Meaningful surface roughness and quality tolerances

David M. Aikens
Savvy Optics Corp, 35 Gilbert Hill Rd., Chester, CT 06412
daikens@savvyoptics.com

ABSTRACT

Most tolerances on optical elements can be derived or calculated from the application requirements using computer-aided optical design programs. For surface quality and surface roughness, however, there are few guidelines or tools for calculating appropriate tolerances. Typically, we simply use a legacy specification (e.g. 60-40 and 3 A RMS) with little thought for either the cost of achieving the specification or the penalty for failing to achieve it. Often these legacy specifications are ambiguous, unnecessarily costly and in some cases completely meaningless. This paper provides some basic rules and equations for calculation of the real or perceived impact of these specifications, and some guidelines for the initiate (and for some of us veterans as well) as to how to compose a meaningful tolerance.

Keywords: roughness, texture, quality, scratch-dig, scratch, imperfection, specification, standard, training

1. INTRODUCTION

When I started designing lenses in the late 70's and early 80's, there was very little to choose from in the way of optical drawing formats. MIL-STD-34, published in 1960, which is a notes-based format, was used in the US and much of the world. It survives today in spirit, if not in practice in much of the optics community. In a MIL-STD-34 drawing, surface quality is specified according to MIL-O-13830 (1954) and surface roughness according to MIL-STD-10 (1949).

Times have changed; standards have come and gone. In 1978 MIL-STD-10A was withdrawn and replaced by a voluntary equivalent, ANSI B46.1, which itself has been revised so much that it bears little resemblance to the original standard. In 1986 MIL-STD-34 was also withdrawn and replaced by a voluntary equivalent, ANSI/ASME Y14.18M (1986) which has since been withdrawn without replacement. And in the past 20 years, MIL-O-13830A has been inactivated and then reactivated three times, but still lives on. Dozens of new drawing standards have been released, many of which use the same notation, but interpret the notation differently. Meanwhile, optics manufacturing has progressed from pitch and slurry to diamond machining and CNC systems. Metrology has moved from visual check to megapixel interferometers, and from a simple stylus to an automated areal optical profiler. Surface texture has gone from an afterthought to a real concern, and surface quality has become a quagmire of mis-interpretation. In the 21st century, we know a lot more about surface quality and surface roughness, and how these parameters affect performance, or don't, in most optical systems. And yet we are still writing our texture and quality specifications the same way we were in the 1970's. We can, and should, do better.

2. SCRATCH AND DIG IS A COSMETIC STANDARD[1]

The surface quality test of MIL-PRF-13830B (the successor to MIL-O-13830) is based on a visual comparison, under specific darkfield lighting conditions, of a subject surface imperfection and a comparison standard set to determine the visibility or "grade" of the imperfection. The specification references a drawing for surface quality standards, C7641866. This drawing, in turn, references a set of master scratches kept at Picatinny Arsenal, and provides a range of polarization angles for each of the scratches, based on a micro-image comparator. Both the SIRA built micro-image comparator and the master scratches are kept at Picatinny Arsenal, which supplies the comparison standards to their suppliers.

This approach to surface imperfections was first proposed by McLeod and Sherwood in 1945[2]. They offered up comparison standards numbered from 10 to 120, to be used in this comparison method. As early as 1945 they recorded that "there is little correlation between the appearance or visibility of a scratch and its measured width." Frankford Arsenal documents dating to the same period declare that "these numbers are arbitrary, and are not to be assumed as denoting the width of the scratch." Scratch morphology is a better predictor of scratch visibility or brightness, than width[3]. Since the purpose of the scratch and dig standard is to control cosmetic imperfections, and the primary criterion of a cosmetic imperfection is its visibility, this is a perfectly reasonable approach, and continues to be used by the US Armed Services to this day. While dig comparison sets, based on a particular size and shape, are relatively easy to manufacture and calibrate, the scratch visibility standard is highly subjective. The master scratch set consists of a set of five pairs of scratched pieces of glass. Each set corresponds to one scratch number, and represents the miminum and maximum visibility for the comparison standards, under the illumination conditions established by MIL-PRF-13830B. These scratches are the master set from Frankford Arsenal, and are still being used to certify sub-master comparison standard sets which are sent to the field to be used in inspections. Today, they are at Picatinny Arsenal in New Jersey, and have remained more or less intact and virtually unchanged in 50 years.

