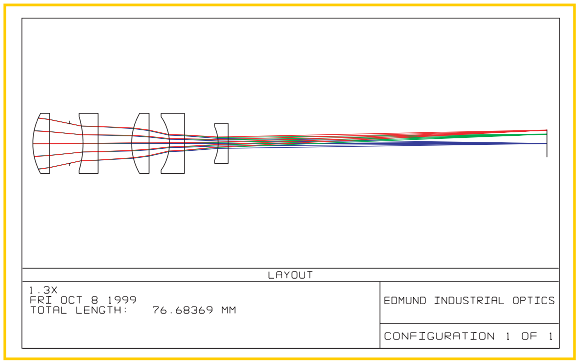


FIGURE 1: Zemax Optical Design Software.

As most designers know, optical design software can be a powerful tool. But it's just that, a "tool". The proper interpretation of the optimized results is just as important as the information inputted. This is why experienced designers will weigh the advantages and disadvantages of using one lens design code over another prior to any actual design. But with growing industry demands, designers need to incorporate all aspects of production into their design in order to ensure that the final product will be brought successfully to market. They need to not only be aware of the nuances of fabrication, assembly, coating, etc., but also with how to integrate cost with the demands of the intended application. Unfortunately, no software program provides a subroutine to assure that costs are minimized.

This introduces the concept of designing with off-the-shelf catalog lenses, which have the dual advantage of being inexpensive (compared to a small custom production run) and immediately available. Clever designers can often integrate stock lenses into custom multi-element designs; by sacrificing marginal performance issues, a significant cost saving can be achieved. Even though stock lenses may not be practical for a required application, they may be suitable for fast prototyping requirements. In addition, the readily available prescription data for most lenses & even many multi-element lenses are encouraging many to use stock lenses (see Figure 1).

This article will attempt to clarify some typical optical manufacturing practices and emphasize the need to monitor costs during the design process. With a keen knowledge of manufacturing practices, lens designers can guide the optimization to an economical solution. By investing some time at a local optics shop, designers can experience firsthand fabrication techniques employed by an optician. Choices made during the design stage that appear to have no effect on production could eventually prove otherwise.

As an example, the simple act of making elements equi-convex or equi-concave could eliminate problems in a seemingly unrelated process such as assembly. Ask any assembler how they feel about lenses that have nearly the same radii on their outer surfaces, and they will tell you horror stories of multiple tear-downs to correct for lenses mounted in the wrong direction. In fact, selecting symmetrical lenses can often introduce cost savings by reducing the cost of test plates and production time.

Any design starts with a given application, and thus some known values. It's the designer's job to solve for the unknowns, typically setting such lens specifications as the radii as variables and constraining others by initially pre-selecting them; namely, the diameter, center thickness, and glass material.

Selecting the diameter

Once clear apertures have been determined, it is important that

continued >

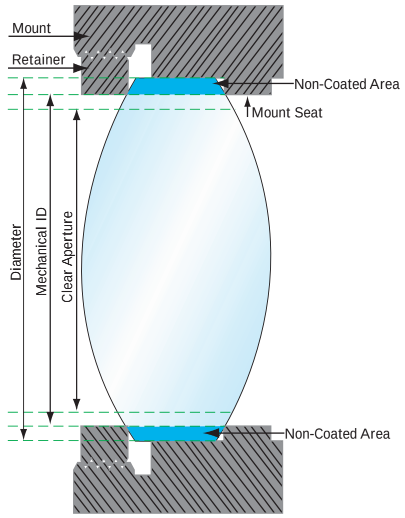


FIGURE 2: *Mechanical mounting considerations.*

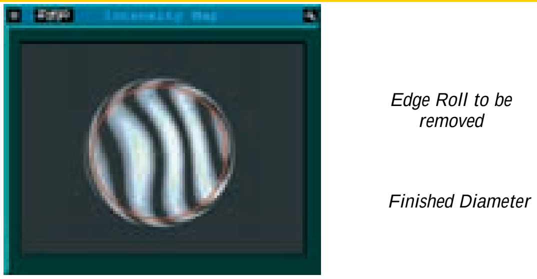


FIGURE 3: *Interferogram of PCX lens showing "edge-roll".*

designers understand how the lens will be mounted, as well as ground and polished. The final lens diameter should be chosen to accommodate the lens mounting (see Figure 2).

When mounting on a mechanical inner diameter (based on contact points with the radii), glare may be introduced from light reflecting off of a spacer, retainer ring or mounting seat/shelf. In comparison, light that reflects off of a larger inner diameter (I.D.) will be cut-off by the aperture of the system. If the element is coated, the diameter of the coating area should be larger than the mounting I.D. in order to avoid exposure of uncoated lens surface areas. Typically, a diameter 3mm larger than the clear aperture diameter is needed for elements in the 20 – 40 mm diameter range.

In order to produce repeatable lenses, manufacturers often use lens blanks (glass in pre-fabricated state) that are typically 2mm larger than the selected lens diameter. This method of "oversizing" allows the optician to remove defects during the final centering process. One common defect, called "edge-roll," (see Figure 3) is a surface deformation that results from excessive wear that the polishing tool exerts on the edge of the lens blank.

Another defect, often referred to as "wedge", occurs when the optical and mechanical axes of an element do not coincide. This centration error can be corrected by aligning the centerline of the lens surfaces with a spindle that rotates about the mechanical axis. The blank is then ground down to the final lens diameter, while being aligned with the optical axis. This in turn defines the diameter tolerance. The deviation angle specification is used to limit the amount of centration error. It is important for a designer to consider this value when reviewing the effect of the compounding errors on the alignment of a multi-element system. Not only must each lens be axially aligned to each other, but the optical assembly must also be aligned to the housing.

The main consequence of working with oversized blanks is that the edge thickness of a bi-convex or plano-convex element will be smaller than at the final lens diameter. The designer can incorporate this knowledge into the design process by using lens diameters that are typically 10% — 20% larger than the final diameters and include a minimum edge thickness operand in the merit function of their chosen software program.

Selecting the center thickness

Typically, a designer will steer designs away from large center thickness values in order to control the material volume, and thus the weight of the final product. Usually as a result of color correction, design software will favor thin lenses with high diameter:center thickness ratios. If kept below 10:1, the diameter:center thickness ratio rarely effects cost. When the ratio approaches 15:1, costs begin to rise for low power lenses with longer radii, as well as meniscus lenses. These types of lenses exhibit "springing" during conventional and high-speed manufacturing. In conventional polishing, lenses are placed on a blocking tool with hot sticky pitch. After polishing, the lenses are removed from the polishing block by chilling the pitch to a brittle state, allowing easy separation from the lens surfaces. Surfaces can deform when stress, introduced in the blocking process, is removed. For high-speed manufacturing, the effect is manifested differently. Increased speed and pressure causes the lens to oscillate, resulting in deformities and making it difficult to control the irregularity (surface shape).

continued >

The effect of the diameter:center thickness ratio on cost can vary due to the lens shape and is actually less cost sensitive for large negative power lenses. In addition, these lenses have large edge thickness values that provide support to handle pressures and stress.

