ARMY RESEARCH LABORATORY

A Generic Validation Methodology for Multispectral Synthetic Scene Generator Models—Interim Report

by Marcos C. Sola, Mark W. Orletsky,
Quochien B. Vuong, and Charles R. Kohler

ARL-TR-1446 October 1997

Approved for public release; distribution unlimited.

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

Citation of manufacturer's or trade names does not constitute an official endorsement or approval of the use thereof.

Destroy this report when it is no longer needed. Do not return it to the originator.

Army Research Laboratory
Adelphi, MD 20783-1197

ARL-TR-1446 **October 1997**

A Generic Validation Methodology for Multispectral Synthetic Scene Generator Models—Interim Report

Marcos C. Sola, Mark W. Orletsky,
Quochien B. Vuong, and Charles R. Kohler
Sensors and Electron Devices Directorate

Approved for public release; distribution unlimited.

Abstract

The establishment of a sufficient, field-measured database to support the analysis of automatic target recognition (ATR) algorithms, sensor fusion effectiveness, and sensor system performance for multiple combinations of targets, environments, sensors, and locations will severely challenge the limited, available resources currently within the U.S. Army research community. However, the use of a high-resolution, synthetic scene generator model (SSGM) for time-independent applications can alleviate the database requirement. We propose a methodology for a robust validation of SSGM that will consist of defining sets of images (real and corresponding SSGM imageries) and using human observers to define a baseline. First-order comparisons of a real scene to a synthetic scene will be performed with the use of the filters in the Tank-Automotive Research, Development and Engineering Center (TARDEC) [1] model or a comparable computational vision model (CVM). The similarity of target-to-background histograms as a function of various CVM filters will need to be analyzed to define first-order effects. Second-order metrics are defined in terms of probability of detection, detection timeline, and false alarm rate. A metric for the target signature will be mathematically defined to test these second-order effects. For a given application, the necessary and sufficient metrics are discussed.

[1] *Gary Witus and Thomas Meitzler, TARDEC Visual Perception Model, Calibration and Validation, Seventh Annual Ground Target Modeling and Validation Conference, Warren, MI (20–22 August 1996).*

Contents

1. Introduction **1**

2. Approach **2**

3. Present Methodology **5**

3.1 A Region-Based Method of Comparing Images *7*

3.2 The Symmetric Difference Method of Comparing Images *11*

4. Goals **12**

5. Acknowledgments **13**

6. References **13**

Distribution **17**

Report Documentation Page **19**

Figures

1. Examples of different images that have the same gray-level histograms 3
2. A tank and a helicopter with their identical gray-level histogram 3
3. Diagram of an approach to image validation 4
4. M60-A1 tank 6
5. CAMAELEON histograms 6
6. Image similarity 7

1. Introduction

Many Army programs require a significant number of signature databases in order to satisfy future program development. A partial listing of these programs includes Intelligence Signature Assessment, Battlefield Visualization Test Bed (BVTB), Automatic Target Recognition (ATR) Algorithm Development, Sensor Performance Analyses, Multi-Sensor Fusion, Countermeasure / Counter-Countermeasure (CM / CCM), Target Acquisition (TA) Modeling Improvement, Required Operational Capability Requirements, Tri-Service Smart Missile / Munitions Testing, and Computer War Gaming Input.

Any of these programs would require signatures related to target acquisition. Target acquisition is a function of many variables; among them are the target, background, environment, geographical location, time, and sensor. Each of these sets is composed of subsets. The number of field measurements required or needed is given by

$$F_{\text{Measured}} = \prod^{n} i\, T_A(v_i) \ , \tag{1}$$

where

F_{Measured} = field test measurement,
T_A = target acquisition function,
v_i = v_1 = target = T (type, aspect angle, engine history, static, dynamic, …),
v_2 = background = B (type, homogeneity, clutter, …),
v_3 = environment = E (real, smoke / obscurant, battlefield, …),
v_4 = season = S (summer, fall, winter, spring),
v_5 = location = L (Europe, Mid-East, …),
v_6 = diurnal cycle = D (time),
v_7 = sensor = s (type, spectral band, field-of-view mode, …),
.
.
.
v_n.

To physically measure signatures in the field to satisfy a wide dynamic range of signature conditions would be a substantial budgetary challenge for any project manager [2]. In addition, the timeliness of having the data and the danger associated with the acquisition of a particular set of data (foreign targets in hostile environment (location and weather)) would have to be factored into the data requirements. From equation (1), it is evident that there would not be enough manpower or dollars to completely quantify the signature of one target over all the possible combinations of sets and subsets for target acquisition. The physically measured data must not only answer the requirements of the project, but also must be general and of sufficient resolution to take into account other near-term signature requirements [3]. Hence, there is the need to make use of synthetic databases to augment, supplement, and / or complement field databases. However,

this requires a robust validation [4] of the particular synthetic scene generator model (SSGM) [5] that is used so that it gains credibility and acceptance in the modeling and simulation community.

2. Approach

An integral step in the validation of an SSGM is comparing a set of image pairs under the criteria that are relevant to the given application. We define an image pair as the physically measured scene (target in background / clutter) and the corresponding synthetically generated scene based on model input requirements. For example, if we had a real-synthetic pair of images of a simple block world, with well-defined lighting, angles, etc, we would almost definitely do well by comparing the images on a pixel-to-pixel basis. While not every pixel of the synthetic image would have the same value as the corresponding pixel of the real image, a sufficient number would be close enough to indicate the quality of the synthesis. In this case, the simplicity of the "world" under consideration would almost definitely permit synthesis of images that are, pixel-wise, very similar to real images.

For most applications of synthetic images, however, a bit-wise comparison to a corresponding real image is impractical as a validation criterion. Take, for example, the task of comparing images of a natural scene containing trees, grass, sky, water, etc. Here the variations that could be expected between the images would be large. For example, the wind may sway the objects in different directions, the clouds may cast shadows in different places, and so forth. Thus, the use of a pixel-to-pixel comparison would almost definitely be too exacting to effectively evaluate the quality of the synthesis.

An alternative method of comparison that has seen substantial use involves the comparison of statistics [6–11] that have been derived over the entire image. Such statistics include gray-level histograms, local-energy histograms, and many others. These statistics frequently give information about the quality of a synthetic image, but rely on obtaining the statistics from the image as a whole. Thus, they can be misleading, as the following examples show.

