

# Resolution requirements and the Johnson criteria revisited

Jon C. Leachtenauer\*

J/M Leachtenauer Associates Inc. 1281 Still Meadow Ave  
Charlottesville VA 22901

## Abstract

Since the 1950s, numerous studies have been performed within the surveillance and reconnaissance (S&R) and target acquisition (TA) communities in an attempt to predict information extraction performance as a function of image collection and quality parameters. In general, the work followed two separate paths. The TA community developed models to predict probabilities of detection, recognition, and identification as a function of target size, range, and collection system design/performance parameters (e.g., MRT, FLIR92, NVTHERM, MRC). The S&R community developed models to predict National Imagery Interpretability Ratings (NIIRS) as a function of system design and collection parameters (e.g. IR GIQE). More recently, efforts have linked the two approaches such that NIIRS can be predicted from TA models and probabilities of identification can be predicted from NIIRS.

With both approaches, resolution is a dominant term. A considerable amount of variability and uncertainty results from target differences. The criteria used to define the NIIRS generalize target type, size, and level of identification specificity. The TA predictions use the Johnson recognition criteria to relate lines on the target to recognition performance.

A recent paper found that TA predictions differed substantially between the visible and IR. Further, the paper reported substantial differences among vehicles in terms of a confusion matrix. This finding was not surprising in light of other research, but suggested the need for a more detailed examination and explanation of results. Accordingly, the current effort was undertaken. Data from a variety of past studies dealing with target recognition were examined relative to the Johnson criteria, along with a more detailed analysis of data from two recent TA studies. A hypothesis of target recognition performance was generated and partially validated using available data.

Key words: Johnson criteria, resolution, recognition performance

## 1. INTRODUCTION

The Johnson criteria provide the basis for current target acquisition models. Johnson presented various military targets to observers through electro-optical viewing devices.<sup>1</sup> Target range was increased until the target could barely be identified (e.g., M-48 tank), recognized as to the type of target (e.g., tank, APC, truck), or detected. A bar pattern was placed in the same field of view and spatial frequency increased until it could just be resolved at the same range as the target. In this manner, the number of resolution cycles required to achieve some level of performance for each target was defined. In subsequent Night Vision and Electronic Sensors Directorate (NVESD) studies, prediction models were refined.<sup>2-4</sup> Corrections were made for the length of the target and two dimensional criteria developed using the geometric mean of the target height and width (as seen by the sensor or observer). Table 1.1 shows the Johnson (cycles across minimum dimension) and current NVESD criteria (cycles across geometric mean)..

A key element of the Johnson criteria is the requirement for adequate contrast and signal-to-noise ratio. In early work by others, this requirement was sometimes overlooked.<sup>5</sup> The criteria are defined at the 50% performance value, an equation is used to define performance at other levels. It was originally defined as:

$$P(N) = \frac{\left(\frac{N}{N_{50}}\right)^{2.7+0.7\left(\frac{N}{N_{50}}\right)}}{1 + \left(\frac{N}{N_{50}}\right)^{2.7+0.7\left(\frac{N}{N_{50}}\right)}} \quad (1.1)$$

\* jcleachr@aol.com; phone 1 434 973-9582, fax 1 434 973-9582

where  $N_{50}$  is the required number of cycles for a 50% level of performance and  $N$  is the actual number of cycles.<sup>2</sup> The exponent is also defined as 1.73<sup>6</sup> and 3.73.<sup>7</sup> With the exponent in Eq. 1.1, the number of cycles required to achieve 90% performance is 1.75 times the number for 50% performance. With the 1.73 exponent, the ratio is 3.56, with the 3.73 exponent it is 1.8.