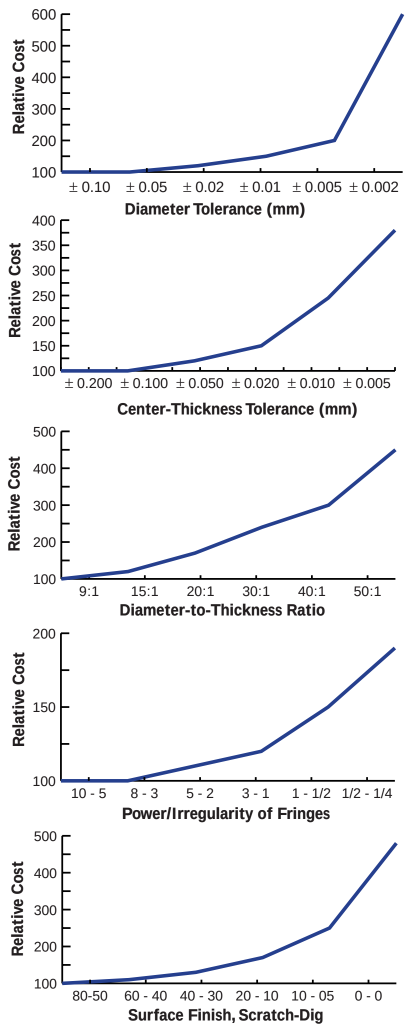


FIGURE 4: *The effect of relative costs are shown for various parameter and tolerance specifications. The value 100 represents the cost of a basic element.*

Source: See Reference #2

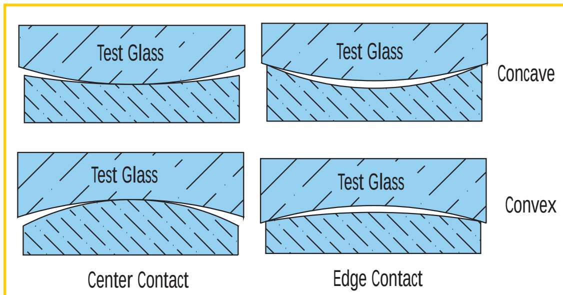


FIGURE 5: *Polishing on the "low side".*

Selecting the glass material

There is almost as much selection in types of glass materials as there is in cost. For example, using a relative price comparison with the most commonly used optical grade BK7 glass as a value of 1, then SF11 is 5 times more expensive, while LaSFN30 is almost 25 times more expensive. Properties of a material that can drive up costs include high staining and softness, which are often difficult to work with and require careful handling. It is important to note that these characteristics can affect production during both fabrication and coating procedures.

Many design software programs provide an option to "model" a glass type, allowing the index and dispersion values to vary continuously. Although this will usually produce quicker results, caution should be used. If this modeling option is selected, the designer must diligently monitor the design to steer it away from expensive and difficult-to-control glass types. Many optical designers will use a personalized glass catalog, usually containing glass types that are less expensive, readily available and possess other desirable characteristics. This method, although slower, may provide for an easier means to produce an inexpensive design.

Using tolerancing schemes

Once the initial design is completed, the designer's next task is to assign appropriate tolerances for the various parameters. Diameter, wedge, power/irregularity and center thickness tolerances all need to be assigned for each element. Design performance will be more sensitive to some of these tolerances, while others have little effect at all (see Figure 4). The designer can limit the use of tight tolerances to the sensitive areas and permit them to broaden or loosen in others. Additionally, many optical shops have varying degrees of success controlling specific tolerances. By getting to know the strengths and weaknesses of various optical shops, as well as the associated costs, designers can streamline the process by directing designs to appropriate vendors.

Tolerancing runs performed by most design software programs assume Gaussian distribution, with errors equally distributed about the nominal value. However, some parameters tend to be skewed either to the plus or minus end of the scale during manufacturing. Opticians tend to polish lenses on the plus side of a center thickness tolerance. By leaving extra material, the optician can rework lenses should they be damaged during later stages of fabrication.

Another trend is the practice of polishing surfaces on the "low" side. When using a test glass to monitor the power tolerances, the optician will avoid center contact in favor of edge contact in order to prevent scratching the polished surface, as well as the test glass (see Figure 5). As a result, the power tolerance is cut in half and thus convex/concave surfaces will be flatter/sharper than the nominal value.

Finally, the presentation of the tolerancing must be interpretable by opto-mechanical designers. By emphasizing the sensitive areas of a design, a designer can help ensure a successful opto-mechanical design. Emphasizing axial position over individual spacing tolerances, for instance, can better control fixed flange distance requirements that

REFERENCES

[1] Tech Spec Bulletin Article, Understanding Optical Specifications, Issue 5, Volume 4, Winter 1998-99

[2] Russell Hudyma and Michael Thomas, Reasonable Tolerancing Aids Cost-effective Manufacture of optics. LASER FOCUS WORLD, May 1991, pgs. 183-193

[3] Warren J. Smith, Optical Component Specifications: Avoiding Pitfalls in Setting Tolerances for Optical Components. The Photonics Design and Applications Handbook 1999, pgs. 346-349

may suffer due to the "stacking" of individual errors.

There are several other topics that have not been discussed due to the scope of this article, but nonetheless should be addressed. They should include but not be limited to coating, surface accuracy (power/irregularity), and surface quality (scratch-dig).

Conclusion

The goal of this article was to bring to light some of the key factors that effect cost after a design has been completed. By being aware of what goes on after a design is handed off, a designer can be better prepared to integrate the relevant issues before and during the actual design. This results in less redesigning and optimization and should lead to a better final product.

TECH TIP ON SURFACE QUALITY

Surface quality refers specifically to the cosmetic condition of the surface of an optical element. During the grinding and polishing stages of fabrication, small defects can occur, such as scratches and digs. A scratch is any mark or tear and a dig is any pit or divot in the surface. The specification used for the maximum allowable flaws is denoted by a combination of numbers, the scratch number followed by the dig number; for example 60-40. The lower the number, the higher the level of quality. For example, a 60-40 value is common for research and industrial applications, whereas a 10-5 value represents a high quality standard for laser applications.

It is important to note that both the scratch and dig numbers do not actually correspond to a specific number of defects. Instead, they reflect the quality of an optical surface by means of visual comparison to a precisely manufactured set of standards. This is in accordance with the MIL Spec. Scratch and dig evaluation as defined by the U.S. Military Specification for the Inspection of Optical Components, MIL-O-13830A.

There is no direct correlation between the scratch number and the actual size of the scratch. As a common reference, the scratch number relates to the "apparent" width size of an acceptable scratch. However, there is some ambiguity since it also includes the total length and number of allowable scratches. Dig numbers do relate to a specific size. For example, a 40 dig number relates to a 400µm (or 0.4mm) diameter pit. Coating quality inspection is also held to the same Scratch-Dig specification as the surface of an optic.

Surface Quality inspection typically includes additional criteria, such as staining and edge chips. Overall cosmetic inspection also includes defects within the material, such as bubbles and inclusions, including striae. Imperfections of this nature can contribute to scattering (i.e., in systems involving lasers) and image defects (if at or near the image plane). Inspection for surface accuracy and quality specifications are limited to within the component's clear aperture.

Building a WDM From Component Parts

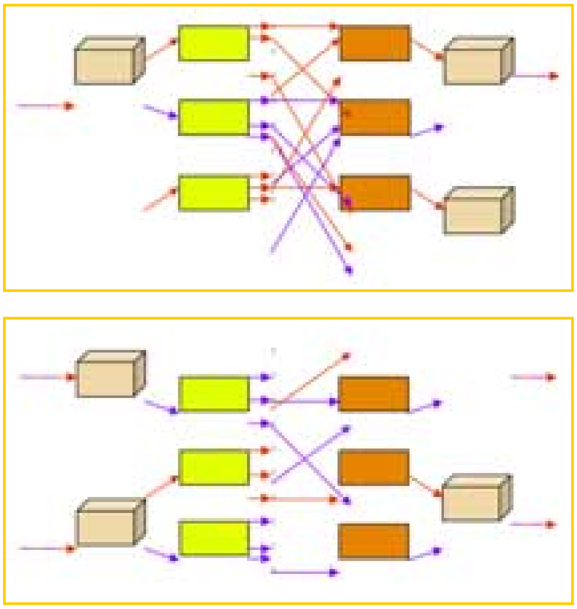


FIGURE 1: *A 3x3 port system architecture.*

This tutorial provides an example of telecommunication product integration using the components required for network systems. The overall objective of this primer is to make your choice of components as easy as possible.

Such a sub-system device is also known as an optical cross connect. In WDM systems, the ability to cross connect or route a wavelength on an all-optical domain is a desired function of optical communication networks. Such a device not only simplifies but speeds networking. Traditionally, this was done by electronically processing the input optical information. Problems associated with this technique were noise, cross talk, over/under signal amplification or modulation of the signal when being routed. As a result, the ability to switch optical signals in one port and out another without electronic conversion is advantageous. Removing the electronic step then requires a cross-connect that has the ability to process light information only. The example below will explain how to construct such a device.

This article will allow users to design and construct a 3 x 3 port device capable of separating 1310nm signals from 1550nm signals, and routing them accordingly.