An intensity histogram (also called gray-level histogram) is a necessary but not sufficient tool for comparative assessment of an image pair. Figure 1 shows three different image pairs that would produce the same intensity histogram for both images in each of the given pairs, while figure 2 shows two different weapon platforms with exactly the same intensity histogram. Figure 2 was produced by a C-program that repositions the pixels of one image to approximate the other. Thus, virtually any image can be slightly modified to have a gray-level histogram of another image while still retaining its original "look." Therefore, given this example, sole reliance on histogram distribution as a similarity metric for image comparison can lead to a wrong conclusion.

Figure 1. Examples of different images that have the same gray-level histograms.

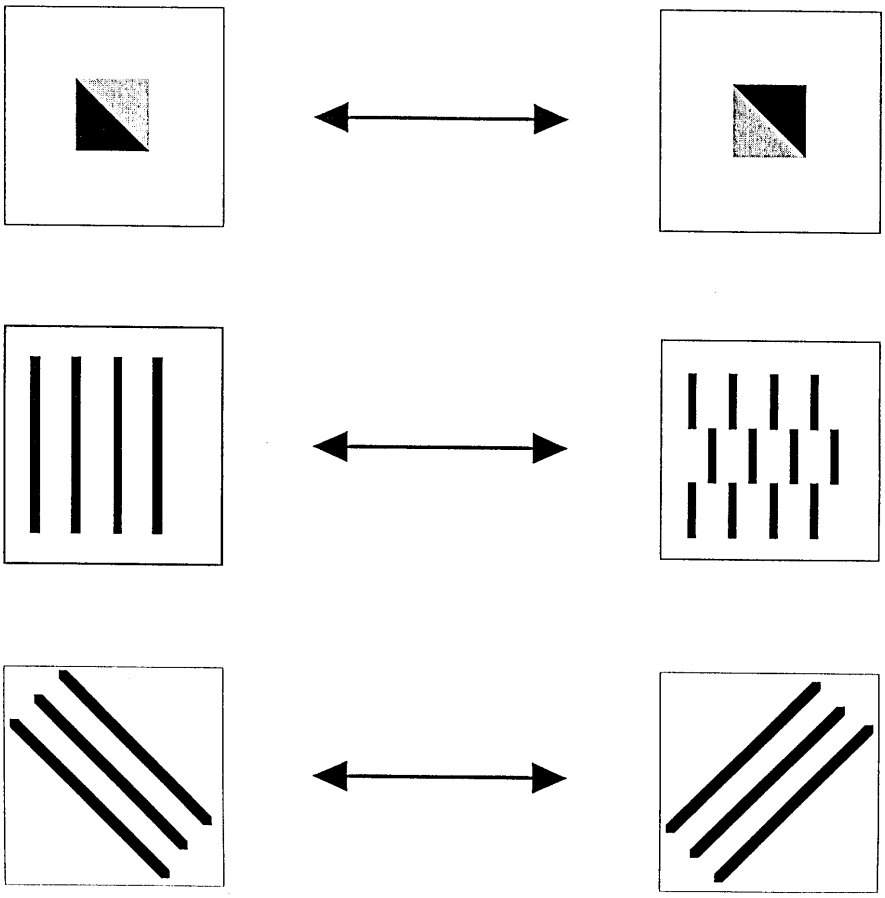
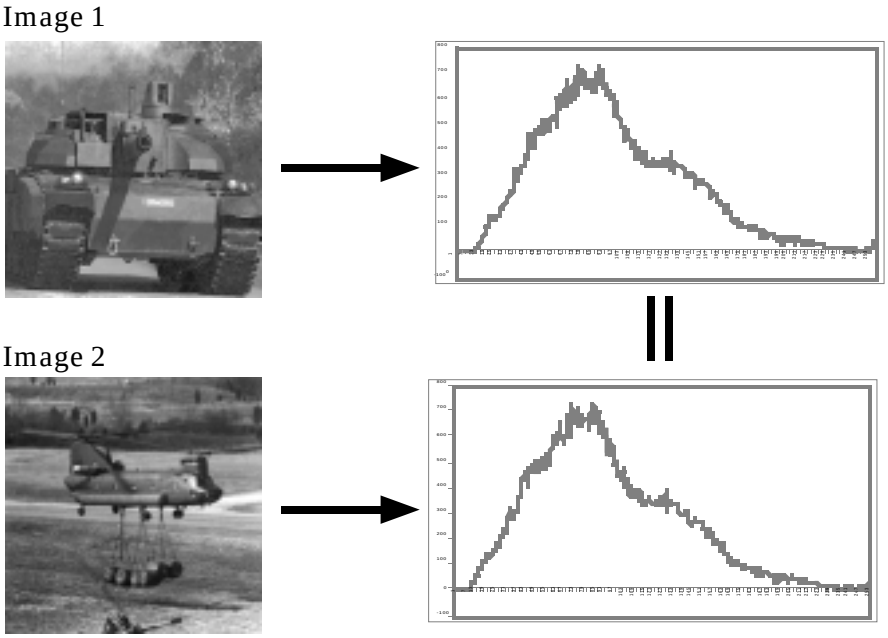


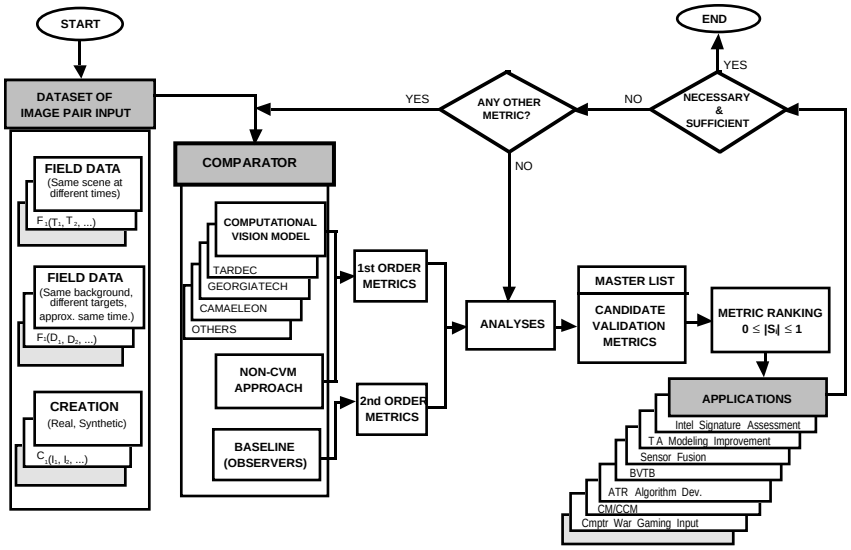
Figure 2. A tank and a helicopter with their identical gray-level histogram.



