# INTEROPERABILITY ISSUES ASSOCIATED WITH THE USE OF DISSIMILAR SIMULATORS

**Richard Purdy and Brian Goldiez**
**Institute for Simulation and Training (IST)**
**University of Central Florida**
**Orlando, FL**

## ABSTRACT

Interoperability issues associated with the use of dissimilar simulators within a DIS network range across a diverse spectrum of concerns. These include differing levels of simulation fidelity, terrain data rendition and correlation, and their relationship to purpose, activity, training efficacy and realism. The Institute for Simulation and Training (IST) is undertaking a two-pronged approach to the study of the interoperability problem. The first prong identifies the primary simulator parameters affecting interoperability and then quantifies them on existing simulators. This is coupled with polling the user community as to whether or not the simulators are perceived to be 'interoperable.' By doing so IST intends to identify the magnitude of parameter variation deemed acceptable by the user community. The second part of our approach is a research program that quantitatively determines the magnitude of parameter differences at which interaction outcome is affected. The initial investigation in both prongs is limited to consideration of the visual rendition of terrain on image generators within the simulators. In-house studies include tests determining the effect of various magnitudes of terrain mis-correlation on the outcome of combat engagements. The results of both investigations are ultimately combined to define an interoperability test procedure that is applied to networked simulators.

This paper addresses the rationale and approach for assessing interoperability and summarizes the results to date of the experiments and data gathering activities of our research program.

## ABOUT THE AUTHORS

Richard Purdy is a Research Associate at the Institute for Simulation and Training. Mr. Purdy has worked for 15 years as a systems engineer and systems analyst in the aerospace industry. He received his BA degree in Chemistry and his MS degree in Electrical Engineering from the University of Arizona. Mr. Purdy is a member of the National Council of Systems Engineers (NCOSE).

Brian Goldiez is Deputy Director and Manager of Research & Development, responsible for managing the overall technology and research at the Institute. Mr. Goldiez has over twenty years of experience in simulation and training working in government, commercial, and academic environments. He received his BS degree in Aerospace Engineering from the University of Kansas and MS in Engineering Math and Computer Systems from the University of Central Florida. Affiliations include the Society for Computer Simulation, Institute of Electrical and Electronics Engineers and American Institute of Aeronautics and Astronautics.

# INTEROPERABILITY ISSUES ASSOCIATED WITH THE USE OF DISSIMILAR SIMULATORS

**Richard Purdy and Brian Goldiez**
**Institute for Simulation and Training (IST)**
**University of Central Florida**
**Orlando, FL**

## 1.0 INTRODUCTION

### 1.1 Interoperability problem overview.

When simulators are exercised jointly in a DIS network there have been significant problems that are lumped under the heading of 'interoperability'[1,2,3,4]. These problems run the gamut from network communications problems to disparities in the rendition of the virtual environment. Over time, solutions to most of the problems will be agreed on and implemented. There is however, a class of problems which are particularly difficult; correlation of parameters between and among simulators. An example of this class is the correlation of the virtual environment and its rendition on dissimilar simulators. This problem is particularly vexing because it is so visible; tanks appearing in mid-air because a hill in one simulator's virtual environment does not occur in another and helicopters flying below ground for the same reason. Efforts have been made to mitigate [3] and compensate[5] for the phenomena described, but the problem persists because of different source data, algorithms and image generator pipelines between networked simulators. Other examples of this class of problems for a given entity type involve differences in hardware and software implementation of simulators. An example of this is two M1A1 tank simulators, one of which considers soil type in its mobility model and the other does not. Obviously the difference between the two simulators impacts any combat engagement in which the two participate. Correlation of the outcome of such an engagement to the real world is dubious. The task of achieving 'interoperability' for this class of problems involves determining the level of mis-correlation that can occur without degrading the utility of the simulators.