Step 1: System architecture

When designing a WDM sub-system, it is important to start by arranging or mapping the sub-system. Mapping the sub-system's architecture will give the designer a visual model of the transmitted signal's path. For our example, the 1310nm and 1550nm signals are input into a device that separates the two signals. The separate signals are then transmitted to another device that further splits each signal into three individual signals.

These three individual signals are then processed to meet specific criteria such as wavelength, polarization, etc. Once the correct criteria are met these signals are routed to another station where data content is extracted. The independent signals are then recombined and output for further transmission. Such a device is also known as an optical cross connect. The ability to cross connect or route a wavelength is a key function of communication networks (see Figure 1).

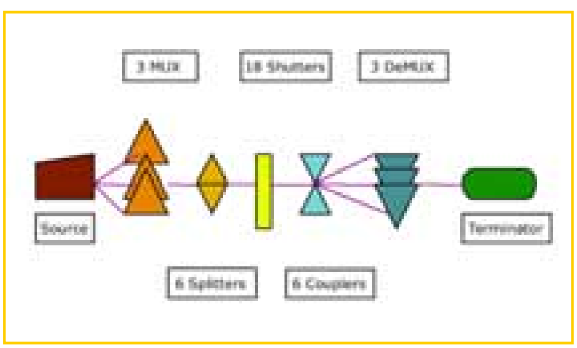


FIGURE 2: *Flow chart of needed components.*

Step 2: System conceptualization

As the designer, once you have traced the signal path, you need to conceptualize the sub-system. Unless it is for a customized or metro application, it will need to be feasible and need to be constructed from off-the-shelf components. The architecture map not only allows system visualization but also helps the designer determine what components are required. Figure 2 illustrates a flow chart of the types of components we expect will be required to construct the 3 x 3 port. For this port, the basic components required are:

- Waveguides

continued >

FIGURE 3: Some of the Edmund Optics components used to construct a routing device. These components are listed on our print and on-line catalog and include the following:

Multi-Channel IR Laser Light Source
Fiber Connector Patchcords
SingleMode WDM
Fiber Optic Coupler

- Source capable of transmitting 1310nm and 1550nm signals
- Device to separate/combine signals
- Device capable of converting one signal into three independent signals (and vice versa)
- Device to route signals

Step 3: Component selection

Now that we have determined our needs and have mapped our sub-system, our final task will be to select the components. On the next page are stock devices Edmund Optics provides which could be used in our 3 x 3 port switch. These devices are: Multi-Channel IR Laser Light Source, Fiber Connector Patchcords, SingleMode WDM, and Fiber Optic Coupler.

Conclusion

In review, this sub-system will separate a composite signal into three individual signals. Once this information is processed and routed to its correct destination, the signals will be recombined and sent back into the network. By mapping and charting, we have designed and selected the proper components required to construct a 3 x 3 port which will act as a signal routing device.

TECH TIP ON BALL LENSES

Ball lenses are great tools for improving signal coupling between fibers, emitters and detectors.

The effective focal length of a ball lens is very simple to calculate (Figure 1) since there are only two variables: the ball lens diameter, D, and the index of refraction, n. The effective focal length is measured from the center of the lens. Therefore, the back focal length can also be easily calculated.

$$BFL = F - \frac{D}{2} \qquad EFL = \frac{nD}{4(n-1)}$$

The Numerical Aperture, NA, of a ball lens is dependent on the focal length of the ball and the input diameter, d. Since spherical aberration is inherent in ball lenses the following equation begins to fall off as d/D increases.

$$NA = \frac{2d(n-1)}{nD}$$

You can see from the graph (Figure 2) how the NA changes as input beam diameter increases. Please note that this graph includes the effects of spherical aberration.

When coupling light from a laser into a fiber, the choice of the ball is dependent on the NA of the fiber and the diameter of the laser beam. The diameter of the laser beam is used to determine the NA of the ball lens. The NA of the ball lens must be less than or equal to the NA of the fiber in order to couple all of the light into the fiber. The ball lens is placed directly onto the fiber as shown in Figure 3.

To couple light from one fiber to another fiber of similar NA, two identical ball lenses are used. Place the two lenses in contact with the fibers as shown in Figure 4.

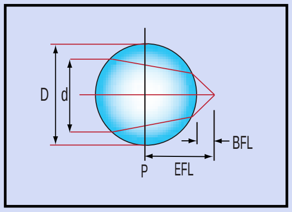


Figure 1

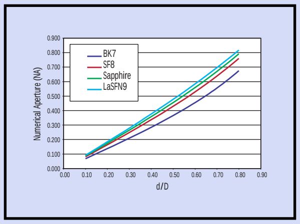


Figure 2

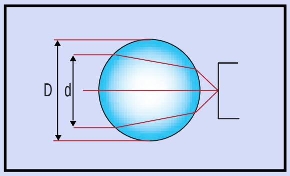


Figure 3

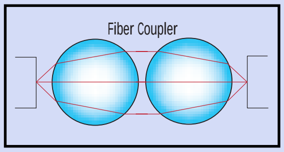


Figure 4

Lens Selection and MTF: Demonstrating the Relationship Between Contrast and Resolution

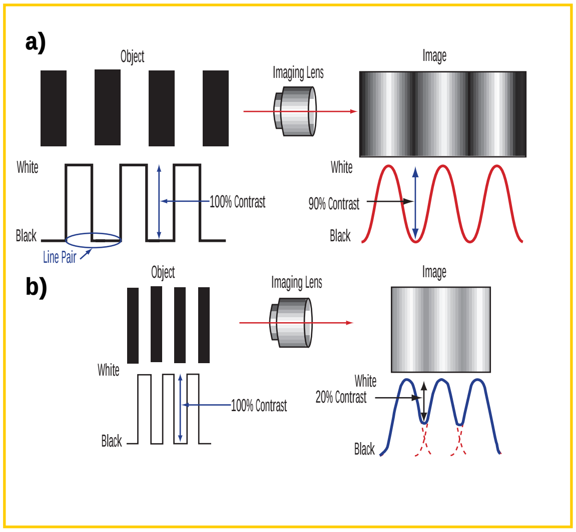


FIGURE 1: At increasing frequencies, optical information passing through a lens loses contrast. (a) The 100% contrast information becomes 90% contrast. (b) The higher-frequency information that starts at 100% contrast becomes 20% contrast after passing through the same lens.

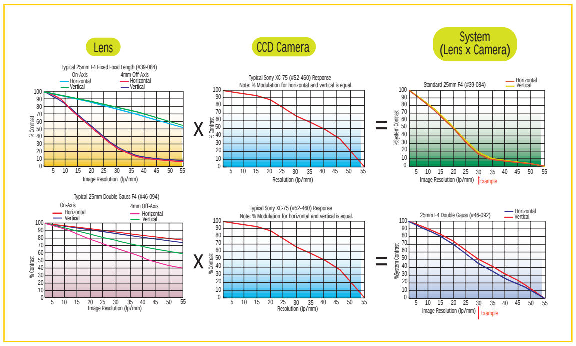


FIGURE 2: The MTF graphs for a 25-mm fixed focal length lens and a 25-mm double-Gauss lens show how contrast varies with the image resolution of each lens. Multiplying the worst-case MTF curves for a lens by the MTF curve for a camera yields an MTF curve for the system (the lens-camera combination). The MTF for the camera is equal in horizontal and vertical directions.

The performance of an imaging system is determined by its ability to provide images of a given quality. Image quality requirements vary depending on application, and are determined by the amount of information that is needed about an object in the image. The variables comprising image quality are: resolution, contrast, distortion, and perspective errors. There is also a measurement that combines resolution and contrast into a single specification.

A low-resolution image contains blurry scenes in which objects lack detail. A high-resolution image provides crisp edges and includes much detail. Contrast also factors into image quality because it expresses how well an image differentiates between an object's shades of gray. An image with low contrast will appear "washed out" because it lacks vivid blacks and whites.

Resolution and contrast are closely related. To understand this, think of imaging a target with alternating equal-width black-and-white lines (Figure 1a). This target represents 100% contrast. No lens — not even a perfect one — at any resolution can fully transfer this contrast information to the image because of the inherent diffraction limit dictated by physics.

Now imagine that the width of the line pairs on the target decreases (that is, the frequency increases). As the frequency increases, the lens is less and less able to transfer the contrast, so the resulting image has less and less contrast (Figure 1b). (A line pair is one black and one white line of equal width. The "frequency" of these line pairs is often defined as the number of line pairs per millimeter, or lp/mm.)

