Standards for Visual Acuity

A Report Prepared For

Elena Messina
Intelligent Systems Division
National Institute for Standards and Technology

In Support Of
ASTM International Task Group E54.08.01 On
Performance Measures for Robots for Urban Search and Rescue

Prepared by
John M Evans LLC
Newtown, CT
john_evans@snet.net

Under Contract 2006-02-13-01 Through
KT Consulting

Standards for Visual Acuity

1. Summary

Visual acuity is defined as the ability to read a standard test pattern at a certain distance, usually measured in terms of a ratio to “normal” vision. Multiple standard test patterns have been developed and are in use today.

Closely related are measures of image quality or imaging or printing system quality, generally referred to as optical resolution. Again, standard test patterns can be used, or the modulation transfer function (MTF) can be measured. The MTF defines the ability of an imaging system to reproduce a given spatial frequency; the MTF of components of a system can be multiplied to give the MTF of the composite system.

The problem for defining the quality of vision systems for robots for application in Urban Search and Rescue (USAR) involves viewing a scene with a camera (which includes a lens and sensor chip and digitization circuitry), transmission of the digital data and conversion to an image on a screen that can be viewed by the operator. Hence, either approach is relevant. We will discuss both approaches with emphasis on the first.

The recommended approach is to use the “Tumbling E” eye charts for evaluating visual acuity of robots for urban search and rescue applications.

2. Standard Eye Charts

Attempts to define visual acuity in quantitative terms using eye charts date to the early 1800’s in Germany. The term “visual acuity” itself dates to Donders in 1861 who defined it as the “ratio between a subject’s performance and a standard performance” in distinguishing details of a test pattern. Snellen published his famous eye chart in England in 1862 and only relatively minor variations and improvements have been made since then.

2.1. The Snellen Chart

Prior to Snellen’s work, eye charts had used printing fonts. Snellen defined a new font, which he called “optotypes” and which he laid out on a 5 x 5 grid [1,2]. Using the standard of dividing a degree into 60 minutes (the use of base 60 actually dates to the Babylonians), he defined “standard vision” as the ability to recognize his optotypes from a distance of 20 feet when they subtended an angle of 5 minutes of arc. The detailed features of his optotypes, one grid element, were then 1 minute by 1 minute of arc for “standard” vision. Standard vision is thus the ability to distinguish features separated by 1 minute of arc; eye charts generally are scaled to be used at a distance of 20 feet.

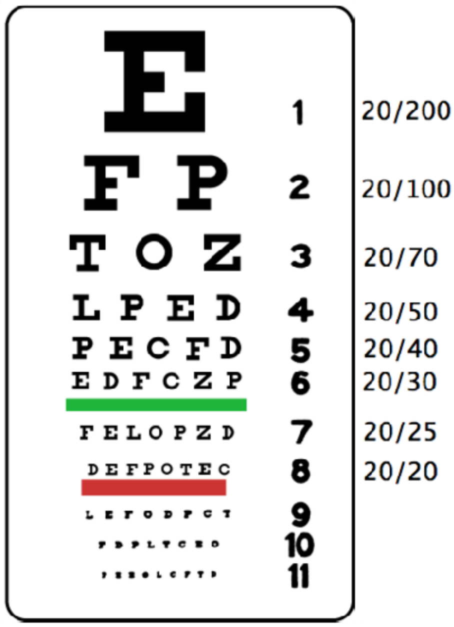

Figure 1: Snellen Eye Chart

Snellen's charts were labeled by the ratio of performance to "standard" performance, so we are all familiar with the use of 20/20 vision, 20/40 vision, and so on. The choice of 20 feet as a measurement distance was arbitrary. Given his preference for round numbers such as 1 and 5 for the optotypes, one might imagine that this was the distance near the largest dimensions of his office or laboratory where he could read his standard size optotypes that was a nice round number, although this is only speculation.

It turns out that 20/20 is not perfect human vision. Indeed, it is near the average for adults in their 60's as vision degrades. Good vision in young adults with no visual impairment is generally between 20/16 and 20/12, much better than 20/20. 20/20 vision has come to be interpreted as the limit of "normal" vision with which an individual can cope well enough in school or industry and hence does not require correction. Vision beyond 20/20 is generally improved with corrective lenses.

Note that 20/12 corresponds to an angular resolution of .01 degrees. If Snellen had used decimal fractions rather than the established convention of dividing a degree into 60 minutes, standard vision might be defined quite differently. The application of angular resolution to relevant applications is explored in the appendix.

Snellen's optotypes included only 9 letters with strong serifs: C, D, E, F, L, O, P, T, and Z. In general these are easy to distinguish, although F and P can be mixed near the limits of resolution.