The approach taken by the U.S. Army Research Laboratory (ARL) is to develop a robust SSGM validation methodology for any synthetic rendering model. In particular, we would like to test this methodology against our own model for generating synthetic images—CREATION [12–17]. Figure 3 is a generic approach to the problem of image validation. An image pair is passed through a comparator, and its output is statistically analyzed to identify a validation metric.

The comparator has three components: a computational vision model (CVM), a noncomputational vision model (non-CVM) approach, and a group of human observers used as a baseline. First-order validation metrics are obtained from a CVM such as the Georgia Institute of Technology Vision Model [6,7], Tank-Automotive Research Development and Engineering Center (TARDEC) Vision Model [1], German CAMAELEON Model [8,9,11], or other suitable candidates. These metrics define the fundamental attributes of the scene and are ***not*** the final output of the model. Second-order metrics are obtained by using trained military observers to provide probability of detection, highest level of acquisition (classification, recognition, identification) given detection, detection timeline, and false alarm rate. The aforementioned models also provide some of these second-order metrics as model output. In figure 3, note that military observers are used as a baseline. We purposely elected not to use an ATR algorithm as a baseline because the threshold for correct acquisition can possibly change from algorithm to algorithm. Instead, we chose to use the ATR as one of the comparators belonging to the non-CVM class. The non-CVM approach can provide either first-order or second-order metrics. If an image has the same second-order metrics as another, this does not necessarily mean that their validation metrics will be the same. Consider as an example two images that provide all the same second-order metrics as previously mentioned. Such a condition can be satisfied, for example, by a low-observable tank at close range versus a high-contrast tank at long range (all other variables being equal). A robust set of validation metrics should provide an indication of two different scene conditions.

Figure 3. Diagram of an approach to image validation.



From the comparator, a master list of candidate validation metrics is compiled. A sufficient set of image pairs is parsed through the particular component of the comparator that is being tested for statistical analysis. These metrics are then rank-ordered from zero to one in terms of how similar the synthetic scene is compared to the real scene. A similarity of one is a perfect fit, while a similarity of zero signifies no correlation between the two images being compared. The number of validation metrics required is a function of the SSGM application. For high-resolution applications, such as target acquisition modeling improvement, most if not all the validation metrics with high similarity values may be required. For low-resolution applications, such as real-time computer war gaming, only a certain portion of the validation metrics may be needed, and their similarity metric requirement will be less stringent. It is therefore necessary to rank-order the validation metrics from the master list for a given synthetic scene-rendering application. In the CREATION model, we lack enough field-measured data from which we can generate a synthetic scene. Part of the problem is that the CREATION model requires a diurnal cycle target signature history for us to be able to create a synthetic scene. Although ARL needs to validate the CREATION model, it must first be able to develop a robust validation methodology. To alleviate the image pair database problem, some pairs could be created from real-field data, for example, scenes at different times or for the same time and background, but different targets.

3. Present Methodology

Our current validation approach is to postulate a similarity metric that is defined as

$$0 \le |S_i| \le 1 \ , \tag{2}$$

where

$$S_i = \begin{cases} 0, \text{ no match} \\ 1, \text{ perfect match} \end{cases}$$

and i = statistical validation metric (0, 1, 2, …n) in the master validation metric list.

A wide dynamic range of signature image pairs is created from field-measured data and their corresponding synthetic scene or from field data with differences in either time, background, target, environment, or other measurable variables. Each image set is parsed through a comparator with, in this case, the German CVM CAMAELEON. This model outputs the first-order metrics such as gray-level, frequency, orientation, and local energy distributions. For each of these metrics, a statistical analysis is applied in terms of mean, median, mode, variance, standard deviation, absolute deviation, skew, kurtosis, and entropy. Figure 4(a) shows the field-measured data of an M60-A1 tank scene taken with a DL calibrated infrared sensor at FT AP Hill, VA. A data artifact was introduced unintention-

ally because every other field was missed during the digitization process, resulting in a lower quality image than what a high-resolution, calibrated DL forward-looking infrared (FLIR) instrument [18] is capable of showing. Figure 4(b) is the synthetic infrared rendering by the CREATION model. It contains some statistical sampling rather than first-principle rendering of background data. Figure 5 shows the comparison between the real and the synthetic scene based on the output of the CAMAELEON model. Figure 6 shows our use of the similarity metric with our present validation approach. Two non-CVM approaches developed by ARL for comparing images, the region-based and the symmetric difference methods, are discussed next.

Figure 4. M60-A1 tank.

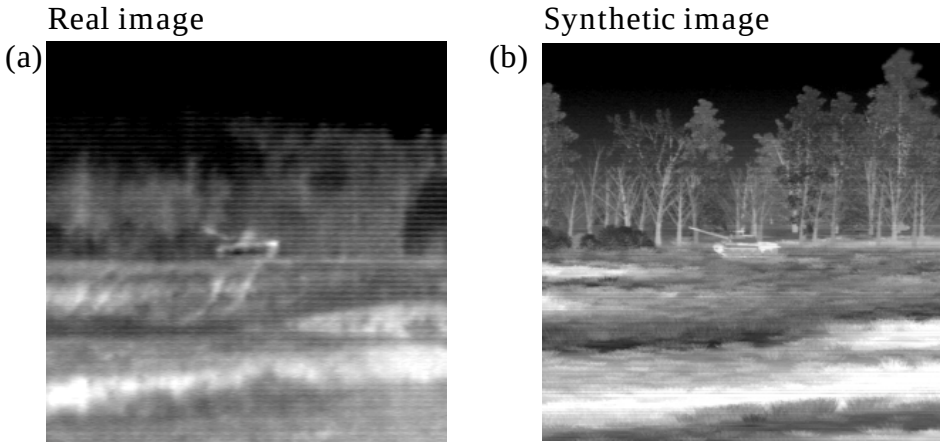


Figure 5. CAMAELEON histograms.