Table 1.1  
NVESD -Modified Johnson Criteria

Task	Description	Cycles across minimum dimension	Cycles across geometric mean
Detection	Target is military vehicle	1.0±0.25	0.75
Recognition	Vehicle type (tank, APC, truck)	4.0±0.8	3.0
Identification	Vehicle model	6.4±1.5	6.0

The General Image Quality Equation uses sensor resolution and other physical quality factors to predict NIIRS values.<sup>8,9</sup> NIIRS values define the tasks that can be performed on a given image in terms of object type, size, and level of identification.<sup>10</sup> Recent studies have related the NVESD and GIQE predictions for the IR.<sup>11-13</sup> A relationship has also been shown between the NIIRS and the Johnson criteria.<sup>14</sup>

A recent study using both visible and IR imagery showed substantial differences between the two image types relative to the Johnson/NVESD criteria.<sup>5</sup> The same study also showed substantial variability as a function of target type. This prompted a review of the Johnson criteria, both in terms of this recent study, as well as several previous studies relating resolution and recognition.

## 2. BACKGROUND

Johnson's criteria were initially published in 1958 [1]. Many similar studies were performed after Johnson's work, most investigators apparently unaware of the Johnson data. Whereas Johnson was concerned with direct viewing through optical systems, most of the later studies dealt with television and electro-optical line scan systems. Johnson, as well as several other investigators, was concerned with target acquisition from the ground. Others were concerned with observation from aerial platforms. A review of these other studies is of interest because of both similarities and differences relative to the Johnson criteria.

Although Johnson's study was one of the first relating resolution and recognition, efforts along these lines continued well into the seventies. A wide spectrum of target types was studied, ranging from simple shapes and Landolt Cs to industrial targets. Viewing aspect ranged from elevation views to plan views and included both low and high oblique views. It was recognized that visual subtense was an important factor and thus target (or resolution line) subtense was varied in many of these studies. Target subtense is implicitly treated in the NVESD/Johnson approach by the requirement to resolve bar patterns. Results of these previous studies are briefly reviewed in the following sections.

### 2.1 Symbol Identification

Baker and Nicholson studied Landolt C and alpha-numeric symbol recognition.<sup>15</sup> For the Landolt Cs, they found a performance threshold (point at which performance no longer improved) at 5 lines and 7 minutes of arc per C. Performance decreased at 4 lines and 5 minutes of arc. For letters and numbers, the threshold was 16 lines and 15 minutes of arc. In a third study, they used a set of silhouettes with objects ranging from a table fork to a heavy jet aircraft. Symbols varied in height/width ratio and were presented at two orientations (major axis vertical and horizontal). The performance threshold was reached at about 7 lines through the minor axis and 15 lines through the major axis.

Erickson and Hemingway studied geometric symbol recognition at three levels of lines per symbol and visual angle per symbol.<sup>16</sup> At the smaller visual angles, all of the lines could probably not be resolved and performance did not reach an asymptote. Performance on individual symbols varied from 50% to 97% correct with performance seemingly varying as a function of symbol similarity to other symbols. For example, the circle and hexagon were the most frequently confused symbols. The differences which distinguished the symbols were much smaller than the symbols themselves, on the order of 20% of the symbol height or less.

In an unpublished study, this author studied Snellen letter recognition (10 letters) at three visual angles over a range of lines/symbol.<sup>17</sup> Snellen letters have a 1:1 aspect ratio and stroke width is 20% of overall width. Substantial

differences occurred between different letters with performance ranging from 9 errors (“L”) to 106 errors (“N”). For letter pairs, errors ranged from 0 (F vs. H, L, and T) to 60 (F vs. P). Error frequency was related to differences between letters. The differences between letters ranged from 0.5 to 3 stroke widths. If the difference could not be seen, recognition could not occur. The data indicated a requirement of one raster line per stroke (5 lines per letter) and 2 minutes subtense per line or stroke.

## 2.2 Vehicle Identification

Rosell and Willson studied the effects of subtending scan lines on identification using an approach similar to Johnson’s.<sup>18</sup> Using oblique views of a sample of five tanks, they found that 13 TV lines (~6.5 cycles) were needed for identification. This value is virtually identical to the Johnson value of 6.4 cycles. However, a 42% difference in the number of required lines as a function of tank model was reported. It was attributed to differences in model similarity.