In 1875 Snellen created a new set of charts that used six meters instead of twenty feet as the "standard" measurement distance. The Snellen fractions were then 6/6, 6/12, and so on. Undoubtedly he was criticized by proponents of the English system of measurement, who fought doggedly against the French-backed metric system [3], but metric measurement prevailed and 6 meters is the standard in England today while 20 feet is in use in the United States.

Monoye in 1875 proposed to change the Snellen fractions to a decimal to make it easier to compare values irrespective of the original measurement distance. The decimal value of 0.5, for example, could be derived from 20/40 or 6/12 or similar results using any other "standard" measurement distance. This points out how arbitrary the 20 feet or 6 meter distance is. Despite the logic, the Snellen fractions remain in common use today.

Another equivalent measure of visual acuity is the logarithm of the minimum angle of resolution, LogMAR. This converts the progression of the Snellen chart to a linear scale. 20/20 vision is thus 0.0, 20/40 vision is 0.3, 20/100 is 0.7 and 20/200 is 1.0. The Snellen chart intervals are not a geometric progression (i.e. there is not a constant ratio of letter heights between adjacent lines) so the lines on a typical Snellen chart (20/200, 100, 70, 50, 40, 30, 25, 20, 15, 10) do not produce equal intervals on the LogMAR scale.

2.2. Variations on the Snellen Chart

Louise Sloan in 1959 defined a new set of 10 optotypes without serifs on a 5 x 5 grid, with the “standard” size to be used at 1 meter. She also proposed the approach of using all ten letters on each row which standardized the effect of crowding between letters and avoided any problems of some letters being more recognizable than others and made letter size the only variable measured. The top line of these charts is much longer than the bottom line, so the chart is an inverted triangle. In fact the largest letters were printed on more than one line. Many modern versions of the Snellen chart use Sloan optotypes.

In 1868 Green[4], who had worked with Snellen, proposed a chart with a geometric progression of letter sizes and proportional spacing between letters. This approach was re-invented by Bailey and Lovie in 1976 [5] using British letters (4 x 5 grid) and a six meter test distance.

In 1982 the National Eye Institute adopted a chart used in the Early Treatment of Diabetic Retinopathy Study[6], which used the Bailey and Lovie layout with Sloan optotypes and a standard test distance of 4 meters, back illuminated to a calibrated light level, and which required a detailed protocol of counting each correctly identified letter. This chart, called the ETDRS chart after the initial study, and the associated protocol is cited by the International Council of Ophthalmology as the “gold standard” for visual acuity testing and is used as a research standard.

The Snellen chart, with either Snellen or Sloan optotypes, remains the common standard in schools, businesses and optometrists’ offices.

2.3. Alternative Approaches

2.3.1. Landolt C

Not all of Snellen’s optotypes are equally recognizable. Landolt in 1888 addressed this problem by proposing an eye chart that had only one symbol, a ring with a break at top, bottom, left or right, and 45 degree positions in between, basically the letter C in various orientations. To match Snellen’s results, the “standard” size of the C was 0.35” (which subtends 5 arc minutes at 20 feet) with a gap of 0.07” or 1 arc minute [7].

Figure 2: Landolt C in various sizes and orientations

The International Council of Ophthamology considers the Landolt C the purest research standard and requires all other research approaches to be calibrated against the Landolt C [8].

2.3.2. Lea Test

The Lea test was developed in 1976 for testing preschool children and is named after the inventor, Lea Hyvärinen of Finland [9]. She used a set of pictorial optotypes that are symbolic outlines of an apple, a house, a square and a circle. Various versions for testing near vision, far vision, contrast sensitivity, amblyopia and brain damage have been used.

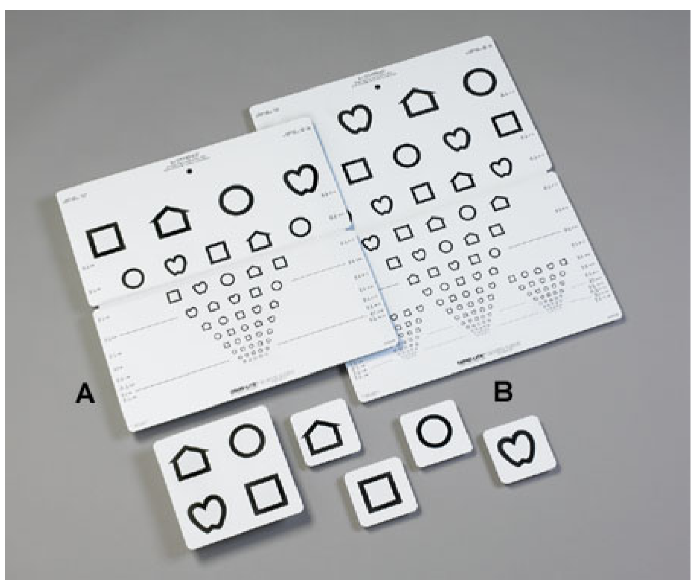

Figure 3: Lea Test Cards

2.3.3. HOTV Chart

Another test for preschool children uses the HOTV chart and HOTV cards [10]. A set of four optotypes, *sans-serif* versions of the letters H, O, T, and V, are used. These optotypes are represented to be equally recognizable and maximally distinguishable. A child is first taught the four symbols using the flash cards and then tested against the eye chart.