MTF incorporates resolution and contrast

When you must characterize the resolution and the contrast provided by a lens, you can refer to modulation transfer function (MTF) supplied by the manufacturer for a specific lens. You do not have to measure the MTF for a lens. The MTF describes the ability to transfer contrast at a particular resolution (frequency) from an object to an image. In other words, the MTF indicates how much of the object's original contrast gets lost as the frequency in the object being imaged increases. In this way, the MTF combines resolution and contrast in a single specification.

Manufacturers measure the relationship between contrast and resolution and then plot the results as shown in Figure 2 for two lenses. The points on the lines provide the MTF values. Specifically, the graphs plot the percentage of transferred contrast vs. the frequency (lp/mm) of the lines. As mentioned above, the contrast in the image decreases with increased frequency. The MTF illustrated in Figure 2 was measured both on axis (at the center of the image) and for the

continued >

full field (toward the corner edges of the field, or off axis). These measurements tell you how well the lens can resolve features throughout a field of view. Also, notice that the plot includes both horizontal and vertical performance. The difference between these two measurements indicates the amount of astigmatism present in the image.

To understand the importance of the MTF specification, consider a conventional technique used to predict a system's performance. For a typical machine-vision system, a designer might estimate the system's performance using the "weakest link" rule of thumb. The rule holds that the system's resolution depends mainly on the component with the lowest resolution. This approach proves useful for quick estimates, but systems tend to have lower resolution than predicted by this rule of thumb, because all of the optical and electronic system components reduce resolution to some extent. And the quick estimate includes no consideration of contrast, which is also critical to image quality.

To accurately predict the image quality of the optical system, you must combine the effects of each component to determine how the overall system will affect resolution and contrast. Within a system, every component — the lens, the camera, the cables, the capture board, and so on — has an MTF. The system MTF is the product of all of the component MTF curves.

To accurately determine whether a particular lens provides sufficient image quality, you must multiply its MTF by the MTF for each component in the system. You can observe how MTF affects system performance by comparing the resulting MTF for two different lenses used with the same camera. The examples in Figure 2 compare a 25-mm fixed focal-length lens with a 25-mm double-Gauss lens, each mounted on a Sony XC-75 CCD monochrome camera. (This example simplifies the "system" to cover just the camera and the lenses to illustrate how lens MTFs can affect performance.) By analyzing the lens-camera MTF curves for each combination, you can determine which combination will yield sufficient performance for a specific machine-vision application. For this application, assume that you require a minimum contrast of 35% for an image resolution of 30 lp/mm. The double-gauss lens is the better choice.

Watch lens MTF specs

Lens manufacturers often can provide theoretical or nominal MTF graphs for lenses. Although this information can be helpful for planning purposes, it doesn't indicate the actual performance of a manufactured lens. Manufacturing always introduces some imperfections that degrade the performance of a lens. Accurate MTFs can be obtained either from software (as long as it takes the manufacturing tolerances into consideration) or by measuring the actual MTF of the lens after manufacturing.[1] Not all lens manufacturers can provide accurate MTF measurements, however. Be sure to ask for measured MTF data when you're evaluating lenses for machine-vision systems.

REFERENCES

[1] The Web site for the Research Libraries Group (Mt. View, CA) provides more information about how to measure MTF: www.rlg.org/preserv/diginews/diginews21.html.

Using Light to Read the Code of Life

The DNA sequencing field is competitive and fast-moving, partly because so much data must be obtained, and partly because the stakes in the biomedical industry are so high. Optical imaging system is an integral part of most sequencing systems. An incredible amount of proprietary work focuses on optimizing setups for efficiency, accuracy, and cost effectiveness. However, most DNA sequencing systems work on the same general principles.

DNA is a long double-helix molecule made of four nucleotides. In order to “read” the DNA, researchers must figure out the sequence in which the nucleotides are arranged. We describe how DNA is fragmented, tagged with fluorescent molecules, and how the fragments race across gel lanes. We describe the basic optical system required to excite and collect fluorescence from gels. As with most optical systems, the DNA sequencers require robust design and tradeoffs between speed, price, and accuracy.

Intro

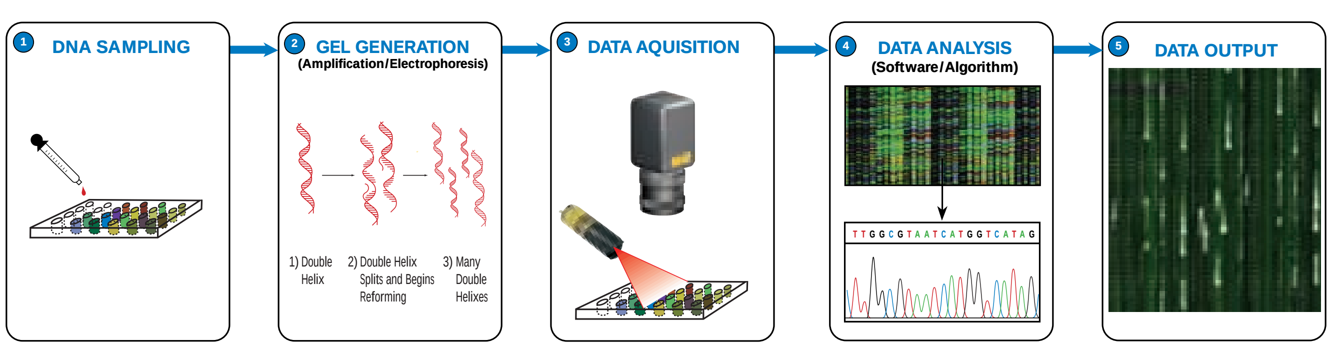
The newest advances in biotechnology require the ability to read DNA - or rather, to read the sequence of the four bases (aka nucleotides) that make up DNA. The ability to sequence segments of DNA accurately and quickly provides a clear advantage to the researcher. Therefore, the makers of DNA sequencers put an incredible amount of proprietary work into optimizing setups for efficiency, accuracy, and cost effectiveness.

DNA can be sequenced in a number of ways. However, most DNA sequencing systems work on the same general principles. To read the sequence of bases, the DNA molecule is fragmented and tags are added to identify the ending. The fragments are separated by size and then the tags are read.

Before sequencing, if there is only a small sample of DNA available, it may be amplified, using a polymerase chain reaction (PCR). An enzyme (DNA polymerase) uses single-stranded DNA as a template to make a new, complementary strand from a stew of available nucleotides. The process can be repeated over and over again.

Similarly, PCR-based sequencing can copy the original DNA strand

DNA SEQUENCING – A General Overview



continued >

- with one important difference. Along with the pure nucleotides, the solution in which the process takes place contains dideoxynucleotides. These dideoxynucleotides lack a hydroxyl group, which means that no more nucleotides can attach to the strand. Each of the four dideoxynucleotides (one for each type of nucleotide) also has a fluorescent molecule attached. Because each of the four different kinds of dideoxynucleotides fluoresces a different color, the base at the end of the chain can be identified.

After this reaction ends, the solution is full of DNA fragments, of many different lengths and with four different endings. If one shone a low-energy laser at the solution and imaged the resulting fluorescence, four different colors would be apparent, but the image still wouldn't provide information about the sequence of the nucleotides - for that, the fragments must be sorted by size. This can be done by taking advantage of the molecules' electrical charge.

Electrophoresis

The solution is introduced onto an agarose gel with an electric potential. The negatively charged DNA migrates across the gel toward the positive terminal. Not surprisingly, the smallest segments move the fastest and farthest down the gel "lane". This process takes some time, but it does sort the fragments by length. The gels can be "read" by lab technicians under ultraviolet light, but it is more efficient to automate the reading. If one shines the laser at the gel and captures the image of the fluorescent tags, the spatial sequence of the colored tags correlates to the sequence of the bases in the DNA (see Figure 1).

Ideally, the system would capture an image quickly that could distinguish fragments very close together with inexpensive imaging equipment. As with most optical systems, however, the design of DNA sequencers balances tradeoffs between accuracy, speed, and price.