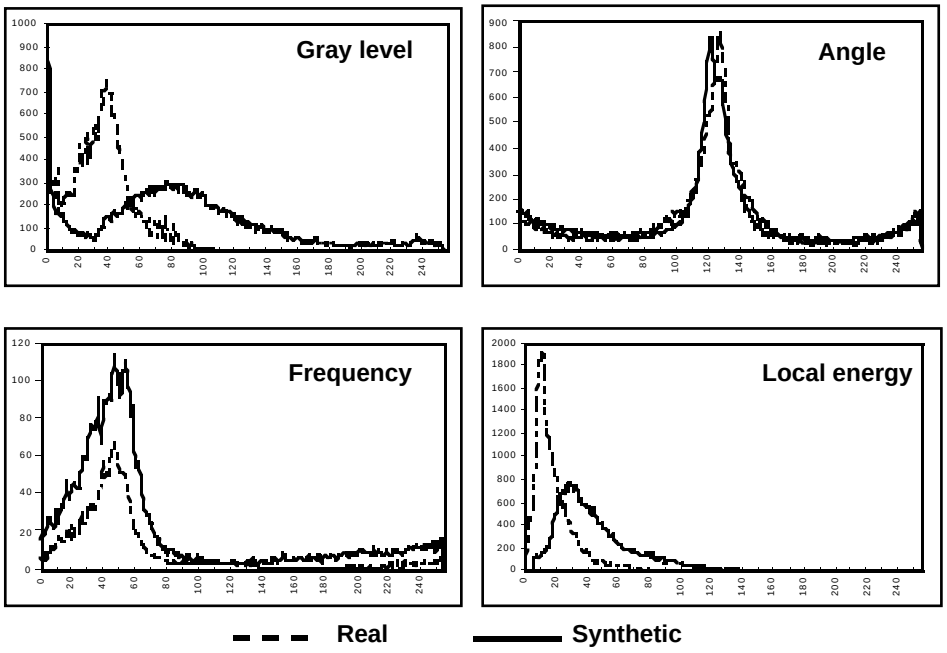
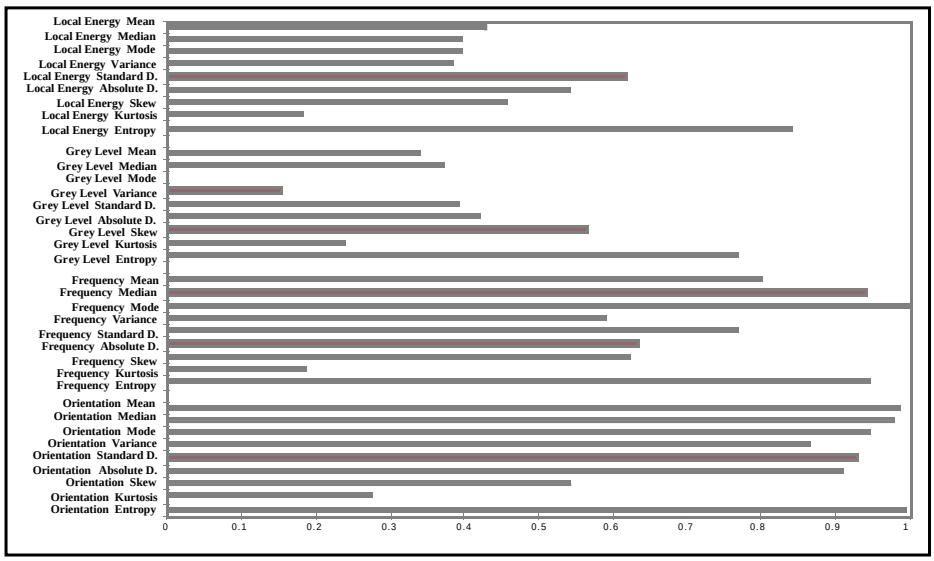


Figure 6. Image similarity.



3.1 A Region-Based Method of Comparing Images

Historically, there have been two different approaches for comparing a real scene to the synthetically rendered scene. One approach relies on very local information such as pixels, and the other relies on global information such as statistics generated over the entire image. Our approach is to use the middle ground between these approaches. We call it a region-based method of comparing images.

In region-based image comparison, we produce a mask that defines nominally 40 to 100 regions in the real or the synthetic image under comparison. This mask is then applied to both the real image and the synthetic image, and image statistics (gray-level histograms, local energy, etc) are computed over each region. Comparisons are then made of the statistics obtained from each region to the statistics obtained from the same region in the other image. A real-synthetic image comparison is then obtained by computing the area-weighted average of the similarity metrics that resulted from each of these local comparisons.

For example, note that the images in figure 2 have identical global gray-level histograms. Thus any comparisons based on global gray-level histograms will indicate that the two images are identical. Suppose, alternatively, that we generated any arbitrary tessellation and applied that tessellation to both images. We could then generate the gray-level histograms over each of the regions produced, and compare the histogram of a particular region to the histogram in the corresponding region in the other image. While the gray-level statistics over the images as a whole would be identical, the same statistics over a small portion of each of the images would most likely not be.

Thus, regionalization has the potential to reduce the negative effects of false matches that can occur when comparisons based on global image statistics are used. Additionally, the use of regionalization cannot produce

results that make images appear to be statistically more similar than those that would be obtained from a global comparison. Specifically, if the images compare favorably on a region-to-region basis, they will also compare favorably on a global basis. In short, regionalization can be used to allow a more rigorous application of common global statistics, and achieves a balanced medium between the very demanding bit-wise image comparisons and the somewhat ineffectual global comparisons.

As to the question of which regionalization to use, we think that any regionalization or tessellation would be acceptable because an arbitrary regionalization has the potential for producing more accurate image comparisons than the corresponding global comparison, and no regionalization has the potential for producing comparisons that are less appropriate than global comparisons. We, however, also see an advantage in producing regionalizations that are consistent with the low-frequency variations that are naturally present in the image.

Next we give a step-by-step description of a region-based approach of comparing two images under a low-frequency mask. Step 1 aligns the two images with each other, step 2 matches the average brightness of one image to that of the other image, step 3 creates a low-frequency mask, and step 4 uses the mask to apply a global image comparison metric.

Step 1: Register and crop the images.