### 1.2 Interoperability problem definition

It is the contention of the authors that one of the most difficult aspects of the interoperability problems lies in the definition of the terms and the problem. Once a quantitative problem statement is achieved, solutions and/or trade-offs can be considered. The currently accepted definition of interoperability in the DIS community is that interoperability is achieved when two or more simulations/simulators, for a given exercise have performance characteristics to support a fair fight to the fidelity required for the exercise. A fair fight is defined to occur when differences in simulator performance are overwhelmed by user action. This is a reasonable high-level definition. The authors have broken this definition down into four sub-classes of interoperability problems to facilitate the development a quantitative definition. The classes were chosen because of their utility. From such a decomposition logically flows the requirements and activities necessary to achieve the defined paradigm.

Case 1: Interaction/virtual incompatibilities. This type of interoperability problem arises when a characteristic or behavior of one type of virtual world entity is not recognized or is incompatible with the same characteristic or behavior of another entity. Firing munitions at an entity that it does not recognize and can not assess probable damage if hit is an example of this problem. A method used by some simulators is to convert the munition into an entity that is known and assess the resulting damage (e.g., treat an AK-47 round as if it were an M-16 round). Such an approach reduces the magnitude of this class of incompatibility but increases the problem of non-uniform simulation fidelity incompatibilities (Case 3).

Case 2: Low interactive fidelity. Virtual world entities should interact with each other and with the environment according to a set of rules such as the laws of physics. Using such laws, vehicles can not drive through solid objects. When entity interactions defy the rules, training efficacy is degraded.

Case 3: Non-uniform levels of simulation fidelity. This interoperability problem occurs when two simulators model objects or behavior at differing levels of fidelity. Such differences skew interaction results and degrade training efficacy. For example, if one tank simulation considers in its mobility model the terrain type (sand versus clay) and adjusts its movement accordingly while another tank simulation does not consider terrain, the differing level of simulation fidelity may result in skewed interactive outcome. Another example of this is two simulators whose image generators differ in rendition capability (e.g., polygon loading, color gamut, & etc.)

Case 4: Differences in the virtual environment. An assumption made when connecting simulators is a single and common virtual environment. When the environment is not common, problems such as intervisibility, floating tanks and subterranean aircraft degrade the realism of the training scenario, skew the interaction outcome, and degrade training efficacy.

Note that all four cases are classes of differences between simulators. The differences are behavioral, capability related or data related.

It should also be noted that the definition is in accordance with the scope of the investigation (see paragraph 1.3); interoperability issues associated with image generators. Although the definition does not directly address issues that are specific to any given training task, they are applicable. For example to implement a specific training task which involves more than one simulator, a certain level of fidelity must be present in all simulators. This is an example of a Case 3 type issue but with only a single-sided limit--on the low side. The interactions between and among the simulators must also have a certain level of fidelity (Cases 1, 2 and 4). By breaking down the interoperability issues into these smaller pieces, it is easier to identify the fidelity parameters relevant to interoperability.
It is the authors position that interoperability will not necessarily be achieved even in the presence of homogeneous image generation systems. The algorithms governing the behavior and performance of entities with each other and with the environment would likewise have to be uniform. Such uniformity is extremely unlikely in the presence of heterogeneous hardware and software systems. Practical interoperability will be achieved when all the relevant parameters are identified and acceptable quantitative difference limits are defined and adhered to.

Because networked simulators are non-homogeneous as to entity, hardware, purpose, simulation model, & etc., instances of all four of the interoperability cases are expected to occur to one degree or another. The critical issue is the degree to which they can occur and still accrue utility from connectivity.

### 1.3 Investigation scope

The scope of this investigation does not (initially) include all the possible interoperability issues. It is limited to those issues associated with the image generator (IG) portion of the simulators. Consequently, Case 1 issues include only an examination of the commonality of model images. Case 2 issues are not dealt with at all. Case 3 issues to be considered are limited to rendition differences. The bulk of Case 4 issues impinge on the IG system; these issues are examined in significant detail. In considering all of the cases, we make no distinction between simulator types; CGF implementations are treated as a stand-alone simulator to the extent possible.