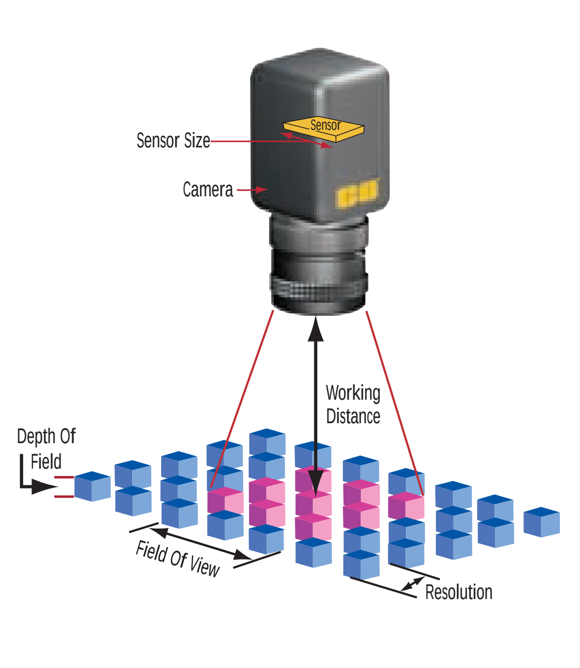


FIGURE 1: Fundamental parameters of an imaging system include the resolution of the object, the field of view, and the depth of field that the user wishes to image. The working distance, from the object to the lens, is also important, as is the sensor size. The primary magnification is the field of view divided by the sensor size.

Optics

Every imaging system has the following elements:
- Field of View: The portion of the gel that fills the camera's sensor.
- Working Distance: The distance from the front of the lens to the gel.
- Resolution: The minimum feature size of interest on the gel. The resolution should be sufficient to differentiate between wells and the spaces between fragments separated by electrophoresis.
- Depth of Field (DOF): The maximum object depth that can be maintained entirely in focus. The DOF is also the amount of object movement (in and out of focus) allowable while maintaining an acceptable focus. The DOF is not typically critical for these systems.
- Numerical Aperture (NA): The light-gathering ability of an imaging lens - important because this is a low-light application. The NA is inversely proportional to the aperture of the lens.
- Sensor Size: The size of the sensor's active area. This parameter is important in determining the proper lens magnification to obtain the desired field of view.

In the most general terms, the imaging system must also:
- provide a light source (a laser, in this case)
- be able to move and focus
- collect, detect and output optical data to a computer

Algorithms

Once you've collected the raw data, how do you process it? Algorithms are tailored to specific systems, but all have to tangle with the same issues:
- Resolving background and peaks
- Contrast (and resolving adjacent bases that are identical)

continued >

• Pattern recognition

Basically, the algorithm must be able to distinguish fluorescing tags from the background. It must also be able to separate one tag from the next: when the tags are for different bases, they are easier to distinguish, but what about when a base repeats? The tags are imaged as intensity peaks - in short, they are somewhat fuzzy colored dots.

The optics of the system, and the algorithm used to interpret the image, dictate the minimum distance between tags (at the time of imaging) for consecutive nucleotides. This, in turn, dictates how long the electrophoresis will take: the closer the tags can be, the shorter the time. And the shorter the time until an acceptably accurate result is obtained, the better.

Contrast is often used as a normalized metric to describe the limit beyond which the system cannot resolve a signal from a background. Contrast is measured at different resolutions. The system must be able to determine whether a block of a single color indicates one, two, or several peaks.

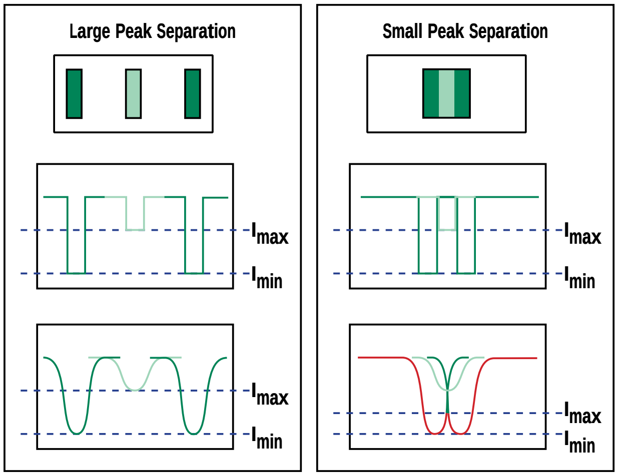
One can improve the contrast by filtering out unwanted wavelengths. There are some other methods that can also minimize the contrast thresh-

Peak Proximity Effects on Contrast

The effects of lowering peak separation is only evident asthe peaks are blurred by non-ideal components within the system. The peak levels are represented by - - - - . The overall resulting curve is shown in ——— .

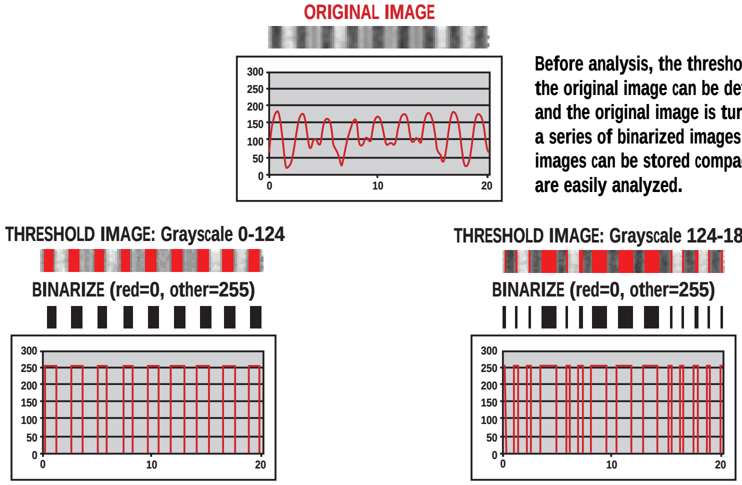
IDEAL SYSTEM
Identical contrast levels of the peak are placed closer together

NON-IDEAL SYSTEM
Although the contrast level is comparable to the ideal system case when the peaks are far apart, as they come closer together, they blend into each other. This causes a break down in overall contrast and make the central peak impossible to discern.



THRESHOLD TECHNIQUES FOR IMAGES

(Note: The Grayscale Values are Inverted)



Before analysis, the thresholds of the original image can be determined and the original image is turned into a series of binarized images. Binary images can be stored compactly and are easily analyzed.

continued >

old and improve the peak separation. The peaks can be accentuated by an algorithm that takes the derivative of the image line profile and plots the resulting slope. Noise, however, can cause problems with this approach.

Another option is to set up the software as a pattern recognition system. This requires active calibration tools, but because the general shape of the signal profile is fairly consistent, it enables quick analysis of patterns. Finally, converting the data to a binary representation can simplify the algorithms -- but only if one sets the threshold with some care.

Calibration: active or upfront?

A critical component to any imaging analysis system is calibration. A system must be suitably calibrated in order to trust the results that are generated. There are two methods to obtain a calibrated system: active and upfront calibration. Both of these methods can ease tolerances and reduce the manufacturing cost of the system.

For example, one method of active calibration is to keep the optical boresight in check across focus movements using calibration software and calibration marks in the electrophoresis gel. However, this and any other active calibration method requires computation time that will slow down the data analysis.

For values that will not drift over time or changing environmental conditions, upfront calibration is more cost-effective and does not slow down processes significantly.

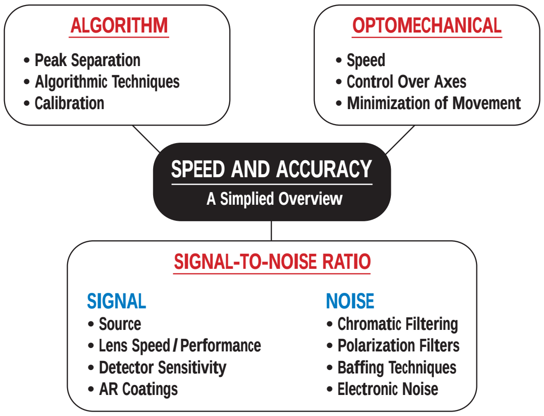


FIGURE 2. *A simplified overview of the various components that affect speed and accuracy*

Designing for speed and accuracy

As mentioned earlier, DNA sequencers are designed to offer faster and more accurate readings than manual sequencing or competing sequencers. Most commercial systems measure performance by throughput. This is measured in tens of thousands of base-calls per run, hundreds of sequencing lanes per day, or thousands of fragment analyses per day, each within some percentage of accuracy.