The first step in comparing images is to make sure they are (1) registered (or aligned) with each other and (2) of identical size. We accomplish this as follows. A human operator locates a well-defined point in each image and determines the pixel coordinates of that point. At present, a tool such as *xv* (x-windows view) is used to determine the coordinates of that point. The real and synthetic images have not been previously registered, so the coordinates of the “well-defined point” will in general be different for the two images. Our program “crop” is then used to produce a new image for each of the two original images. The new images will be of a specific size and centered at the chosen point. We used a size of 256 by 256 pixels for the initial trials. At this point we have a real-synthetic pair of images that are the same size and registered with each other.

This registration process involves translation only. A more sophisticated method of registration would also include rotation and scaling. However, this would necessitate a method of mapping the rectilinear grid of pixels of the original image to another grid of a different angular orientation and having a different scale. Methods for performing such a mapping exist, though they generally introduce artifacts, and the artifacts would have a great potential to pose problems for the comparison. Thus, we insist that the real and synthetic images be made to the same scale and oriented at the same angle, and then the registration is done by performing only a translation.

Step 2: Gamma-match the images.

A process that we call "gamma-matching" is used to change the brightness levels of the synthetic image so that the total brightness of the synthetic image is within a small factor of the total brightness of the real image. This is done with gamma-correction as follows. If the synthetic image is dimmer than the real image, gamma values are successively chosen starting at 1 and increasing in steps of 1 until the brightness of the corrected synthetic image exceeds the brightness of the real image. If the synthetic image is brighter than the real image, gamma values are successively chosen starting at 1 and decreasing in steps of 0.1 until the brightness of the corrected synthetic image is less than the brightness of the real image. When such cross-over points are found, the process is repeated with smaller step sizes, specifically step sizes of one-tenth the size currently being used. This entire process is repeated until the ratio of the total brightness of the synthetic image to the total brightness of the real image is within a small constant of 1. We have been using the constant 0.001.

Step 3: Construct a template for one of the two images.

The process of constructing a low-frequency template for an image involves three steps: low-pass filtering the image, multilevel thresholding the image, and uniquely labeling the regions that result. These three steps are explained next.

Step 3a: Low-pass filter the image.

The image is low-pass filtered by convolving it with a square template containing a normalized Gaussian function. Here, the user specifies the size of the "radius" of the template to be used and then the template is automatically generated. For example, if the template radius is chosen to be 5, then an 11 by 11 template will be generated. This template is then filled with a two-dimensional Gaussian function centered at the center pixel in the template and having a sigma chosen so that the area under the Gaussian curve over the template is 99 percent of the total area under the Gaussian curve over the entire real plane. The entries in the template are then normalized over the template (i.e., scaled so that they sum to 1). When the image has been convolved with such a template, the brightness values of the image will be "smooth" in the sense that there will be no rapid changes in brightness. The image will actually appear blurred. This is done so that the next step of thresholding will not produce many small regions, which would be possible if the image contained high-frequency components.

Step 3b: Apply multilevel thresholds to the image.

After step 3a, the image will have brightness values that do not change rapidly throughout the image. In fact, the image can be thought of as a landscape in which the brightness values represent altitudes. After the low-pass filtering, all the hills and valleys will be smooth and gently rounded. There will be no sharp peaks, no spikes, no cliffs, no abrupt changes of any sort.

We now apply a multilevel thresholding process to the low-pass filtered image. Here, we (arbitrarily) choose four levels of brightness uniformly spaced between the dimmest pixel value and the brightest pixel value. Thus, every single pixel in the image falls into exactly one range. Each pixel is then assigned a number from 1 to 5 corresponding to the range into which it falls. After this thresholding process, the image appears somewhat like a topographic map in that the regions shown correspond to the brightness range of the pixels in the region. This looks very similar to topographic maps that show the altitude ranges between the curves on the map.

Step 3c: Label the regions produced.

The next step is to give the regions that result from the above step a unique label and to give all pixels the label of the region into which they fall. To accomplish this, we initially set all pixels to be unlabeled and then start in the upper left corner of the image and proceed in a raster-like scan. During the scan, we do the following. When an unlabeled pixel is encountered, we use a recursive process that we call flooding to label each pixel in the same region as the newly encountered pixel with the "next available label." When the single raster scan is complete, all pixels in the image will be labeled, either from the scan itself or from the recursive flooding process.

Here we scan the image from left to right and from top to bottom in a raster pattern. If a pixel is encountered that is unlabeled, it is given the "next available label" and the recursive flooding procedure is invoked. This procedure will attempt to recurse one pixel left, one pixel right, one pixel up, and one pixel down from the recently labeled pixel. Specifically it will recurse in those directions if and only if the pixel in that direction has not been previously labeled, the pixel in that direction exists (i.e., it is not at the edge of the image), and the pixel in that direction has the same brightness value (after the thresholding) as the recently labeled pixel. Thus the recursion will proceed to label all pixels in the given region and it will not proceed across region boundaries. At the completion of recursively labeling a region, the raster scan will continue until another unlabeled pixel is encountered, at which point the region containing it will be flooded in a similar manner. This continues until the entire image has been scanned and labeled.

Step 4: Apply a global-metric according to the mask produced.

Upon completion of the mask, or tessellation, that was produced by step 3, we can now use that mask to apply a similarity metric in a region-based fashion. As an example, let us consider using the "common area overlap" of the gray-level histogram as our metric. Recall that if we were using this metric in the global sense, we would generate the gray-level histogram for the real image and also for the synthetic image, normalize each of those histograms, and then determine the area under the histograms that is common to both histograms. This area will be zero if the histograms are completely disjointed and it will be one if the histograms are identical. In the region-based method, we apply this metric not to the images as a whole, but to each of the individual regions of the tessellation. Thus, we will

obtain a number between zero and one for each of the regions, which indicates the extent to which the gray-level histograms of the two images resemble each other in that region. To obtain the final image metric, we simply take an area-weighted average of the similarity metrics that we obtained for the regions. The equation for the region-based metric is

$$v_j = \frac{\left|\phi\left(r_j\right) - \phi\left(s_j\right)\right|}{\max\left[\phi\left(r_j\right),\ \phi\left(s_j\right)\right]}\left(\frac{a_j}{A}\right) , \tag{3}$$

$$V_i = \sum_{\forall_j} v_{i,j} , \tag{4}$$

$$S_i = 1 - V_i , \tag{5}$$

and

$$0 \le S_i \le 1 , \tag{6}$$

where

i = first-order validation metric,
v_j = region-based difference,
ϕ = statistical operator,
r_j = region j of the real image,
s_j = region j of the synthetic image,
a_j = area of region j,
A = area of the whole image,
V_i = difference between real and synthetic over the entire image, and
S_i = similarity of real versus synthetic over the entire image.