In the 1970's, though, a series of enigmatic revisions were made, not to the MIL specification, but to the drawing C7641866, which has created an enormous amount of confusion in our industry. In 1974, revision H of the drawing specified a comparison set wherein the scratch number was to be the width of the scratch in microns. To make matters worse, in 1976 revision J described a comparison set wherein the scratch number was the width of the scratch in tenths of microns. Finally in 1980, revision L made all such width notes for reference only, although in fact the scratch widths are meaningless; scratch width and visibility are uncorrelated. In all this time, the limit masters and the meaning of, say, a #60 scratch, remained unchanged. As a result of these unfortunate revisions, however, there have been many myths and legends in our industry regarding the scratch standard. But the simple truth about the scratch standard is that:

1) The scratch-dig standard is, and has always been, a visibility standard, not a width standard.
2) The scratch number is not the width in microns or tenths of microns.
3) The Army never tightened the scratch specification by 10x.
4) Scratch standards do not "heal" over time. (They do, however, get dirty and need to be cleaned.)

Because of this series of unfortunate events[4], however, and the desire of the industry to have an objective, rather than subjective surface quality standard, the scratch and dig standard has become the most mis-used, ambiguous, and mis-interpreted specification in our industry[5]. Most of these problems can be sorted out, however, by changing the way we call out our surface quality specification.

3. MEANINGFUL SURFACE QUALITY SPECIFICATIONS - COSMETIC

If you are trying to specify the cosmetic quality of your optics, then the scratch-dig number system still works for you. The trick is to specify a scratch and dig that fits your application. Most people are not actually familiar with the visibility of the scratches in the scratch-dig standard. Table 1 is a qualitative assessment of scratch visibility by number[6], and the most common scratch-dig pair for a given level of desired visibility.

Table 1. Scratch visibility by number, and some common applications

Scratch number	Scratch visibility under normal viewing conditions	Typical callout	Application
10	Very, very hard to see	10-5	Not recommended; see below
20	Barely visible	20-10	Not recommended; see below
40	Almost the same as the #20	40-20	Close to image plane, or dark-field illumination
60	Somewhat visible	60-40	Precision Optics
80	Easy to see	80-50	Commercial Optics

I don't recommend using anything tighter than 40-20, and for most applications 60-40 is more than adequate. Even though the #40 scratch is barely visible in normal conditions, there are cases where the lens will be used with significant magnification, or where the surface will be illuminated by a bright light and viewed in dark-field, and the cost of a 40-20 specification is justified. But if your application is sensitive enough to surface imperfections to specify anything tighter, then you are using the wrong standard. More on that in Section 4, below.

For cosmetic applications, though, rather than use the MIL-PRF-13830B, it is wiser and safer to use the voluntary equivalent, ANSI/OEOSC OP1.002[7]. This new standard is still being maintained, is less ambiguous, is not at risk of being withdrawn, and has far less confusion associated with it than the MIL standard. Since it uses the same comparison standards as the MIL, the meaning of the numbers is the same. The language to invoke this standard is "Surface quality 60-40 per ANSI/OEOSC OP1.002." You will want to also specify the comparison standard to be used.

Discounting the ones claiming to offer width-based standards (which is just plain wrong) and chrome on glass reticles (which don't look like scratches, IMHO), there are three commercial manufacturers of comparison standards; FLIR/Brysen, Davidson Optronics, and Jenoptik. This last is the plastic paddle sold by both Edmund Optics and Thor Labs. I'm not sure how repeatable any of these standards are, but I suspect each company does a decent job of certifying them to whatever internal masters they have. Remember though, that only Picatinny Arsenal has the SIRA tool or the limit masters, so all claims to make standards to C7641866 should be taken with a grain of salt. In some cases, these sets are not mutually compatible. The #10 artifact of one brand could be brighter than the #60 artifact of another[8]. So the meaning of the visibility of a scratch is dependent on the comparison samples being used. For all commercial procurements, the brand of the comparison standard should be referenced on the drawing. For example, "Surface quality 80-50 per ANSI/OEOSC OP1.002 using brand x comparison standard," is a reasonably unambiguous approach to a cosmetic surface quality requirement.

One word of caution; adding a complex dielectric or metal coating to a substrate can change or amplify the visibility of any given scratch or dig. I have seen cases where substrates that passed a 40-20 spec just fine couldn't pass a 60-40 spec after coating.