Clinical tests have shown that the Lea Test and HOTV tests provide similar results for 4 and 5 year old children but that testability was better using the Lea symbols with 3 year olds [11].

Figure 4: HOTV eye chart and flash cards

2.3.4. Tumbling E's

In 1976 Taylor created a chart using a single optotype, a stylized letter E, in various orientations to test visual acuity of Australian Aborigines [12]. This has become standard for testing of illiterates and populations not familiar with the Roman alphabet.

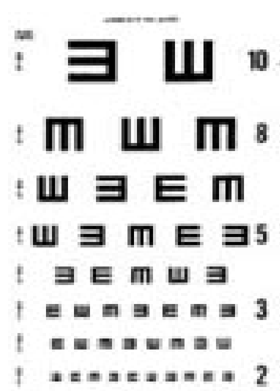

Figure 5: Tumbling E Eye Chart

Taylor's E's are basically a truncated grating with three bars. Multiple studies have shown that grating recognition and orientation is, in general, superior to letter recognition in both foveal and peripheral vision [13]. This can be understood in terms of sampling theory. When too much energy is in higher spatial frequencies, above the Nyquist frequency, letters cannot be distinguished. Letters in general throw more energy into multiple higher spatial frequencies than comparable truncated gratings.

A further study [14] showed that visual acuity thresholds for the Tumbling E were slightly better than the Landolt C, particularly in the presence of any astigmatism. This study concluded that more confidence can be placed in the visual acuity thresholds obtained with the Tumbling E.

Hence, of all the eye charts, the Tumbling E is considered the best choice. It has the further advantage of being easy to produce and use.

3. Test Patterns and Measures of Image Quality

Evaluating the performance of lenses, still cameras, video cameras, television sets, computer monitors and printers requires the quantitative measurement of image quality or optical resolution. This is closely related to visual acuity measurement in which the optical instrument is the eye and the image is created on the retina and processed in the visual cortex.

Two distinct approaches have been used. The first is the use of standard images, which are interpreted by a person, and the second is the use of the modulation transfer function. We will deal with standard images first.

3.1. Spatial Frequency and Lines of Resolution

Spatial frequency is the number of cycles per unit distance, e.g. 10 lines per mm or 100 lines per inch. By Fourier analysis, small patches of complex spatial images can be decomposed into factors at different spatial frequencies that sum together to create the total image patch. The ability of the system to reproduce high spatial frequencies will determine the sharpness or resolution of an image. There is some spatial frequency at which images of closely spaced lines blur together and cannot be distinguished. This is the called the cut-off frequency and the system cannot reproduce higher frequencies. Incoherent illumination is assumed in measuring cut-off frequencies.

Image resolution is often measured not in terms of spatial frequencies but rather in terms of the number of lines that could be resolved across the entire image. The definition is the number of lines across a circle inscribed in the image, so it is the smaller of the dimensions of a rectangular image. So for a normal television picture, with a 4:3 aspect ratio and horizontal scanning, it is the number of horizontal lines that could be resolved that is the measure of resolution. Standard NTSC provides 525 lines scanned, of which 480 horizontal lines are used for the image, but the resolution within a horizontal line is only 440 dots, which by definition provides ¾ (440) or 330 lines of resolution and a picture after transmission and reproduction is around 300 lines resolution. [15]

Most test patterns have sections of grid patterns to test resolution. Most also have grid patterns that taper as wedges, where the ratio of black line to white space is constant but the width of the wedge decreases. At some point the individual lines become indistinguishable, and test patterns are labeled with numbers that are the total number of lines at that spatial frequency that could be distinguished across the whole image and the point at which the lines become indistinguishable is the measure of image resolution. The cut-off frequency (lines per mm) times the image size (in mm) yields the lines of resolution (lines across the whole image). Clearly, if a picture was ever processed digitally, the lines of resolution cannot exceed the number of pixels.

3.2. EIA Resolution Chart 1956

The Electronic Industries Association published the standard resolution chart shown in Figure 6. This pattern was developed in the era of black and white television and is still used for testing cameras.