Speed and accuracy are closely linked precisely because one must often be sacrificed to improve the other. If one can increase the signal-to-noise ratio (SNR), however, then both speed and accuracy can be increased. In this case, the signal is the fluorescence from the tags, while the noise is the background. Improving SNR inherently eases the computational component of the system, and thus affects the throughput. The goal is to maximize the signal by optimizing the light source, the transfer of light through the lenses and lens coatings, and the detector sensitivity.

Because fluorescent energy is typically very low, the signal starts out low. To a certain extent, one can increase it by increasing the power of the laser that excites the fluorescent molecule. If one uses illumination-shaping optics, one can match the laser illumination with the field of view of the imaging system -- this also helps to maximize the signal.

The lens must also pick up enough light to be read by the system. Although lens performance is not directly related to the amount of signal incident on the detector, it does determine whether the system is imaging the energy in the correct location and within a minimum amount of area. By accounting for the fluorescent emission pattern as well as the coating and performance of the lenses, the radiometric transfer can be calculated from the numerical aperture. Coatings on the optics both help increase throughput and reduce noise by reducing lens flare and ghost reflections.

Noise can also be reduced using baffles and filters. Baffles isolate

continued >

stray light from the detector. Proper modeling can predict the problem surfaces and stray light sources. Then the opto-mechanical design can be altered to minimize the problems by including features like strategically placed baffle threads and light-absorbing finishes.

Above, we mentioned using filtering out unwanted colors. There are many types of wavelength-differentiating filters. One must balance the absolute throughput with the throughput of the desired wavelength range. In general, the narrower the wavelength range you choose to sense, and the sharper the filter's boundaries, the less absolute throughput you receive. One can also use polarizing filters to suppress stray light, but again you must find a balance: Polarizers can be extremely powerful at reducing noise, but they also cause significant fading of the signal.

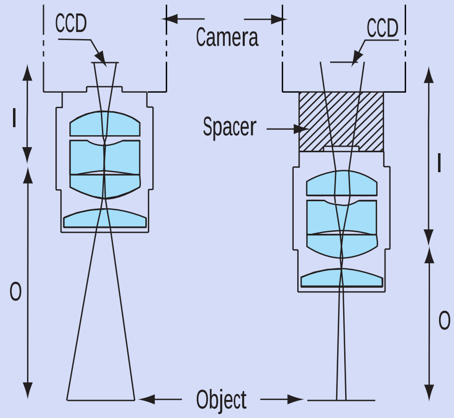
When selecting a detector, sensitivity is not the only criterion but again because the signal level is low - the sensitivity should play a significant role. CCDs produce a linear response and offer high quantum efficiency, which leads to good sensitivity across the spectral band of interest. Speed can sometimes be an issue in arrays, although scanning systems are emerging as an alternate solution.

Conclusion

Design issues for DNA sequencers are similar to many optical imaging systems, in that they involve trade-offs between speed, cost, and accuracy. On the other hand, the technology and economics of DNA sequencing result in systems that push the limits of the imaging systems - and the ingenuity of their designers.

TECH TIP ON WORKING WITH SPACERS

Fixed focal length lenses are an economical solution to many machine vision inspection applications. One drawback of these lenses, however, is that they are typically designed and optimized for an infinite conjugate, which leads to long minimum working distances. The required long distances make mounting these lenses in a bench-top application impractical, and provide fields of view much too large for close inspection. Spacers address these issues by reducing the specified minimum working distance and by increasing the magnification, which decreases the field of view.



It is important to keep in mind, however, that adding spacers forces the lens to focus much closer than its optimized design. This may cause an otherwise well-designed lens to exhibit increased distortion, chromatic and spherical aberrations, reduced depth of field, illumination non-uniformity, and decreased light gathering ability. These problems become more prevalent as additional spacers are introduced, and the lens is forced further and further from its design.

Keeping a Tight Focus on Optics

Problem: The engineers hadn't left provided enough space for the optics.

Recently a semiconductor capital equipment maker was designing a new wire-bonding machine that included a vision system. Engineers there knew the demands on the vision system were not particularly strenuous, so they concentrated most of their efforts on the electronics and motion control. They did, however, leave some space for the camera and optics.

That's when my company got involved. The OEM asked us to design a vision system with certain multiple magnifications that could be changed by the user in the filed and which, of course, would fit in the space available.

Problem: The engineers hadn't left provided enough space for the optics. Too bad, because with better planning, we could have provided the required magnifications with ease using optics that were both inexpensive and off-the-shelf. As it was, the only alternative was a custom system that came at some cost. Moreover, some of the specs had to be relaxed simply because of the mechanical constraints.

Though the engineers involved had the right idea, they hadn't considered that the working distance for lower magnifications tends to be longer than that for higher magnifications. And lack of sufficient room complicated the focusing and compensation methods.
The irony was that optical specifications were not, in general, unreasonable. They would have been easily met had we been able to place the lenses at our discretion.

The situation these engineers ran into is not at all uncommon in semiconductor manufacturing equipment. Few manufacturers in this area have more than a single engineer in charge of optical systems. This despite the fact that it may be unusually difficult to design these systems because they must fit into the design after (and around) other systems.

Often, the integration of a vision system means snaking the optical system through the equipment without interfering with the primary process. This is particularly true for operations that include wirebonding, die packaging, aligning wafers, and lining up registration marks before lithography or metrology.

Optics are often the last systems that engineer's design into their equipment, but it can be crucial. The most cost-effective way to buy optics depends both on volume and specifications of the system. The application and not the manufacturer should determine these specifications. If you take some time, early in the design process, to consider the requirements of your system, you can fulfill them with the best balance of performance, cost and yield available. The application dictates the system performance/requirement, which dictates the specifications. The System requirements should be identified and establish at the same time as the packaging is being decided upon. Allowing the required envelope will reduce optical performance/cost problems later on in the project.

Off-the-shelf or Custom?

The Design

With basic parameters nailed down, the next step is to work out a combination of focal lengths and object/image distances. The bad news is that this usually involves calculating thin lens equations, which you prob-

continued >

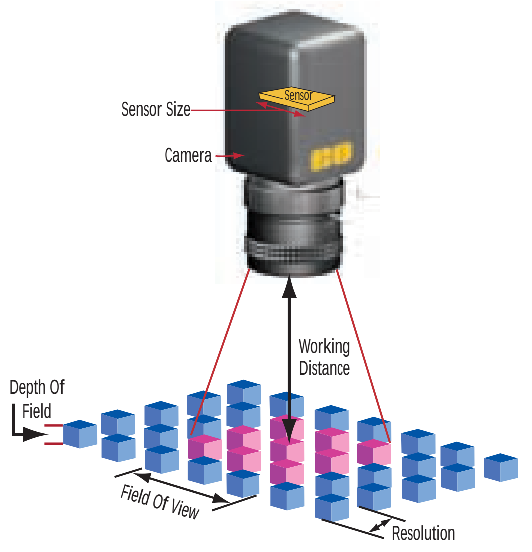


FIGURE 1: Fundamental parameters of an imaging system include the resolution of the object, the field of view, and the depth of field that the user wishes to image. The working distance, from the object to the lens, is also important, as is the sensor size. The primary magnification is the field of view divided by the sensor size.

FUNDAMENTAL PARAMETERS OF AN IMAGING SYSTEM

Field Of View (FOV): The viewable area of the object under inspection. In other words, this is the portion of the object that fills the camera's sensor.

Working Distance: The distance from the front of the lens to the object under inspection.

Resolution: The minimum feature size of the object under inspection.

Depth Of Field (DOF): The maximum object depth that can be maintained entirely in focus. The DOF is also the amount of object movement (in and out of focus) allowable while maintaining an acceptable focus.