4. MEANINGFUL SURFACE QUALITY SPECIFICATIONS - DIMENSIONAL

Some applications require a width specification, or are sensitive to scratches too small to see. In 1985, there were no practical alternatives to scratch and dig in the US. Lots of people have tried to extend the usefulness of the scratch-dig standard by adding higher brightness light sources and magnification, which has caused more confusion. In 1996, with the publication of ISO 10110, there was an alternative. Based on the German DIN 3140 standard, part 7 describes all surface imperfections in terms of a "root area", and not only allows but practically requires direct measurement of every scratch and dig. Lately this approach has been gaining momentum, but aside from a few stalwarts, it has not become common practice in the US.

Even when using this international standard, some careful attention is required. Although the notation is clear and unambiguous, a strict interpretation of the standard requires microscopic inspection and documentation of every surface imperfection greater than 0.16x the size of the maximum allowed imperfection. This kind of inspection can be much more expensive than the optic being inspected, and is usually far more than the engineer had really wanted.

The newest revision of the American National Standard for surface imperfections, OP1.002, retains the visibility method from MIL-PRF-13830B discussed in Section 3, but adds the dimensional method from MIL-C-48497A, which is familiar, easy to use, and effective for people who really require a functional specification. To invoke the dimensional method, a pair of letters is used, such as A-A or E-E, to reference a specific size of imperfections. This new notation is gathering a following in the United States, especially in applications which demand very tight tolerances on surface imperfections, such as in micro-optics or objectives, CCD cover glass, reticles, and high energy laser applications, or on applications which are tolerant of large surface imperfections but are on reflective or partially reflective optics like IR imaging. Thanks to the specification of a precise dimension, the standard is objective, but because grades are used, visual inspection for most optics is sufficient. On marginally acceptable optics with fairly tight specifications, microscopic inspection and documentation is still required to determine if the part passes the accumulation specification.

Table 2 shows the translation from a given scratch width and dig diameter specification to the correct specifications per OP1.002. For grades that are not in the table, a new notation is provided. Any size surface imperfection can be specified using the letter "A" followed by a number. The number indicates the maximum size imperfection in microns. That is, a specification A2-A20 has a maximum allowed scratch width of 2 micron and a maximum dig diameter of 20 micron. If such a non-standard specification is required, some thought should be given to how the supplier is expected to inspect and measure such imperfections. Providing even minimal guidance on the drawing such as "Inspect with 40x stereo microscope and 150 Watt tungsten halogen fiber light source" can go a long way to eliminating unnecessary confusion and cost.

Table 2. Translating scratch width and dig diameter to correct specifications per OP1.002.

Maximum scratch width in microns	Scratch specification letter	Maximum dig diameter in microns	Dig specification letter
120	G	700	G
80	F	500	F
60	E	400	E
40	D	300	D
20	C	200	C
10	B	100	B
5	A	50	A
n	An	n	An

Alternatively, a surface quality specification can be derived based on total scattered light. Care must be taken to estimate the worst possible scenario for accumulation, and assumptions must be made regarding obscuration and scatter, since line width does not correlate well to brightness. If we assume that all scratches and digs are 100% obscuring, and that the part is made to result in the worst possible amount of obscuration, we can calculate the amount of obscured light based simply on the relative area of the surface and the accumulation of all the scratches and digs. Table 3 shows this calculation. It is important to note that, although the allowed number of digs and lengths of scratches scales with part diameter (more or less), the area of the surface is scaling with the square of the diameter. So the impact of a given surface quality specification falls with part size.

Table 3. Calculating the maximum allowed total obscured area by specification and part size.

	Percent obscured area by surface diameter in mm				
	10	20	30	40	50
A-A	0.132%	0.033%	0.022%	0.017%	0.013%
B-B	0.275%	0.069%	0.047%	0.034%	0.028%
C-C	0.589%	0.147%	0.103%	0.074%	0.061%
D-D	1.199%	0.300%	0.210%	0.150%	0.123%
E-E	1.848%	0.462%	0.326%	0.231%	0.191%
F-F	2.537%	0.634%	0.451%	0.317%	0.264%
G-G	4.036%	1.009%	0.727%	0.504%	0.423%

5. SURFACE TEXTURE; THE OTHER FORGOTTEN SPECIFICATION

Surface texture, or roughness, is another area where new specifications improve our options. MIL-STD-10A, which was a very good standard for the time, was not for polished surfaces *per se,* and focused primarily on average roughness, a parameter that is uncommon in the optics community. Today the successor document, ASME B46.1-2002, has grown to some 110 pages, and now includes notations for more than 40 parameters, including RMS, slope, skew, kurtosis, and my personal favorite, PSD. The latest revision adds the areal versions of all these specifications, which is a huge breakthrough for power users who need to specify the 2D PSD and RMS amplitudes by spatial frequency for (for example) diamond turned surfaces. But for most applications, all you need is an RMS surface texture specification.