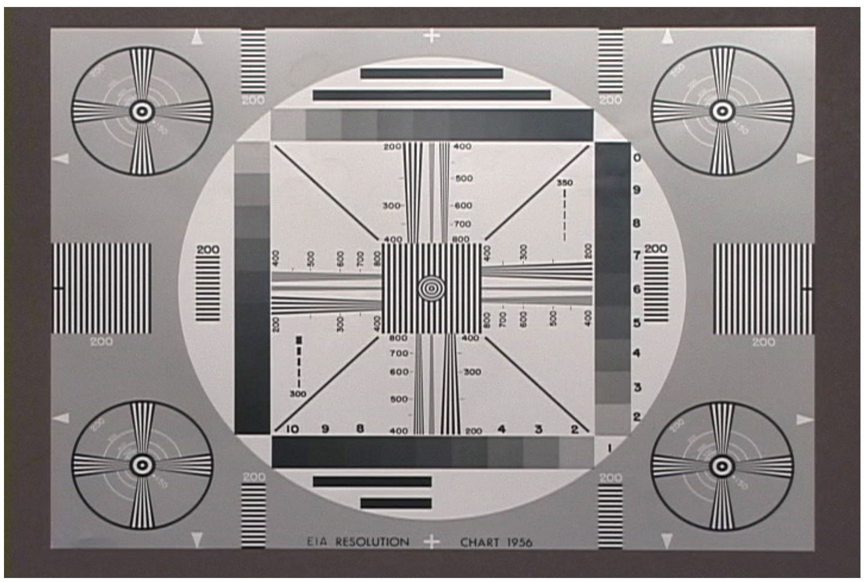

Figure 6: EIA Resolution Chart 1956

Note the use of wedges at multiple points on the chart. When the individual lines of the wedges become indistinguishable, the numbers next to the wedges give the number of resolved lines across the whole image. For broadcast television, this was typically about 300, although studio monitors could resolve all 480 scanned lines. The vidicon tubes and amplifiers and broadcast signals were all analog, so each stage in producing and transmitting and reproducing an image produced further degradation of the image. HDTV provides 1018 lines in an image with digital processing.

There are various proprietary reference target charts based on this chart that add capabilities for testing and adjusting color video cameras. A complete set of charts is sold by Vertex Video, for example, that includes color bar chart, a grey scale chart, a version of the EIA chart above, and charts for checking registration and focus and zoom. Other commercial charts replace the grey scales in the EIA chart with color bars.

3.3. ISO 12233

A more complex and detailed resolution chart was developed for HDTV cameras and promulgated as ISO 12233 [17] which is shown in Figure 7.

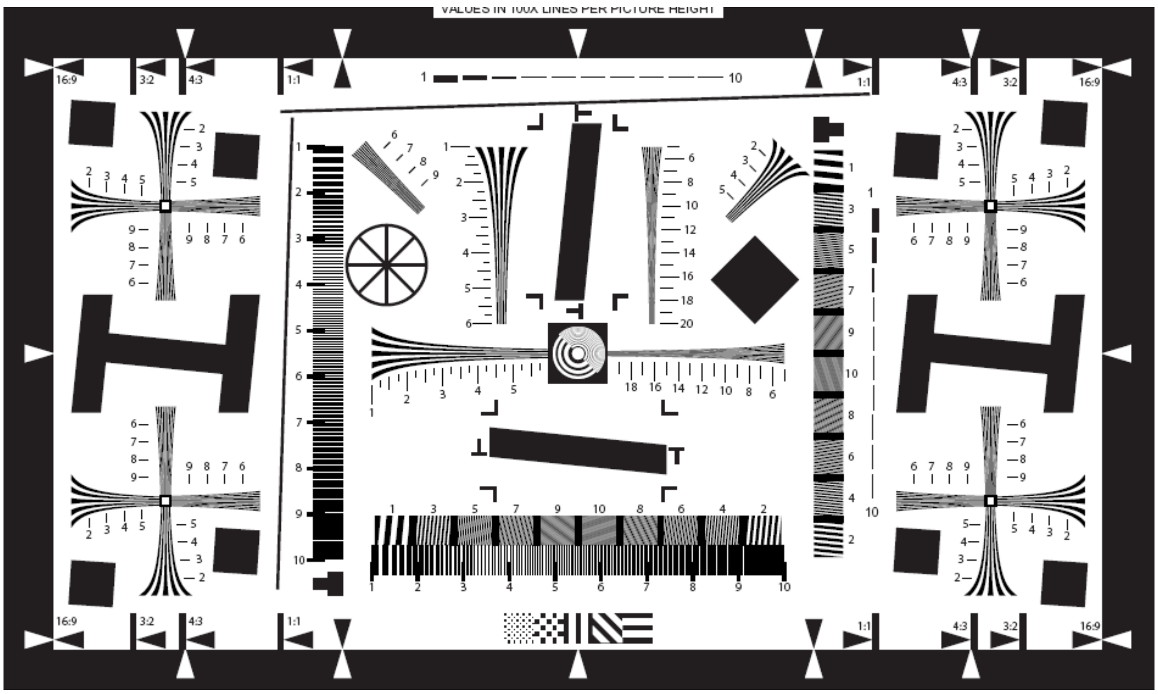

Figure 7: ISO 12233 Standard Chart for measuring resolution of “electronic still imaging cameras”