The examination of interoperability as defined is not limited to inter-simulator issues. Most simulators include abilities above and beyond instantiation of a single entity. Other common items such as computer generated forces (CGF), plan-view (stealth) displays, map displays require examination of interoperability, both within a given simulator and also with similar items in other networked simulators.

### 1.4 Investigation methodology.

IST has observed interoperability problems at various DIS demonstrations and set out identifying the parameters affecting the interoperability of simulators. A search of the available literature yielded only descriptive and anecdotal information of the same. Our intent is to determine the limits over which these parameters may vary while still accruing advantage from networking simulators. To achieve this goal, three steps need to be taken:

(1) Parameters impacting interoperability must be identified.

(2) A means of measuring these parameters has to be developed.

(3) A method of determining limits for these parameters has to be developed for defining ‘interoperability.’

A related activity being engaged in is the development of useful corrective action methodologies. This includes the methods for the discovery of problem sources, prioritization of the problems, and tools and methodologies for automated correction of the problem. For example, the IST discovered that virtual environment terrain mis-correlation occurred as a result of using terrain data from multiple sources, The IST has developed a set of tools that measure the correlation of two terrain databases so that significant discrepancies can be adjudicated prior to an exercise. This was done with significant success in the DIS demonstration performed at the December 1994 IITSEC[4]. The IST is currently developing tools and a methodology for determining the correlation of other items such as culture features and models.

Step 1 (parameter identification) was straightforward. A large number of the parameters affecting interoperability are obvious to the user community. Parameters have manifested themselves in various DIS exercises and were reported in the literature.[4,6]. The more significant parameters selected include the following:

Database and model correlation. Interaction between or among simulators must share a single definition of the environment; hills must be the same shape and in the same place. The occurrence of culture features such as trees, roads and rivers must also have a single common instantiation. This commonality must also extend to simulator entities; a given M1A1 tank should be the same size and color in all simulator image generators.

Viewport position and field of view. For a given viewport on a given entity type, differences in the position or field of view could skew an interaction. If one entity's view port provides a greater field of viewing, it provides an advantage over other similar entities.

Color gamut. Many visual cues for target detection and recognition are color related; a green camouflage tank is more detectable against a brown desert background than it would be in a forest. If two simulators differ in their ability to reproduce color, the visual cues may be degraded or lacking.

Entity detectability and identification. As range increases in the virtual environment, image generators normally reduce the complexity of the image in order to conserve image generating capability for those items more proximate. As detail is lost, the ability to detect and identify entities is also degraded. The ability to detect, identify and react to entities at uniform ranges is necessary if interactions are to be consistent.

Image generator rendition capability. The standard metric used here is the polygon throughput capability of the system. An image generator that has the ability to create a scene with 10,000 polygons and update it at a 60 hertz rate will be able to render terrain and model detail that a lesser system would not. If unequal image generators are used in opposing forces, the operator of one of the systems will have a significant advantage.

Step 2 (Parameter measurement) is to be accomplished by two methods: Acquisition of performance data from hardware and software specifications of the image generators and simulators, and by testing. Data acquisition via specifications is straightforward. Test data is acquired by means of a standard set of tests being developed by the IST. These test procedures, which are available via the anonymous FTP site at the IST, are used to acquire or verify data regarding the interoperability parameters defined above.

Because of the wide range of simulator types, the standard tests are constrained in their design to apply stimuli and acquire data at only two points: Across the DIS network or at the image generator display. No access to the internal parts of any simulator is assumed

Step 3 (Determining the limits for interoperability). Several approaches are suggested to simultaneously determine the parametric limits of interoperability. These include both experimental and empirical approaches. The empirical approach consists of coupling the acquisition of data (step 2) with a assessment of the user community's satisfaction with any given simulator network. The intent here is to discover a correlation between the level of satisfaction and parameter variability.

The second approach is to conduct experiments. The IST has developed and performed several experiments to determine the impact of parameter variations on operator performance. These are discussed in detail in section 2.2 of this paper.