6. MEANINGFUL SURFACE TEXTURE SPECIFICATIONS - SCALE LENGTHS

There as several good papers on the effects of roughness and waviness[10-13]. Unfortunately not all of them are still in print, such as Bennett and Mattson's excellent book. But the equations required to determine a meaningful roughness specification are fairly simple. High spatial frequency surface texture, or roughness, scatters light. The amount of light scattered is proportional to the square of the RMS roughness (assuming a lot of stuff not worth getting into here), and the angles the light is scattered into is determined by the spatial bandwidth of the roughness.

Most people are unaware of the importance of spatial bandwidth (or trace length or window size, on unfiltered data) in evaluating roughness. Simply put, a roughness without a spatial bandwidth is meaningless. For almost all smooth, well-behaved surfaces, if you measure a larger area, the RMS surface error is higher; if you measure a smaller area, the RMS error will be lower. This is shown quite clearly in the references, and is a direct consequence of Fourier Transform mathematics and Parsival's theorem. So the next time someone tries to sell you their super-polished capability with a 100 micron AFM trace, ask for a long-trace profiler measurement instead.

In reading MIL-STD-10A, it is clear this was well understood by the authors, who were concerned about the meaning of a roughness number without a scale length. So if the user did not specify scale lengths, it was assumed that the roughness was measured with a stylus, and that the scan length was 0.03 inches, or about 0.8 mm. This works for applications where we are not so concerned about where the light goes after it scatters from the surface due to our surface roughness. Unfortunately, the newer standards which replaced MIL-STD-10A, such as ASME B46.1-2002 and ISO 10110 Part 8 do not have defaults or if they do, they are not the same as they were for MIL-STD-10A. The result is significant ambiguity, and in some cases, gamesmanship on the part of suppliers. So to have a meaningful roughness specification, we must now indicate the spatial scale length explicitly on the drawing. For example, to invoke the old MIL-STD-10A limit, we need to write "Surface roughness per B46.1-2002 Rq < 3 nm for spatial scale lengths shorter than 0.8 mm." This can be shown pictographically in an ISO 10110 drawing (figure 1) or using the symbology of ASME Y14.36M-1996.

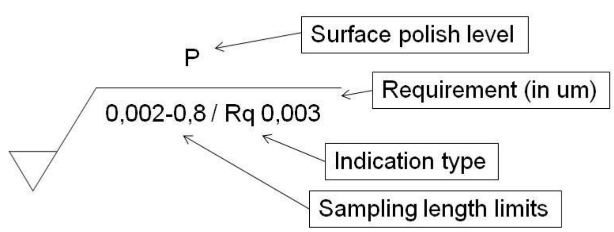

Figure 1. Example of surface texture specification using ISO 10110-8 (2010) to re-create the MIL-STD-10A definition of "30 Angstrom RMS surface roughness".

In the cases where we do care where the scattered light goes, we need to consider selecting our cutoff spatial scale length. It has been shown elsewhere[11] that the scattered light follows the grating equation (again with a bunch of assumptions.) That is,

$$\theta_s = sin^{-1}\left(\frac{\lambda}{L}\right) \quad (1)$$

Where L is the spatial period of interest and λ is the wavelength of light. The shorter the spatial period, the higher the scattering angle. So our roughness can be assumed to scatter light into a host of angles related to the spatial periods covered by the roughness specification. For visible light (say λ = 0.56 micron), the default spatial bandwidth from MIL-STD-10A equates to roughness which scatters light at angles greater than 0.04 degrees. For many applications, this is perfectly appropriate. For other applications, this near-angle scatter may be far more egregious than simple surface texture, and multiple specifications for various spatial bandwidths of surface texture (e.g. Roughness and Waviness) may be required.