3.4. USAF 1951

The Air Force uses a chart for testing visual acuity of UAVs, night goggles and other imaging systems that consists of triplets of lines. The scaling is geometric; each group is related by powers of 2 and each group contains six patterns that are scaled by the sixth root of the power of two. The top numbers on the chart, labeling the group, are the exponents of 2 that give the frequency of the first pattern in the group. For example, -2 means 2**(-2) = ¼ cycle per mm and the 1-6 patterns in that group scale as the sixth root of 2. This means that the pattern 6 in the -2 group is the sixth root of 2 larger than the first pattern in the -1 group, and so on. This is a clumsy notation; a table giving the lines/mm for each element can be found on the Edmund Scientific web site.[18]

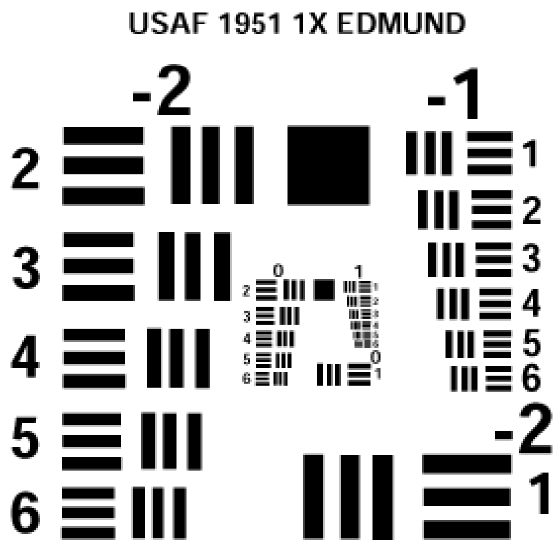

Figure 8: USAF 1951 resolution test chart

A similar test chart was developed by the National Bureau of Standards, NBS 1963A.

The results of testing with these charts can be related to Snellen fractions in the usual way: the ability to distinguish a set of bars separated by 1 arc minute is 20/20 vision, 2 arc minutes is 20/40 vision and so on. The triplets are very similar to the Tumbling E optotypes without the side bar and resolution results should be very similar.

3.5. Other Resolution Charts

There are many other charts, developed by EIA, ISO, IEEE, NBS and private companies, for testing scanners, microfilm readers, microscopes and other optical instruments. These are considered beyond the scope of this report.

4. Modulation Transfer Function

Lenses and optical systems are not perfect, so an image is degraded as the light passes through. The modulation transfer function is commonly used to measure that degradation.

4.1. MTF Defined

Consider a square wave grating. The light coming from (or reflecting from) the grating image can be quantified in terms of luminence. The maximum amount of light will come from the white spaces and the minimum from the dark bars. Modulation is defined as

$$\text{Modulation} = M = (L_{max} - L_{min}) / (L_{max} + L_{min})$$

This is basically a measure of contrast and, when defined as this in terms of light, it is called the Michelson contrast since he used it first in defining the visibility of interference fringes. The modulation function can be equivalently defined in terms of transmittances of the source (viewing a transparency against a uniformly illuminated background) or irradiances on the image.

This modulation function is dependent upon spatial frequency. Higher frequencies are degraded more than lower frequencies.

The Modulation Transfer Function (MTF) is defined as a function of spatial frequency as:

$$\text{MTF}(\nu) = M_i / M_o$$

Where M_i is the modulation of the image and M_o is the modulation of the object.

This ratio is less than one and at some spatial frequency goes to zero. That frequency for an ideal diffraction limited lens is the (incoherent) diffraction cutoff spatial frequency, ν_{ic}. The graph of MTF versus ν is called the modulation transfer function curve and is defined only for lens systems with positive focal length that form real images. The MTF graph is generally plotted in terms of the normalized spatial frequency

$$\chi = \nu / \nu_{ic}$$

The MTF graph for a lens is a line between (0,1) at zero frequency and (1,0) at the cutoff frequency. The graph for a real lens is often compared with that for an ideal diffraction limited lens.

MTF is very sensitive to lens imperfections and as such is a useful diagnostic in lens manufacture. An example on the Melles Griot website notes that a lens with a quarter wavelength aberration, which would be "barely discernible by eye, would reduce the MTF by as much as 0.2 at the midpoint of the spatial frequency range." [19]

MTF has the further value that MTFs of components of an optical system can be multiplied together to find the MTF of the total system. This allows isolation and consideration of each component separately.

4.2. MTF and Fourier Analysis

Fourier analysis allows complex patterns to be decomposed into sums of components at different frequencies. In particular, a step function at zero on the period (-π/2, π2) has a fundamental that is a sine wave with period 2π and a series of odd harmonics (sine waves with periods of 6π, 10π, etc.) that all together sum to a step at zero. If some of those higher harmonics are removed, the result is a curve that approaches a step but has some oscillation and is rounded from a value of -1 before zero, has a positive slope through zero and then rounds up to oscillate near +1 at some point after zero.