Is should be emphasized that the process defined above is an interative one. Parameters are identified, measured and the appropriate limits defined. In the process of doing this, new parameters have been and will be identified while others are abandoned. We expect to ultimately converge on a set of significant parameters and limits which will be useable to define and verify interoperability for most simulators and tasks.

1.5 Rationale and assumptions

The IST believes that interoperability assessment procedures and methodologies can be developed which serve two purposes related to interoperability. Gathering interoperability parameter values is the first purpose. Assessing interoperability using the parameters gathered from other systems is the second purpose. The same procedures and methodologies serve both purposes. The IST made several assumptions that are both the justifications for and the basis of its interoperability parameter gathering and test procedures. These are listed as follows:

1. Interoperability is subject to decomposition and is analytic. This assumes that the lack of interoperability stems from more than one source and that interoperability is achieved by determining each source and taking steps to mitigate it.

2. The parameters contributing to or detracting from interoperability are quantifiable. This is a critical assumption. Obviously any parameter that is subjective rather than objective is probably not

controllable (unless in a statistical sense). The occurrences of such parameters, unless mitigated, preclude achievement of interoperability.

3. Perfect interoperability is not achievable but interoperability may be achieved at a level sufficient for most training tasks.

4. The acceptability of interoperability errors is dependent upon the magnitude and frequency of random and systematic errors, as well as the requirements of the task. For a given task, all players in a network of simulators should be equally interoperable (i.e., CGF players should not receive an advantage overhuman players).

## 2.0 INTEROPERABILITY DATA GATHERING METHODOLOGY

If interoperability is defined as a function of simulator differences per paragraph 1.2, it logically follows that the investigation of interoperability is an examination and quantification of those differences. Interoperability parameters are categorized by application. The purpose of the categorization is twofold; to reduce the range of parameter variability and to aid the analyst in applying parameters to specific simulator uses. We hypothesize two categorization strategies.

The first approach develops classes of tasks and determines the impact of parameters on those task classes. For military training, the classes chosen are: Static offense or defense, dynamic or moving defense and offense, navigation and reconnaissance. The approach involves comparison of performance parameters of simulators involved in these activities with other like or unlike simulators involved in the same activity. An example of this comparison is the determination of the impact of the presence or absence of a 360 degree out-of-the-hatch capability in a Bradley fighting vehicle has when exercising with a M1A1 simulator without such a capability.

The second approach defines the tasks in terms of the distance over which interaction occurs; interoperability issues are different for combat engagement of dismounted infantry than for tanks or artillery. Here the training tasks are categorized by range and the interoperability parameters would be examined and grouped accordingly.

### 2.1 Data acquisition

2.1.1 Hardware performance metrics. As part of the interoperability data acquisition process for a simulator or CIG, the tester (e.g., the IST) acquires all available data pertaining to simulator IG performance specifications. This includes such things as vendor data specifications, original procurement specifications and the results of functional configuration audits performed on military simulators. The list of parameters acquired borrows heavily from the ADST Interoperability Specification[7] that defines the minimum performance requirements for achieving defined ADST interoperability. The parameters included consist mainly of those performance items associated with the ability of the image generator to render complex scenes and update them in a timely fashion. These parameters include such factors as polygon loading, texel loading, rendering delay, number of graphics pipelines and the ability to include special effects such as animations and weather conditions[8]. Secondary functionality parameters are also acquired for comparison. This includes, for example, coordinate conversion algorithms that will then be examined for both precision and accuracy for comparison with other simulators.

2.1.2 Test procedures. The IST developed an initial set of test procedures that are used to acquire data for comparison between and among simulator image generators[9]. Our intent is to use the procedures to gather parameters from interoperable systems and use those parameters to establish criteria for testing other simulators. The procedures, must therefore be applicable to a large set of simulator types. Most of the parameters measured are for comparison of simulators of the same entity type. Most of the tests use the DIS network paradigm for test stimulus and data acquisition.