Equation (4) is defined as in equation (5) in order to preserve the concept of the similarity metric we postulated earlier, i.e., the closer to the value of 1, the better the scene rendering (see eq (2)).

3.2 The Symmetric Difference Method of Comparing Images

The segmentation of images into regions as described above can also be used to produce a figure of merit of comparison between two images. Consider generating one mask based on the real image and another mask based on the synthetic image. Comparison of the masks, then, could be used as a method of comparing the images.

The manner in which we compare the masks is as follows. For one of the two masks derived from the real image, consider a particular region in the mask. Calculate the symmetric difference between this region and every region in the mask of the other (synthetic) image and choose the minimum. Notice that if this region closely coincides with a region in the other image, this minimum will be small. In fact, if there is a region in the other mask that is identical to the region under comparison, this minimum will be zero. Now choose a second region in the real image and repeat this process to obtain another “minimum symmetric difference.” If we continue this for

all regions in the image, we will develop a string of minimum symmetric differences. We use the sum of this sequence as a scene metric.

Notice that this sum will be small when each region in the real image closely coincides with a related region in the synthetic image, and it will be large otherwise. Thus, we have a metric that will be zero when comparing two identical images, will be small when comparing two images with similar low-frequency structures, and will increase for images that have few such commonalities. The non-CVM approach can be expressed by the following figure of merit:

$$\Psi_{SDM_j} = \min\left(r_j \,\Delta\, s_k\big|_{\forall k}\right) \quad * \quad \frac{\left(r_j \cup s_k\right)}{A} \quad , \tag{7}$$

where the same definition for the parts apply as in equation (3) and

Ψ_{SDMj} = area weighted minimum symmetric difference found and
N = total number of comparisons made between real region r_i and synthetic regions s_j.

Thus,

$$\Psi_i = \frac{\sum_{\forall j} \Psi_{SDM_j}}{N} \quad , \tag{8}$$

$$S_i = 1 - \Psi_i \ , \tag{9}$$

$$0 \le S_i \le 1 \ , \tag{10}$$

and

$$r \,\Delta\, s = (r \sim s) \cup (s \sim r) = (r \cup s) \sim (r \cap s) \ . \tag{11}$$

These two proposed figures of merit as part of the scene metrics will have the strength of distinguishing between the two images in figure 2 with identical global intensity distribution. A sufficient amount of co-registered data would need to be analyzed to enhance these scene metrics and properly define their limitations and usefulness in conjunction with the validation metrics and figures of merit previously defined. For image pairs where translation, scaling (magnification), and rotation are factors to be considered, more extensive analyses and resources would be required. Fortunately, for the intended application of this validation methodology (ARL CREATION model), this SSGM can negate the problems described. The threshold for creating the masks is anticipated to be a function of the application of the SSGM.

4. Goals

The near-term goals are to be able to generate a sufficient number of image pairs over a wide dynamic signature range based on high-quality field measurement data. Investigation of other candidate validation metrics [19] will need to be analyzed.

The long-term goal is to compare various target acquisition models in terms of their capability to predict second-order effects, again over a wide dynamic signature range. The prediction results from these various models can be compared to the baseline (human observers), and the differences can be analyzed using validation metrics identified to date from this work.

A vision to allow optimization of resources in this particular research area is to apply this methodology for dual technology application (military and nonmilitary). The methodology being developed for validation will lend itself well for the analysis to allow improvement of target acquisition modeling, understanding of the ATR technical underpinnings via the scene metrics under analysis [20], and broadening of our technical interaction with the research and development community. One pristine area for target acquisition enhancements for the military, as well as for law enforcement, is in the littoral environment [21–23]. A coalition force is being attempted between Department of Defense (DoD) agencies that could be extended to North Atlantic Treaty Organization (NATO) working groups that are interested in this particular area of research and development.

5. Acknowledgments

This effort was made possible by the members of the ARL Electro-Optics/Infrared Image Processing Branch: Khang Bui, Janice Colby, Glenn Dockery, Giap Huynh, Teresa Kipp (Branch Chief), Hung Nguyen, Joseph Penn, Matthew Thielke, and Paul Zirkle, who provided several of the real-synthetic image pairs presented in this report, along with many hours of consultation and technical support. Further technical support was provided by Jim McManamey and Eddie Jacobs, of the Night Vision Electro-Optics Laboratory's Camouflage Modeling Branch, Fort Belvoir, VA, who ran our images through the CAMAELEON software package. This includes the images presented in this report.

6. References

1. Gary Witus and Thomas Meitzler, *TARDEC Visual Perception Model, Calibration and Validation*, Seventh Annual Ground Target Modeling and Validation Conference, Warren, MI (20–22 August 1996).

2. Marcos C. Sola and Tadeusz M. Drzewiecki, *A Proposed Technical Methodology to Reduce the Number of Physical Infrared Signature Measurements*, Eleventh Annual Ground Vehicle Symposium, Keweenaw Research Center, Michigan Technical University, Houghton, MI (22 August 1989).

3. Andrew Zembower and Marcos C. Sola, *Signature Quality Metrics*, Eleventh Annual Ground Vehicle Symposium, Keweenaw Research Center, Michigan Technical University, Houghton, MI (22–24 August 1989).

4. Department of the Army, Army Regulation 5-11, Section 6-2: Verification and Validation, Department of the Army, Washington, DC (1992).

5. Marcos C. Sola, Mark W. Orletsky, Quochien B. Vuong, and Charles R. Kohler, *Proposal for Validation of a Multispectral Synthetic Scene Generator Model*, Proceedings of the 1996 Battlespace Atmospherics Conference, Naval Command, Control and Ocean Surveillance Center, San Diego, CA (3–5 December 1996).

6. T. J. Doll, S. W. McWhorter, and D. E. Schmieder, *Computational Model of Human Visual Search and Detection*, Proceedings of the March 1994 IRIS Passive Sensors Symposium, Albuquerque, NM (1994).