By analyzing the slope and curvature and oscillation of the response to a step function input, the MTF and cutoff frequency of the optical system can be measured. This is the point of the large angled black areas and the angled "H" patterns in the ISO 12233 chart. The reason these patterns are angled is to avoid errors introduced by digital filters in modern television circuitry.

Software tools to analyze images and compute MTF using the ISO chart are available commercially, for example from the International Imaging Industry Association, I3A[19]

Melles Griot notes that MTF should be measured using sine wave gratings, not square wave gratings, since as much as 20% of the light energy is in the third and higher harmonics. [19]

It should be noted that digital systems create additional measurement problems due to the quantized nature of digital sensors. Consider, for example, an image of a grid pattern taken with a digital camera when the grid image falling on the sensor exactly matches the pitch of the sensor pixels. If perfectly aligned, one pixel will see all white and its two neighbors will see all black. Moving the camera slightly, one half pixel in translation, will result in each pixel seeing half black and half white and thus producing a uniform grey. This is another reason ISO 12233 has wedges and patterns at angles.

5. Conclusion

Determining lines of resolution or spatial frequency cut-off requires busy charts that would be difficult to produce for different distances. In particular, the EIA and ISO and USAF resolution charts are comprehensive, but they are complicated. It is desirable to have a simple and easy to use pattern for use for USAR robot visual acuity testing.

MTF has the advantage that a system can be measured analytically and automatically, with no dependence on a human observer. The disadvantage for the current USAR robot exercise is that this is cumbersome and expensive in terms of purchasing the analysis tools and the engineering manpower needed to learn how to apply those tools, and producing the ISO chart or equivalent for long distance vision would be difficult.

Standard eye charts, on the other hand, are relatively inexpensive and everyone knows how to use them. Of the standard eye charts, the Tumbling E charts are considered the best choice.

The Tumbling E eye charts are recommended for a visual acuity testing procedure for ASTM E54.08.01.

6. References

[1] http://en.wikipedia.org/wiki/Visual_acuity

[2] Snellen, H., Letterproeven tot Bepaling der Gezictsscherpte, Utrecht, Weyers, 1862 (Dutch edition). Also published in many languages as Optotypi ad Visum Determinandum.

[3] The Society for the Preservation of the English System of Weights and Measures, founded in London in the 18th Century to combat the pernicious metric system promoted by the French, fought long and hard against metric measures. Their anthem starts out: "A pint's a pound the World around, as everybody knows…"

[5] Bailey, I.L., and Lovie, J.E., "New Design Principles for Visual Acuity Letter Charts", Am J Optom & Physiol Opt **53**: 740, 1976.

[4] Green, J., "On a New Series of Test-Letters for Determining the Acuiteness of Vision." Trans. Am. Opth. Soc. 4th Meeting (1868). P68.

[6] Ferris, F.L., Kassov, A., Bresnick, G.H., and Bailey, I., "New Visual Acuity Charts for Clinical Research." Am J Ophthalmol **94**:91, 1982.

[7] http://en.wikipedia.org/wiki/Landolt_C

[8] International Council of Ophthalmology, "Visual Acuity Measurement Standard (1984)." Italian J Opht, II/I, p1, 1988.

[9] http://www.lea-test.fi/

[10] http://www.preventblindness.org/children/distance_child.html

[11] Hered, R.W., Murphy, S., and Clancy, M., "Comparison of the HOTV and Lea Symbols charts for preschool vision screening". J Pediatr Opht Strab **34**(1):24, 1997.

[12] Taylor, H.R., "Applying New Design Principles to the Construction of an Illiterate E Chart." Am J Optom & Physiol Optics **55**:348, 1978.

[13] Multiple papers, e.g., Alexander, K.R., Xie, W., and Derlacki, D.J., "Spatial-frequency Characteristics of Letter Identification." J Opt Soc Am **11**, 2373, 1994.

[14] Reich, L.N., and Ekabutr, M., "The Effects of Optical Defocus on the Legibility of the Tumbling-E and Landolt-C. Optom Vis Sci **79**(6), 389, 2002.

[15] http://en.wikipedia.org/wiki/Image_resolution

[16] http://www.accu-chart.com/

[17] http://www.iso.org See catalog for ISO 12233:2000. Developed by TC42.

[18] http://www.edmundoptics.com/onlinecatalog

[19] http://www.mellesgriot.com/products/optics/os_2_2.htm

[20] http://www.i3a.org/downloads_iso_tools.html

Appendix: Some Applications of Visual Acuity Measures

Visual acuity measures are rather sterile when examined in isolation, and the subtle distinctions between Landolt C and Tumbling E optotypes seem to be fine nitpicking. There are, however, significant applications where visual acuity measures are very important.