7. MEANINGFUL SURFACE TEXTURE SPECIFICATIONS - AMPLITUDE

The amount of light scattered into the angles defined by the spatial periods is related to the amplitude of the surface texture. For a well-behaved polished surface of a lens, the fraction of light scattered (S) is given by equation (2), where n is the index of the material, σ is the RMS roughness, and λ is the wavelength of light[9]. (For mirror surfaces, set n equal to -1).

$$S = \left(\frac{2(n-1)\pi\sigma}{\lambda}\right)^2 \quad (2)$$

From this we can see a direct meaning to our roughness value. Scatter goes as the square of the ratio of the roughness and wavelength. For example, a Silicon surface (n = 3.42) which has an RMS roughness of 10 nm over the scale-lengths of interest and used in the MWIR band (λ = 4 micron) will scatter 0.1%; not a significant factor. The same surface used as a mirror in the visible (n = -1, λ = 0.56 um) would scatter 5% of the incident light. This is why residual surface texture errors are not a problem in the IR but are often a concern in the visible, and become extremely important in the UV.

We can therefore calculate a meaningful roughness tolerance for an allowable level of scatter in a lens system. In table 3, we show various roughness specifications for various scattering requirements for a typical visible application, assuming we have ten surfaces in a lens (n = 1.7, λ = .56 um), such as a double Gauss. Clearly, specifying anything below about 1nm RMS is unnecessary for most visible applications.

Table 3. The RMS roughness amplitude corresponding to the allowable level of scatter, assuming a visible band lens, n = 1.7, l = 0.56, and ten surfaces.

Total Loss	Scatter per Surface	RMS, nm
0.001%	0.0001%	0.1
0.02%	0.002%	0.5
0.05%	0.005%	0.9
0.25%	0.025%	2
1.0%	0.10%	4
1.5%	0.2%	5
3.0%	0.3%	7
5.0%	0.5%	9

One must be a little careful here about coatings and roughness. Just as with surface imperfections, applying a complex dielectric coating can some-times exaggerate the effects of an underlying surface texture. This usually isn't an issue, but when it is, it can be quite problematic.

8. SUMMARY

For some applications, very careful attention needs to be applied to determine the appropriate surface quality or roughness specification. But for most applications, these specifications are intended more to control workmanship and prevent unwanted artifacts than to serve as functional tolerances. Nevertheless, it is irresponsible of us to simply write down an incomplete or ambiguous specification in the 21st century. This paper has provided the basic rules and mathematics that will allow the optical engineer to write a meaningful tolerance on surface quality and roughness, with just a few minutes of thought.

9. REFERENCES

[1] Young, M., "The scratch standard is only a cosmetic standard," Proc. SPIE Vol 1164, 185-190 (1989).

[2] J.H. McLeod and W.T. Sherwood, "A proposed method of specifying appearance defects on optical parts," J. Opt. Soc. Am. Vol 35, pp 136-138 (1945).

[3] G. White and J. Marchiando, "Scattering from a V-shaped groove in the resonance domain", Appl. Opt. Vol 22, No 15, pp 2308-2312 (1983).

[4] R. E. Fischer, "An interview with Matt Young on the scratch and dig standard for optics," OE Reports, Vol 44 (1987).

[5] Aikens, D. M., "Understanding Scratch and Dig Specifications," Course Notes for SPIE SC700, Rev 8 , 19-27, (2010).

[6] From discussions with Allen Krisiloff, Gordon Boultbee and others at the OEOSC ASC/OP1 TF2 Meeting, August 11, 2008.

[7] ANSI/OEOSC OP1.002-2009, "American National Standard For Optics and Electro-Optical Instruments-Optical Elements and Assemblies-Appearance Imperfections." American National Standards Institute, Inc. (2009).

[8] Aikens, D., "The truth about scratch and dig", OF&T technical digest, Optical Society of America (2010).

[9] Lewis, I. T. et. al., " Stray-light implications of scratch/dig specifications", Proc. SPIE Vol 1530, pp 22-34 (1991).

[10] J. M. Bennett and L. Mattsson, *Introduction to Surface Roughness and Scattering,* Optical Society of America, (1989).

[11] J. R. Harvey, A. Kotha, "Scattering effects from residual optical fabrication errors", Proc. SPIE Vol. 2576, pp.155-174 (1995).

[12] R. N. Youngworth and B. D. Stone, "Simple estimates for the effects of mid-spatial frequency surface errors on image quality", Applied Optics **39**(13), 2198-2209 (2000).

[13] Aikens, D., DeGroote, J., and Youngworth, R., "Specification and control of mid-spatial frequency wavefront errors in optical systems", OF&T technical digest, Optical Society of America (2008).