Sensor Size: The size of a camera sensor's active area, typically specified in the horizontal dimension. This parameter is important in determining the proper lens magnification required to obtain a desired field of view.

ably saw last in a college physics textbook. The good news is that you don't have to do it yourself. Most modern optical companies have optical design software that can quickly and easily provide a preliminary solution. Unless the problem is extremely complex, this service is usually free.

This is the point before the design is finalized at which to consider whether you need lenses custom-made or if off-the-shelf optics will do. Custom lenses are almost always used to correct aberrations and or package requirements. But sometimes correction is not important or a combination of off-the-shelf and custom lenses will work as well.

Off-the-shelf vs. Custom

Economy of scale is everything when it comes to the price of optics, because of how lenses are made. Low volume will always favor off-the-shelf elements. But the advantages diminish as volume rises and other factors take over. As a general rule of thumb, the custom approach makes economic sense only when one needs thousands several hundreds of lenses. (As with any rule of thumb, there are always exceptions.) But if a custom approach is absolutely necessary, it can be done at a reasonable cost for 100 to 1,000 and up pieces.

Off-the-shelf optics are made in quantity, in continuing production, and kept in stock by suppliers. These stock lenses are typically designed into standard matrixes in a wide variety of sizes and focal lengths.
Prototype using off-the-shelf components: It is fast, inexpensive, and lets you confirm image quality requirements. Moreover, custom lenses are astronomically expensive in the small quantities normally used for prototyping. Custom lenses made by traditional methods may require long lead times. If one uses lenses made with deterministic grinding and polishing, the lead-time is less but the cost will be high. Of course, if prototyping shows the design must change, any benefit from the expenses is lost.

Where are the breaks

Both custom and off-the-shelf elements are viable options for between 1,000 and 100,000 pieces. No one stocks off-the-shelf lenses at such a large volume without a forecast from a specific customer. But it's easier to handle increases in volume with an off-the-shelf option, because there is less risk in overstocking these lenses than in overstocking a custom lens. And However, if a custom lens can reduce the number of elements in the design, this approach becomes even more cost-effective at such volumes.

Custom lenses are almost always the choice above 100,000 pieces. At these volumes, the benefits of eliminating elements become more pronounced. There are real economies of scale. And the 100,000-piece level is a significant breakpoint: The cost/piece of 200,000 lens is not significantly less than that for 100,000.

Other factors

Additional factors that govern the choice between stock and custom lenses include whether or not designs need minimal weight, small size, tight tolerances, or strenuous specifications for the focal length or aberration correction. If the design is already complete and assumes custom lenses, changes to incorporate off-the-shelf lenses may be expensive. Changing the lens inevitably means different mounting to accommodate any changes in focus. Even lenses with identical focal lengths can mount differently because a change in radius alters where the lens must sit. All in all, the cost for engineering these changes may outweigh the savings of an off-the-shelf lens.

continued >

FIGURE 2: *An assortment of custom and off-the-shelf optics and optical assemblies available from Edmund Industrial Optics.*

Though custom lenses can reduce the number of elements in a design, this may or may not cut costs. Additional elements, however, inevitably add weight to the system. Most off-the-shelf lenses are larger then than if designed specifically for an application, as the components are meant to cover a broad base of applications.

In a tight space, a custom lens can be a real advantage. Another consideration is tolerancing. Tolerance stack-up can accompany the use of numerous elements (rather than custom optics) to correct aberrations. A marked decrease in performance results. In addition, some results require that a certain element have a specific tolerance that may not be standard for off-the-shelf versions.

There are also situations that demand a very specific focal length and there is just no easy way to get around it. Some designs may need a specific form of optics, such as a meniscus lens, to correct aberrations; these types of lenses are often unavailable not available off-the-shelf.

Finally, there are many ways of customizing off-the-shelf elements in lieu of going full custom: Two Some of the most widely used include edging down a lenscomponent, cutting it to a specific size or custom coating it.

The easiest customization is changing the diameter of a stock lenselement. It is easy to edge down or cut an lens element even in small volume. This is often important for mounting in an existing housing or in cramped quarters. "Edge downs" can be quick and inexpensive. Special coatings are frequently a motivation for a custom lens. Sometimes designs require low reflectance at a specific wavelength or an antireflection coating in the UV or near-IR range. Lens suppliers are accustomed to fielding requests for special coatings on batches of uncoated lenses. As with edge downs, the costs are usually quite low and the lead time is short. The cost are reasonable depending on the lot size and the turn around time required.

Before signing a PO, take some time to think about your system's minimal requirements, and how you can fulfill them in the most effective manner. This includes some design considerations, an understanding of imaging in general and an understanding of your specific application needs. This understanding of the design elements, in turn, will affect the decision to design/buy with custom made or off-the-shelf optics. These decisions may have weighted factors and need to be evaluated early in the design cycle to be effective.

The Complexities of Creating High-Power Optical Coatings

High-power thin-film optical coatings are typically required for optics that must handle sustained high-power output from lasers. These coatings can be reflecting, transmitting, polarizing, or beamsplitting; it is important to note that "high power" may have different meanings depending on the application. A reasonable definition is that the term "high power" applies to any coating that requires special attention and processing to avoid damage during irradiation. As a rule of thumb, any design drawing that includes a power specification (that is, for which the standard processing is insufficient) is considered a high-power coating.

The optical coating is generally the limiting factor in the output of a high-power laser system. The most common failure mode of high-power laser coatings results from the presence of absorption sites within the coating or at the coating's interface with the substrate or air. These absorption sites are usually in the form of gross defects that absorb laser energy, resulting in generation of heat that causes localized melting or thermal stress fractures. Failure by this mechanism is usually catastrophic (see Figure 1).

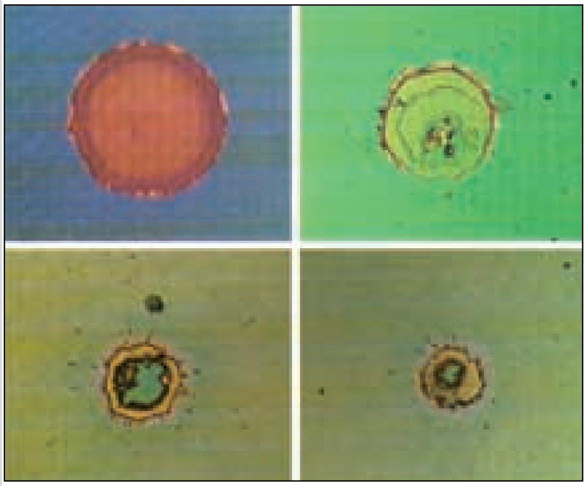


FIGURE 1. *Coatings suffer catastrophic damage when defects absorb laser energy, generate heat, and cause melting or thermal stress fractures. A coating fails at relatively low thresholds of 11.77 (top left), 12.92, (top right), and 14.3 J/cm² (bottom left) for 20ns pulses at a 1064nm wavelength, due to poor coating process control. A coating fails at 73.3 J/cm² due to a coating defect (bottom right).*

Noncatastrophic failure, such as plasma burn, is typically the result of unoxidized 1 to 5μm metallic nodules within the coating. (Some manufacturers will intentionally subject their coated elements to powers sufficient to trigger plasma burns to remove the defect nodules.) Finally, intrinsic material properties determine the laser threshold that an otherwise defect-free film will sustain.

Before deposition

Making high-power laser coatings requires tight control of every aspect of production, from initial substrate manufacture to final packing. Before the substrate even reaches the coating chamber, its surface quality and cleanliness must be assured. A clean coating chamber, appropriate choice of thin-film materials, and good control of process parameters are also essential. After deposition, coating makers must control contamination; even at this stage, surface contamination may cause the element to fail when subjected to high powers. For this reason, meticulous cleaning procedures are also required at the assembly stage, typically under strict cleanroom operating conditions.

Substrates for use with high-power laser coatings must be made of high-quality materials. This is particularly important for transmitting optics—these substrates must demonstrate extremely low intrinsic absorption at the relevant wavelengths. Surface defects are potential damage sites and surface quality is speci-fied in terms of a scratch and dig value (scratch numbers do not directly correlate to scratch size; dig numbers are in units of 0.01 mm). High-power laser optics typically specify 20-10 or 10-5 scratch-dig surface values.