Consider television and computer monitors. Standard broadcast television provides about 300 lines of resolution (NTSC provides 480 scan lines in the image; the ratio of lines of resolution to raster lines is called the Kell factor). For a 30" screen, fairly large for older vidicon tube sets, the vertical height is about 18", so each distinguishable feature is about 0.06" in height. When viewed at 10', that is an angle of ½ mrad or about 2 minutes of arc. If this were the finest resolution that could be seen, it would be slightly worse than "standard" human vision, equivalent to about 20/40 in fact, so a person with good vision will notice that the picture resolution is not very good. For larger sets, the resolution degrades proportionately and HDTV becomes increasingly attractive.

For a computer monitor, similar numbers hold. A 17" monitor with SVGA display provides pixels that subtend ½ mrad, about 2 minutes of arc, at about 20" viewing distance. The phosphor element pitch of 0.28 mm nearly matches this resolution.

NIST conducted a test of resolution for ladar scanners for the Army Research Laboratory for unmanned ground vehicle navigation. These tests are relevant to estimating necessary visual acuity for AUVs performing aerial surveys for USAR applications.

Figure A1: Two mannequins and a road sign as a test target group.

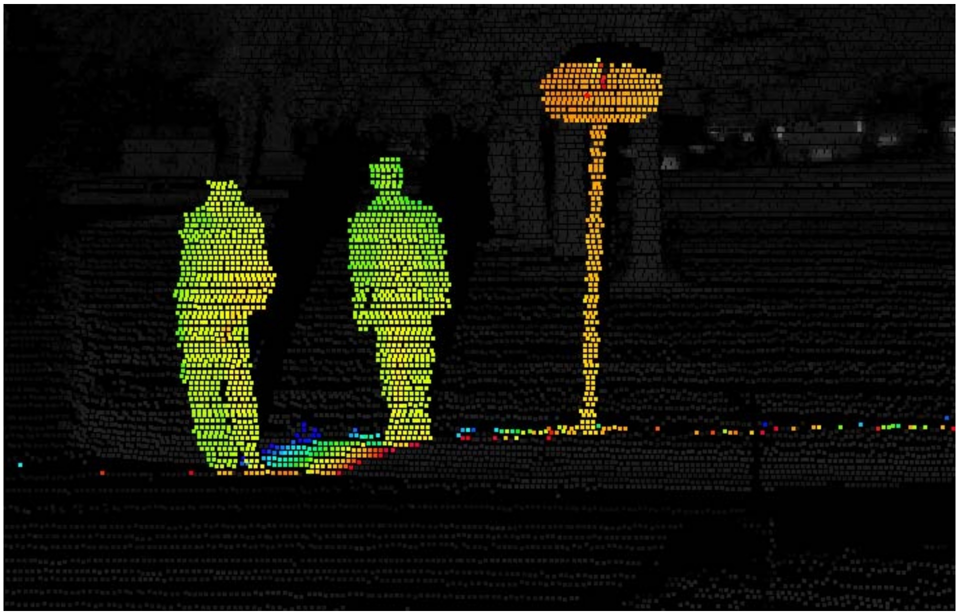

Figure A2: Targets at 100 meters with resolution of 0.02 degrees (20/24 visual acuity) Approximately 600 pixels on one human target.

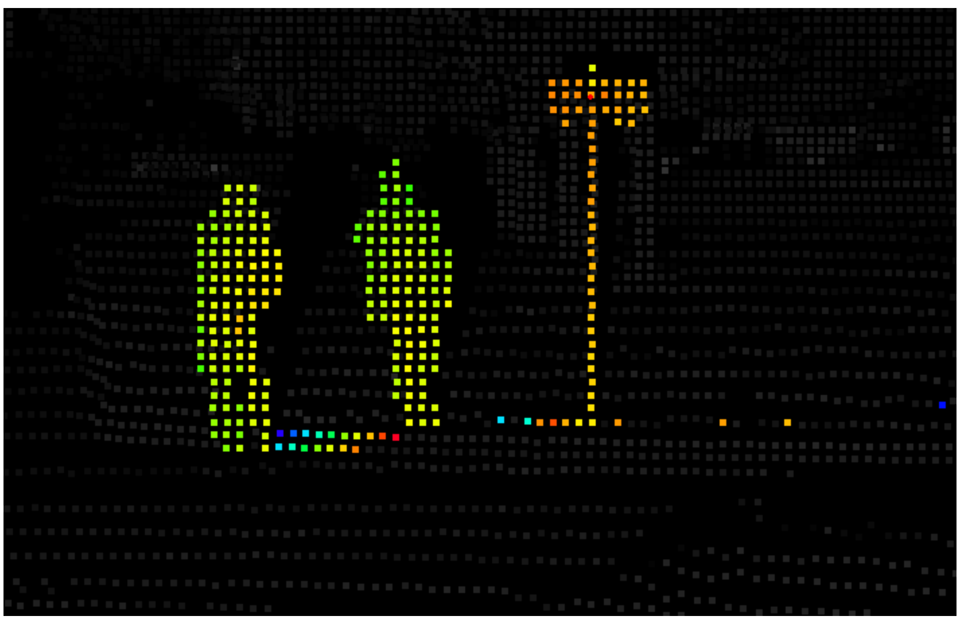

Figure A3: Targets at 100 m with resolution of 0.05 degrees (20/60 visual acuity). Approximately 100 pixels on one human target.