2.1.2.1 Terrain skin, culture and model correlation test procedure. A significant impediment to interoperability is variances in the virtual environment between simulators. These differences occur when either the source database definition is different, coordinate conversions are different, data base processing is different, or as a result of vendor-unique graphic pipelines[3]. We assume that the latter two differences are always be present. This test procedure measures the differences present in simulator virtual environments that exhibit a reasonable degree of interoperability to the user. The procedure uses those measurements to empirically define the limits of interoperability. The method utilized is statistical wherein parameters associated with the virtual environment (e.g., terrain altitude, extents of culture and model artifacts) are measured against a standard database and the resultant critical values are then compared[3]. Significant tests have been conducted on terrain skin which bear out the viability of our methods [2,3,6].

2.1.2.2 View port locations test procedure. The view ports of any real-world hardware (tank, helicopter & etc.) are rigorously defined by the technical data package. The location of such view ports on an entity simulator should be located likewise and should be consistent between simulators of a given entity type. This procedure determines the location of the view-ports in XYZ space. It consists of placing entities (via DIS ESPDU messages) at known locations relative to a viewport, observing their location within the view port and then triangulating to determine the exact location of the viewport. Each view port location is then compared to analogous port locations on other simulators of the same type to determine the consistency present as a parameter for interoperability. The assumption made here is a threshold of consistency is necessary for interoperability.

2.1.2.3 Field of view measurements test procedure. For any given view port in real-world hardware, a certain field of view is present. The simulator equivalent needs an identical field of view. The field of view is measured by this test procedure by placing an entity via a DIS network within the field of view of each viewport, moving it around and observing the boundaries of the field of view. The boundaries of each FOV are then compared between simulators of a given entity type to determine the level of consistency present.

2.1.2.4 CGF field of view test procedure. We test a CGF entity's observation capability for comparison with other CGFs and human operators. The test consists of instantiating a given entity type at a given location with a given attitude. An opposing force-type entity is then instantiated in the same vicinity and moved about until it is detected. Given the knowledge of view ports of the CGF entity type, comparisons are made to determine if the detection locus is consistent with other simulators of the same entity type.

2.1.2.5 Color space and contrast test procedure. Differences in the ability to render color effect interoperability. This is especially true in the case of camouflaged entity detection. To determine differences in the ability of simulators to render color, a set of tests was developed to measure the color range of the image generators for comparison.

2.1.2.6 Human factors/entity detection test procedure. The ability to detect opposing entities within a virtual environment needs to be uniform among simulators of a given type. To determine the degree of uniformity, a human factors test series has been created wherein an entity of a given type is approached by the simulator entity across the virtual terrain. The simulator operator identifies the entity type at the furthest range possible. We note the range at which detection and identification occurs and compare it to other simulators and other viewports to determine the degree of uniformity present. This test is statistical in nature and must utilize multiple entity types and multiple operators to ascertain a reliable average for the measured parameters.

The procedures have been written and executed at the IST their viability. The procedures are available for fielding in a "Beta Test" version. Only through widespread use can the procedures be refined and parameters necessary for interoperability be obtained.

2.2 Human factors research

A component of the IST research program is determining the limits of interoperability for IG-related parameters. The determination is achieved using human-performance metrics as the criteria for interoperability. The results of this research are combined with the data acquired from existing simulators to define limits for each parameter identified necessary for interoperability.

2.2.1 Networked battle experiment. This is one of the major efforts in our current interoperability investigation. Through it, the IST hopes to establish and experimentally justify acceptability limits on parameters impacting interoperability. It is doing this by a series of experiments that measure the impact of parameter variation on human performance in a simulator. Two of the experiments have been completed but the series is not complete and will not be complete until relevant parameters have been identified (see paragraph 1.1). The IST created a simplified tank simulation on each of a pair of Silicon Graphics Indigo Extreme work stations that communicate via the DIS paradigm. In the current experimental suite, the parameter being measured is terrain correlation. The virtual environment in which the simulations operate is varied between the work stations by a quantified amount (i.e., terrain mis-correlation). Two operators engage in a series of exercises (target shooting and combat engagement). Various metrics are being gathered to determine the most relevant. This includes such things as engagement outcome, time of engagement, number of shots fired, target detection and identification ranges and hit-to-miss ratios. The results are being compared to baseline cases (e.g., an engagement where there is no terrain mis-correlation between the simulations) to determine at what level of terrain mis-correlation are the human performance metrics significantly impacted. It Is our plan to continue this approach to defining interoperability parameter limits to cover all major parameters. It is also our intent to simulate and study various solutions to mitigate interoperability problem when all of these tests are complete [5].