Erickson and Hemingway studied vehicle identification using target silhouettes.<sup>16</sup> The number of raster lines per vehicle (3.7, 7, 10.8) and vehicle subtense (4.4, 6.0, and 10.2 minutes of arc) were varied. A confusion matrix indicated that performance as a function of vehicle type ranged from 61% to 99%. A performance threshold appeared at 7 lines per vehicle (3.5 cycles) for the 4.4 and 6.0 minute subtense, maximum performance was achieved at the 10.8 lines/10.2 minute subtense condition.

In a follow-on study, they used oblique photos of vehicle models on two different backgrounds (sand and foliage).<sup>16</sup> Vehicle subtense (minimum dimension) was 6, 10, and 14 minutes; the number of subtending lines was 6, 10, and 15. A performance asymptote was observed at 10 lines (5 cycles) for all angular subtenses, performance was highest at the 14 minute subtense. Performance was better for the foliage background than the sand. Performance as a function of vehicle ranged from 30% to 93%. Note that Erickson and Hemingway reported TV lines whereas Johnson and Rosell and Willson reported the number of scan lines required to achieve 50% performance. Further, Johnson’s criteria were in terms of cycles; the other studies reported in terms of scan lines. Although it is convenient to one cycle as two scan lines, this is only a crude approximation.

Wagenaar and van Meeteran studied vehicle identification on line scan imagery.<sup>19</sup> They found a requirement for 5-10 lines per minimum vehicle dimension, but concluded that target type differences were important and were related to characteristics of the target other than overall size.

Scott, Hollanda,, and Harabedian studied vehicle identification using vertical and oblique views of 25 different vehicles.<sup>20</sup> The vehicles were oriented at 10 to 30 degrees from the scan lines so as to maintain a constant number of subtending scan lines across the minimum dimension. The number of subtending scan lines ranged from 4 to 30; scan line subtense was 9 minutes. It was concluded that about 20 scan lines were required to achieve 80-90% performance. In a follow-on study, Hollanda, Scott and Harabedian studied the effects of scan lines and signal-to-noise ratio.<sup>21</sup> A total of 20 vertical vehicle views were used. Ten of the vehicles were tanks and ten were “miscellaneous” vehicles (trucks, engineering equipment). Results showed a requirement for 30 lines for the tanks and at least 45 lines for the miscellaneous vehicles.

The Scott/Hollanda vehicles and those used in the other referenced studies differ significantly. The Scott and Hollanda vehicles tended to be plan views, whereas the other studies used side profiles or low oblique views. The plan views required use of internal detail; the vehicles used in the other studies could be identified largely on the basis of outline shape. Contrast and possibly size of the identifying details thus differed.

## 2.3 Aircraft Identification.

Lacey conducted a study in which televised aircraft model photos were presented to 15 observers.<sup>22</sup> The observers were asked to identify the photos as one of six different fighter aircraft. The 18 photos consisted of head-on, side view, and oblique views of the six aircraft against uniform backgrounds. Observers were shown the original photos and asked to identify the televised images as one of the 18 aircraft/orientation combinations.

The camera zoom lens was set to display the aircraft images at 7.2, 10.1, and 14.4 scan lines (active TV lines) per aircraft height. Observers were seated at a defined distance from the display (using a head rest) such that the vertical dimension of the targets subtended 6, 10.2, and 14.4 minutes of arc.

Results indicate that performance was improving at the best viewing condition, implying the need for additional scan lines or target subtense. Of equal interest was the effect of target type and orientation. Performance (across all conditions) ranged from 65% to 85% correct as a function of aircraft model and from 59% to 87% as a function of orientation. Performance was lowest on the oblique orientation and highest on the side. For the head-on and side views, height was measured from the bottom of the fuselage to the top of the tail. For the oblique view, it was

measured from the lowest point on the wing or horizontal stabilizer tip of the nose. The oblique views were at a smaller scale than the side and end views.

At each orientation, certain aircraft showed far more errors than did others. It is apparent that some aircraft were frequently confused for others. In the side view, for example, the F-14 and MiG-21 were frequently confused for each other and accounted for 27% of the total side view error. The MiG-21 and MiG-23 confusion accounted for 26 % of total errors. In the head on view, the A-4 and F-4 were most frequently confused (44% of total errors). In the side view, they accounted for only 8 % of the total errors.