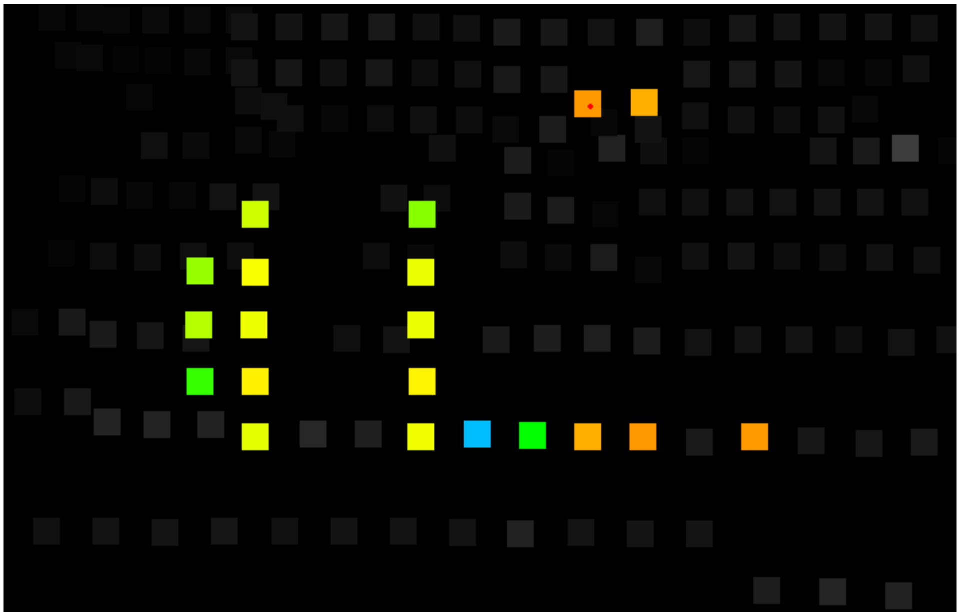

Figure A4: Targets at 100 m with resolution of 0.2 degrees (20/240 visual acuity). Approximately 6 pixels on one human target.

Figure A2-A4 show the target group at a distance of 100 meters (a typical altitude for a UAV) with pixel resolutions of .02, .05 and .2 degrees (1.2, 3, and 12 arc minutes). These resolutions correspond to visual acuity measures of 20/24, 20/60 and 20/240. Clearly 20/240 is useless, and 20/60 is marginal, but 20/24 is quite good. These are ladar range images, not visual images, so color and visual details are not seen, but the goal here was to determine the number of pixels on target.

Army target recognition experts say that you need a minimum of 600 pixels for target identification, so 20/24 vision for a UAV at an altitude of 100 m should be good and beyond 20/60 would probably be unacceptable.

Note that field of view becomes an important parameter in choosing cameras for distance vision. A sensor chip will have a certain number of pixels. Using a 60 degree wide angle field of view will mean that each pixel subtends an angle that is twice that for a 30 degree field of view, so visual acuity for a 60 degree lens will be half that for the 30 degree lens.

20/24 vision is angular resolution of 1.2 minutes of arc. For a camera with a lens that provides a 30 degree field of view, that is 1/1500 th of the field of view (30 degrees x 60 minutes/degree / 1.2). So one would need 1500 pixels across the image to resolve 1.2 minutes of arc. Assuming a 4/3 ratio of horizontal to vertical pixels, this is a 1500 x 2000 pixel camera which is 3 megapixels. Further, the communication system must transmit all 3 megapixels at frame rates to maintain the possibility of creating a display with the same resolution.

Providing bandwidth to transmit 3 megapixel images is quite difficult. However, high resolution can also be achieved with a zoom lens, reducing the effective field of view to examine details of interest. A 10/1 zoom would reduce a 30 degree field of view to a 3 degree field of view, and the number of pixels to provide 1.2 minute of arc would be reduced by a factor of 100, to 30,000 pixels. This is now less than 256 x 256 resolution, the communication support for which is in widespread use.

The recommended approach for wide area survey missions for ground robots is to use a 10/1 zoom (optical or electronic) to provide limited high visual acuity at a distance for items of interest, such as hazmat labels or humans, while still providing a large field of view for situational awareness.