7. T. J. Doll, S. W. McWhorter, D. E. Schmieder, A. A. Wasilewski, and G. Welch, *Georgia Tech Vision (GTV) Model, Version GTV94a*, Analyst's Manual, Georgia Tech Research Institute, Atlanta, GA (1994).

8. R. Hecker, *CAMAELEON—Camouflage Assessment by Evaluation of Local Energy, Spatial Frequency, and Orientation*, Proceedings of the Third Annual Ground Target Modeling and Validation Conference, Anne Marie L. LaHaie, ed., U.S. Army Tank Automotive Command, Warren, MI (1993).

9. R. Hecker, "CAMAELEON—Camouflage Assessment by Evaluation of Local Energy, Spatial Frequency, and Orientation," *Characterization, Propagation, and Simulation of Sources and Backgrounds II*, Dieter Clement and Wendell R. Watkins, eds., Proceedings of SPIE 1967 (1992), 342–349.

10. G. H. Lindquist, G. Witus, T. H. Cook, J. R. Freeling, and Grant Gerhart, *Target Discrimination Using Computational Vision Human Perception Models*, Proceedings of SPIE 2224 (1994), 30–40.

11. James R. McManamey, *An Initial Investigation of the CAMAELEON Model*, Proceedings of the Sixth Annual Ground Target Modeling and Validation Conference, Houghton, MI (August 1995), 239–248.

12. Charles R. Kohler and Marcos C. Sola, *CREATION—Multispectral Synthetic Scene Generation*, Seventh Annual Ground Target Modeling and Validation Conference, Warren, MI (20–22 August 1996).

13. G. H. Kornfeld, *Various FLIR Sensor Effects Applied to Synthetic Thermal Imagery*, Proceedings of SPIE (April 1993).

14. Joseph A. Penn, Hung Nguyen, Teresa Kipp, Charles R. Kohler, Giap Hyunh, and Marcos C. Sola, *The CREATION Scene Modeling Package Applied to Multispectral Missile Seekers and Sensors*, Proceedings of the 1995 Conference on Multispectral Missile Seekers and Sensors, Redstone Arsenal, Huntsville, AL (1–2 November 1995).

15. Joseph A. Penn, H. Nguyen, M. Sola, C. Kohler, J. Weber, and S. Hawley, *The CREATION Scene Modeling Package Applied to Theater Air Defense Fire Control Situations*, Proceedings of the 1995 National Fire Control Symposium, Monterey, CA (31 July–3 August 1995).

16. J. Weber and J. A. Penn, *CREATION User's Manual* (draft), Version 1.27, U.S. Army Research Laboratory, Signature Modeling Branch, Fort Belvoir, VA (November 1994).

17. J. Weber and J. A. Penn, *CREATION and Rendering of Realistic Trees*, SIGGRAPH 95 Computer Graphics Conference Proceedings, Annual Conference Series 1995 (August 1995), pp 119–128.

18. William Stump, notes from *M60-A1 Tank Diurnal IR Signatures*, Night Vision and Electro-Optics Division report (1995).

19. James P. Welsh, *Smart Weapons Operability Enhancement (SWOE), Joint Test and Evaluation (JT&E) Program Final Report*, SWOE report 94-10 (August 1994).

20. George Singley and Jasper Lupo, *Technology Area Review and Assessment Briefing*, Naval Research Laboratory (18–21 March 1997).

21. Marcos C. Sola, *Active/Passive Signature Enhancer*, Patent Application, Army Research Laboratory Docket # 96-34 26 (October 1996).

22. Marcos C. Sola, *Application of Optical Augmentation Technique for Maritime Search and Rescue Operations*, National Target/Threat Signature Data Systems Conference, Naval Air Warfare Center, Point Mugu, CA (23–25 July 1996).

23. Marcos C. Sola, Joseph A. Penn, E. Glenn Dockery, Paul Zirkle, Charles R. Kohler, Teresa Kipp, and Janice F. Colby, *A Proposal for an Active/Passive Signature Enhancer (APSE) for IFF*, accepted for publication in 1997 Joint Service Combat Identification Systems Conference (CISC-97), Coronado Naval Amphibious Base, San Diego, CA (15–17 April 1997).

Distribution

Admnstr
Defns Techl Info Ctr
Attn DTIC-OCP
8725 John J Kingman Rd Ste 0944
FT Belvoir VA 22060-6218

Defns Intel Agency
Attn CMO-2 J Glover-Jones
Bolling AFB
Washington DC 20340

Natl Ground Intel Ctr
Attn IANG-RSG B Reinhold
220 Seventh Stret
Charlottesville VA 22902-5396

Ofc of the Dir Rsrch and Engrg
Attn R Menz
Pentagon Rm 3E1089
Washington DC 20301-3080

Ofc of the Secy of Defns
Attn DASD/C3I B Leong-Hong
The Pentagon Rm 3D228
Washington DC 20301-7100

Ofc of the Secy of Defns
Attn ODDRE (R&AT) G Singley
Attn ODDRE (R&AT) S Gontarek
The Pentagon
Washington DC 20301-3080

OSD
Attn OUSD(A&T)/ODDDR&E(R) J Lupo
Washington DC 20301-7100

CECOM
Attn PM GPS COL S Young
FT Monmouth NJ 07703

CECOM RDEC Elect System Div Dir
Attn J Niemela
FT Monmouth NJ 07703

CECOM
Sp & Terrestrial Commctn Div
Attn AMSEL-RD-ST-MC-M H Soicher
FT Monmouth NJ 07703-5203

Dpty Assist Secy for Rsrch & Techl
Attn SARD-TR R Chait Rm 3E476

Dpty Assist Secy for Rsrch & Techl (cont'd)
Attn SARD-TT D Chait
Attn SARD-TT F Milton Rm 3E479
Attn SARD-TT K Kominos
Attn SARD-TT R Reisman
Attn SARD-TT T Killion
The Pentagon
Washington DC 20301-0103

Hdqtrs Dept of the Army
Attn DAMO-FDT D Schmidt
400 Army Pentagon Rm 3C514
Washington DC 20301-0460