The substrates must also be pristine. Any organic or particulate residue from polishing or cleaning may absorb the laser energy and is therefore a potential damage site. For this reason, the substrate and coating interface is a critical area in achieving high damage thresholds. Mirror elements, however, reflect most of the laser energy from the layers closest to the incident media (normally air), and as a result are less sensitive to the presence of defect sites at the substrate surface than transmissive elements.

continued >

A clean room helps

Either way, this sensitivity to organic or particulate residue tests the cleaning process. Cleanroom conditions help because there is less risk of recontamination after cleaning the substrate. Most coatings companies use lint-free wipes without silicone constituents when cleaning manually in their final clean process. Solvents used are of extremely high purity—typically methanol, isopropanol, or acetone.

Ultrasonic cleaning, when it works efficiently, can be useful and is more effective at dislodging residual polishing compounds than cleaning by hand. Certainly, it is less prone to error.

A typical multistage manual process includes a surfactant wash, several wipes with an ammonia solution, and on the final stage a drag-wipe technique using high-purity solvents. The drag-wipe technique produces very high shear forces, resulting in the removal of any remaining contaminants from the surface.

Contaminants from several parts of a coating chamber can migrate onto the optical surfaces. If the tooling is not meticulously kept clean, it can contaminate the glass. Backstreaming can occur with an inefficient diffusion pump, resulting in organic contamination.

Finally, the walls of the chamber itself can contribute to contamination of the glass. Material evaporated from a target deposits on both the substrate and on the walls of the chamber. After several runs, the material on the walls builds up until it begins to flake off. During the pump-down sequence, loose particulates can be transferred from the walls of a dirty chamber onto the optic.

The solution is to maintain the cleanliness of the chamber. Many chambers are lined with aluminum foil (made by rolling without oil), while other coaters prefer to use removable steel liners. Cleaning the chamber consists largely of replacing the foil or liners and removing coating buildup from any uncovered areas within the chamber.

Design

For high-power applications, coating designers must choose materials with intrinsically low absorption at the relevant wavelengths, which leaves the designer with only a few material choices in each of the spectral regions. Coatings for use with high-power ultraviolet (UV) light are made of different materials from those for use in the visible and near-infrared (IR). Materials for use in mid-and far-IR coatings are a third group.

Dielectric metal oxides are preferred materials for UV, visible, and near-IR laser applications. Silicon dioxide (SiO_2) is the generally accepted and ubiquitous choice for low-index layers. Choosing a material for high-index layers is not as straightforward: oxides of titanium, tantalum, zirconium, hafnium, scandium, and niobium are popular high-index materials.

The design of a coating can significantly alter the damage threshold. In the case of high-reflection coatings, the core structure is typically a repeating stack of high- and low-index layers, each a quarter-wavelength thick. Simply adding a half-wave of low-index material (normally SiO_2) as the final layer can result in measurably higher damage thresholds.

According to some groups, laser-damage thresholds can be increased even further by manipulating the coating layers in at least one of several ways. The electric-field distribution can be averaged across several layers, thereby avoiding a high electric-field concentration within a relatively small number of layers. The high-intensity resonant peaks can be shifted from layer interfaces to locations within the film continuum (see Figure 2). The highest-intensity resonant peaks can be positioned within the layers of the thin-film material demonstrating the highest damage threshold. Reported results for these techniques, however, are mixed.

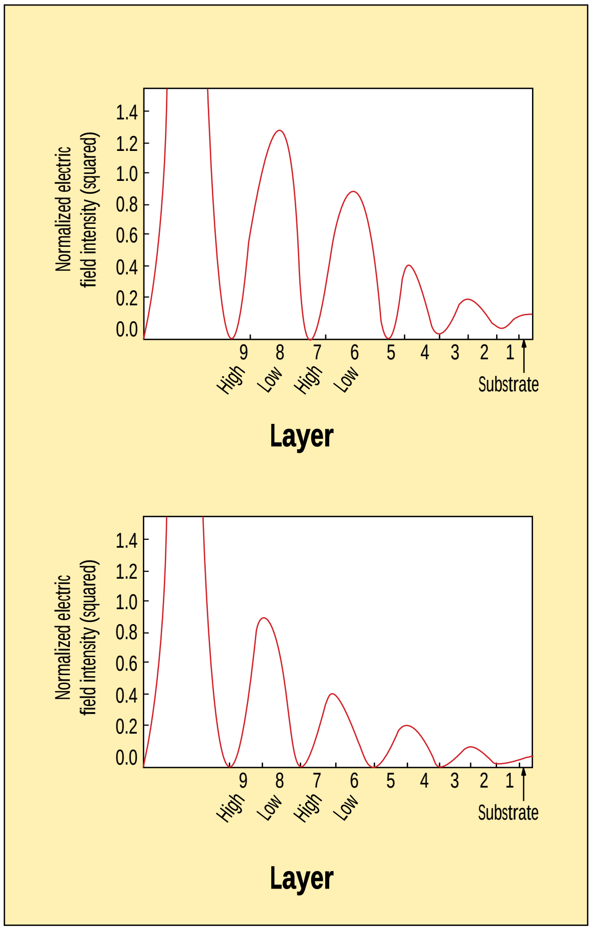


FIGURE 2. *The normalized electric-field intensity (EFI) squared within a reflecting quarter-wave dielectric stack shows peak EFI at layer interfaces and high est EFIs occurring at the layers closest to the air boundary (top). For clarity, the total number of layers shown is a less typical high-reflector design. The thickness of the four layers closest to air in a nine-layer stack is modified to reduce EFI in the high-index layers (bottom).*

continued >

Process control

Many parameters play critical roles in the deposition of a high-power laser coating, including the rate of deposition, substrate temperature, oxygen partial pressure (used in designs including dielectric metal oxides), thickness calibration, material-melt preconditioning, and electron-gun sweep. A poorly controlled evaporation process produces spatter from the source, resulting in particulate condensates on the substrate surface and within the depositing coating. These condensates are potential damage defect sites. Unfortunately, some materials that can be used for high-damage-threshold coatings are difficult to deposit smoothly. The settings applied to the electron-gun sweep can be the difference between the production of a clear, high-damage-threshold coating or the production of a high-scatter coating with a much lower power capability.

The rate of deposition, substrate temperature, and oxygen partial pressure (for dielectric oxides) determine the stoichiometry of the growing film, which significantly affects the metal oxide chem-istry in the depositing film. These parameters must be optimized and controlled to produce a homogeneous layer with the desired metal-oxygen content and structure.

In producing antireflection coatings, thickness accuracy of the depositing films is an important factor in meeting the desired low reflectance. Mirrors are generally less sensitive to small thickness errors as a result of the relatively broad reflectance band afforded by the refractive index ratio of the high and low index layers. Deep-UV mirrors are an exception, however, because material limitations in this spectral range produce relatively narrowband reflectors.

Fabricating high-power thin-film optical coatings challenges every step of the manufacturing process and requires great attention to cleanliness.

Ion beams

Ion-beam technology is now a recognized and widely used tool in the manufacture of thin-film coatings, either as an enhancement to thermal evaporation (ion-assisted deposition) or as a sputtering technology (ion-beam sputtering). While these methods produce more compact and durable films with properties closer to those of the bulk materials, conclusive evidence may not exist that ion-beam technology produces higher damage thresholds.

TECH TIP ON ANTIREFLECTION COATINGS

As light passes through an uncoated glass substrate, approximately 4% will be reflected at each surface. This results in a total transmission of only 92% of the incident light. Antireflection coatings are especially important if the system contains many transmitting optical elements. Coating each component will increase the throughput of the system and reduce hazards caused by reflections traveling backwards through the system (ghost images). Many low-light systems incorporate AR coated optics to allow for an efficient use of the light.

The transmission properties of a coating are dependent upon the wavelength of light being used, the index of refraction of the substrate, the index of refraction of the coating, the thickness of the coating, and the angle of the incident light.

The coating is designed so that the relative phase shift between the beam reflected at the upper and lower boundary of the thin film is 180°. Destructive interference between the two reflected beams occurs, cancelling both beams before they exit the surface. The optical thickness of the coating must be an odd number of quarter wavelengths ($\lambda/4$, where λ is the design wavelength or wavelength which is being optimized for peak performance), in order to achieve the desired path difference of one half wavelength between the reflected beams, which leads to their cancellation.