Some insight can be gained by inspecting schematic drawings of the aircraft pairs. Figure 2.1 shows recognition errors for front views. Lines connecting aircraft pairs show the percentage of recognition errors. The A-4 and F-4 have two engine intakes above the wing; the F-14 and MiG-21 have rectangular intakes on both sides of the fuselage below the wings. This is reflected in the front view error rates.

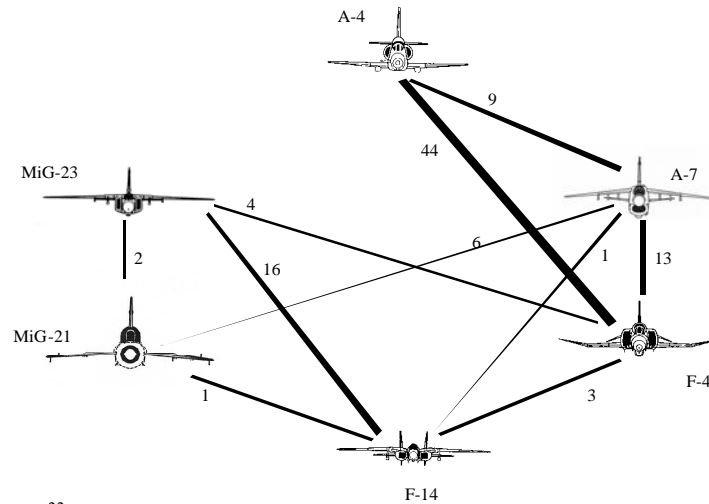


Figure 2.1 Front view error rates.<sup>22</sup>

In two related studies (Jones, Leachtenauer and Pyle<sup>23</sup>, Leachtenauer and Jones<sup>24</sup>), observers were asked to determine the equivalent ground resolution (one cycle or two lines) at which aircraft identification features could be identified. Equivalent ground resolution was defined in terms of visual angle subtense (cycles/degree), assuming a 53 cycle/degree visual resolution capability. Equivalent ground resolution was varied by changing viewing distance. Plan view silhouettes were used in the first study, aerial photos in the second. The features required to uniquely identify 16 fighter and attack aircraft were defined and the ground resolved distance (GRD) for those features specified. GRD is the width of a cycle (bar and space) on the ground. Some aircraft could be identified at 16 foot GRD, 2 foot GRD was required to identify all but two of the aircraft.

The aircraft ranged in length from 37 to 95 feet and in wing span from 27 to 70 feet. The geometric mean of wingspan and length ranged from 34 to 74 feet. Using the geometric mean and one resolution cycle as two lines, the number of resolution lines needed for identification ranged from 3 to 44. There was thus no simple relationship between identification performance and number of subtending lines.

In a later study, aircraft silhouettes were displayed on an 8 bit CRT at 0.5, 1, and 2 minutes per raster line.<sup>17</sup> Observers (three) were required to identify which of 10 aircraft were presented using a hardcopy key for comparison. The silhouettes were presented one at a time so relative size could not be used as a cue. An ascending and descending method of limits was used where the number of subtending raster lines was both increased and decreased until 100% correct performance was achieved.

Results showed that the number of subtending lines did not accurately predict performance. A confusion matrix showed significant target differences as well as response biases. Some targets were more often misidentified than others and where target "A" might frequently be called target "B". the reverse was not always true.

## 2.4 Other Target Types

Brainard, Hanford, and Marshall used models of buildings, bridges, storage tanks, and aircraft imaged by a TV system.<sup>25</sup> Observers were required to identify targets (using a photo as a briefing aid) on a simulated flyover. The range

at which correct identification occurred was recorded and translated to visual subtense and number of subtending TV lines. Target subtense for 50% identification performance ranged from 28 to 35 minutes. The required number of scan lines ranged from 5.8 to 7.