Night Vision & Elec Sensors Dir
Attn AMSEL-RD-NV-LWS D Jenkins
Attn AMSEL-RD-NV-OD J Ratches
Attn AMSEL-RD-NV-OV J Pollard
Attn AMSEL-RD-NV-VISPD L Obert
Attn AMSEL-RD-NV-VISPD M Lorenzo
Attn AMSEL-RD-NV-VISPD C Hoover/
C Walters
Attn AMSEL-RD-NV-VISPD J
MacManamey/
E Jacobs
10221 Burbeck Rd Ste 430
FT Belvoir VA 22060-5806

US Army Matl Cmnd
Dpty CG for RDE Hdqtrs
Attn AMCRD BG Beauchamp
5001 Eisenhower Ave
Alexandria VA 22333-0001

US Army Matl Cmnd
Prin Dpty for Acquisition Hdqrts
Attn AMCDCG-A D Adams
5001 Eisenhower Ave
Alexandria VA 22333-0001

US Army Matl Cmnd
Prin Dpty for Techlgy Hdqrts
Attn AMCDCG-T M Fisette
5001 Eisenhower Ave
Alexandria VA 22333-0001

US Army Spc and Strat Defns Cmnd
Attn CSSD-SD-T B Rogers
Huntsville AL 35807-3801

Distribution

US Army Tank-Automtv Cmnd
Rsrch Dev & Engrg Ctr
Attn AMSTA-TR-S G Gerhart
Attn T Meitzler
Warren MI 48397-5000

US Military Academy
Dept of Mathematical Sci
Attn MAJ D Engen
West Point NY 10996

USA STRICOM PMITTS, TMO
Attn AMCPM-ITSS-QG M Haack
Redstone Arsenal AL 35898-7458

USAASA
Attn MOAS-AI W Parron
9325 Gunston Rd Ste N319
FT Belvoir VA 22060-5582

USACRREL
Attn SWOE G Koenig
Attn SWOE P Welsh
72 Lyme Rd
Hanover NJ 03755

Nav Surface Warfare Ctr
Attn Code 723 P Ostroski
Carderock Div
Bethesda MD 20084-5000

Nav Surface Warfare Ctr
Attn Code B07 J Pennella
17320 Dahlgren Rd Bldg 1470 Rm 1101
Dahlgren VA 22448-5100

NAWCPNS
Attn Code 4KM400E P Yu
521 9th Stret
Point Mugu CA 93042-5001

Ofc of Nav Intel
Attn ONI-235 J D Camp
4251 Suitland Rd
Washington DC 20395-5720

GPS Joint Prog Ofc Dir
Attn COL J Clay
2435 Vela Way Ste 1613
Los Angeles AFB CA 90245-5500

Joint Munitions Test & Eval Program
Attn 46 OG/OGML M Heard
Eglin AFB FL 32542

Marine Corps Intelligence Activity
Attn MCIA 104 B Wong
3300 Russell Rd Ste 250
Quantico VA 22314

Special Assist to the Wing Cmndr
Attn 50SW/CCX Capt P H Bernstein
300 O'Malley Ave Ste 20
Falcon AFB CO 80912-3020

US Air Force
Attn 46 TW/OL-AG LT COL L Palmer
MZ2161 Box 731
FT Worth TX 76101-0371

US Coast Guard Hdqtrs
Attn G Hover
10892 Shennecossett Rd
Groton CT 06340-6096

DARPA
Attn B Kaspar
Attn L Stotts
3701 N Fairfax Dr
Arlington VA 22203-1714

DOT&E-1700
Attn A Manriquez
Pentagon Rm 3E333
Washington DC 20301-1700

US Army Rsrch Lab
Attn AMSRL-SL-EG C R Cundiff
Attn AMSRL-SL-ET W Watkins
Attn BED IS&TD R Shirkey
White Sands Missile Range NM 88002

US Army Rsrch Lab
Attn AMSRL-CI-LL Techl Lib (3 copies)
Attn AMSRL-CS-AL-TA Mail & Records Mgmt
Attn AMSRL-CS-AL-TP Techl Pub (3 copies)
Attn AMSRL-SE J M Miller
Attn AMSRL-SE-E J Pellegrino
Attn AMSRL-SE-SE M Sola (10 copies)
Attn AMSRL-SE-RS A Sindoris
Adelphi MD 20783-1197

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	October 1997	Final, October 1996 to September 1997

4. TITLE AND SUBTITLE

A Generic Validation Methodology for Multispectral Synthetic Scene Generator Models—Interim Report

5. FUNDING NUMBERS

PE: 62120A

6. AUTHOR(S)

Marcos C. Sola, Mark W. Orletsky, Quochien B. Vuong, and Charles R. Kohler

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

U.S. Army Research Laboratory
Attn: AMSRL-SE-SI
2800 Powder Mill Road
Adelphi, MD 20783-1197

8. PERFORMING ORGANIZATION REPORT NUMBER

ARL-TR-1446

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Research Laboratory
2800 Powder Mill Road
Adelphi, MD 20783-1197

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

AMS code: 622120.H16
ARL PR: 7NE4NN

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The establishment of a sufficient, field-measured database to support the analysis of automatic target recognition (ATR) algorithms, sensor fusion effectiveness, and sensor system performance for multiple combinations of targets, environments, sensors, and locations will severely challenge the limited, available resources currently within the U.S. Army research community. However, the use of a high-resolution, synthetic scene generator model (SSGM) for time-independent applications can alleviate the database requirement. We propose a methodology for a robust validation of SSGM that will consist of defining sets of images (real and corresponding SSGM imageries) and using human observers to define a baseline. First-order comparisons of a real scene to a synthetic scene will be performed with the use of the filters in the Tank-Automotive Research, Development and Engineering Center (TARDEC) [1] model or a comparable computational vision model (CVM). The similarity of target-to-background histograms as a function of various CVM filters will need to be analyzed to define first-order effects. Second-order metrics are defined in terms of probability of detection, detection timeline, and false alarm rate. A metric for the target signature will be mathematically defined to test these second-order effects. For a given application, the necessary and sufficient metrics are discussed.

[1] *Gary Witus and Thomas Meitzler, TARDEC Visual Perception Model, Calibration and Validation, Seventh Annual Ground Target Modeling and Validation Conference, Warren, MI, 20–22 August 1996.*

14. SUBJECT TERMS

Modeling validation, synthetic scene generation, image matrics

15. NUMBER OF PAGES

24

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

DEPARTMENT OF THE ARMY
U.S. Army Research Laboratory
2800 Powder Mill Road
Adelphi, MD 20783-1197

An Equal Opportunity Employer