Leachtenauer and Boucek studied detail analysis tasks on aerial photos at three levels of ground resolution and three levels of magnification.<sup>26</sup> Observers were required to respond to questions regarding details in the scene such as aircraft wing shape, number of engines, location of roof vents, etc. The sizes of the cues needed to answer the questions were determined. It was concluded that the cue needed to be subtended by one to two cycles and a cycle needed to subtend one to two minutes of arc.

## 2.5 Summary

Taken together, the studies indicate a rather wide range in the number of resolution lines required for identification. Table 2.1 summarizes results. A wide variation in the required number of subtending lines is evident. Within some of the studies summarized, the range is even greater (3-44 for Reference 24). The variation appears to be greater when top or plan views of the target are used. It is also apparent that a resolution line must subtend 1-2 minutes of arc. The exception is the Brainard et al data (Reference 25), target motion may have increased the required value.

Table 2.1  
Results Summary

Target	View	Dimension	Req. Lines <sup>1</sup>	Subtense	Sub./Line %	Correct <sup>2</sup>	Lines@50% <sup>3</sup>	Reference
Letters & numbers		Height	16	15	0.9	95	7.6	15
Symbols		Height	10	20	2	95	4.8	15
Symbols		Height	7	10	1.4	98	2.9	15
Snellen letters		Height	5	10	2	95	2.4	16
Military equip.	Side	Min. dimen.	12.8	N/A	N/A	50	12.8	17
Vehicles	Side	Min. dimen.	13	N/A	N/A	50	13	1
Vehicles	Side	Min. dimen.	7	10	1.4	95	3.3	11
Vehicles	Oblique	Min. dimen.	10	14	1.4	75	7.4	16
Vehicles	Top	Min. dimen.	20	N/A	N/A	90	11.4	16
Misc. vehicles	Top	Min. dimen.	32	N/A	N/A	98	13.3	21
Tanks	Top	Min. dimen.	48	N/A	N/A	92	26	21
Aircraft	F/S/Obl.	Min. dimen.	14.4	15	1	84	9.3	22
Aircraft	Top	Geom. Mean	18	20	1.1	90	10.3	24
Aircraft	Top	Wing span	20	40	2	100	7.33	17
Various Tgts.	Oblique	Min. dimen.	5.8 to 7	28 to 35	5	50	5.8 to 7	25
Features/cues	Top	Min. dimen	3 to 4	3 to 4	1	100	1 to 1.43	26

<sup>1</sup> Lines are raster lines, TV lines, or 1/2 resolution cycle. Divide lines by 2 to obtain cycles.

<sup>2</sup> A value of 99% correct was used to compute the 50% threshold.

<sup>3</sup> The 50% threshold was computed using Eq. 1.1.

Johnson,<sup>1</sup> Rossel and Willson,<sup>18</sup> and Brainard et al<sup>25</sup> defined requirements at the 50% level of performance. The remainder of the referenced studies attempted to find a performance asymptote. Not all of these studies reached the asymptote and few showed performance at the 50% correct level used by Johnson and Rossel and Willson. A conversion was defined by Eq. 1.1 and is shown (lines @ 50%).

In several of the studies, it appeared that variations in performance were related to the features that distinguished a given target from the remainder. Target confusion matrices were decidedly non-uniform. Although it appears obvious, targets which looked alike were more often confused with each other than those that did not. Logic would suggest that it is the ability to see distinguishing features that defines identification performance, and that this ability may not be uniformly related to overall target size.

Studies using symbols and letters appear to show that the feature distinguishing one letter or symbol from another must be subtended by at least one raster line (and this raster line must be of a size sufficient to be resolved). Where sampling is an issue, the requirement may increase to two lines or one cycle. It thus appears logical to extend this reasoning to more complex objects. Any two objects will have a set of distinguishing features at a given orientation. The features may relate to external shape, or to the presence of internal (to the object outline) features or detail. In order for identification to occur, the features must be resolved by both the imaging system and the observer.

# Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

## Real-Time Litigation Alerts



Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

## Advanced Docket Research



With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

## Analytics At Your Fingertips



Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

## API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

## LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

## FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

## E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.