2.2.2 Complex scene study. Image generators vary between vendors. Images will also vary. To achieve interoperability, these differences must be kept below some threshold. The thesis of the complex scene study was that the threshold is at that level where the differences are not noticeable. To determine that threshold, an experiment was designed and executed by the authors to determine the threshold and relative importance of the following classes of differences: Color, position, texture and luminance. This study has been completed. Two scenes (Figure 2.2.2-1) were created containing a variety of natural and man-made objects (houses, trees, bushes, rocks and clouds of various sizes). The following classifications of objects were made: Houses, large trees, medium trees, small trees, rocks and clouds. Within the original scene there were at least three instances of each change class that occupied the same amount of area on the display for each object class. In the altered scene, these objects were modified in various ways (e.g., size, position, texture, color) by quantified amounts. In the case of position changes, only horizontal changes were made. In a two-dimensional display it is impossible to differentiate between a Z-axis (i.e., toward or away from the viewer) and a change in size. Vertical shifts were not used for a similar reason; there would be ambiguity between such a shift and a Z-axis shift. These scenes were then rendered on a pair of Silicon Graphics Indigo Extreme workstations and presented to a series of twenty subjects who were directed to identify the differences. The noted differences were recorded in the order in which they were noticed. The subjects were university students and staff members working at the Institute for Simulation and Training. Each subject was screened for color blindness prior to participation in the test. Each subject was also screened to eliminate those persons with exceptional training in visual observation (e.g., law enforcement personnel, visual artists, & etc.) who, because of heightened awareness or sensitivity to visual cues and details, are non-representative of the population at large and would skew the test results.

The scenes were created utilizing a terrain database about point “Linus” at Ft. Hunter-Liggett with additional objects (e.g., houses, trees & etc.) whose attributes were controlled and quantified between the original and altered scenes. The scenes were created with Multigen software on a Silicon Graphics Onyx Reality Engine workstation and captured with the resident snapshot utility. The actual test was performed on side-by-side CRTs in a room with no light source other than the CRT’s.

**(a)**

(b)

Figure 2.2.2-1 Complex scene experiment (a) baseline image and (b) image with alterations.

2.3 User community perspectives

Having acquired performance data through both test and literature, the IST is polling the user community to determine the degree of satisfaction that exists in a given system of networked simulators. If satisfaction exists (e.g., if the user considers the networked system 'interoperable'), the IST gathers specific values for each interoperability parameter. Limits on interoperability are determined by sampling a wide variety of simulators under various usage conditions. Limits are then compared to the in-house experimental results establishing specific limits on performance parameters defining an interoperable state. This activity is in process.

## 3.0 RESULTS

3.1 Networked battle experiments.

3.1.1 Mis-correlation versus observation range. The first experiment quickly performed as part of the Networked Battle experimental suite involved a determination of the range at which a constant terrain bias could be detected. An entity was placed in a position at extreme range from the observer with an altitude offset from ground level of between plus and minus one meter. The observer/subjects were then moved towards the entity in virtual space until he could correctly identify the offset. This experiment was repeated under various conditions (e.g., with and without terrain texture, shadowing and with high or low image resolution). The results are shown in Figure 3.1.1-1 and documented in a report at the IST[10]. The data indicates that a change in entity vertical position relative to ground level of as little as 0.5 meters can triple the defilade detection range from 400 to 1200 meters.

This experiment constituted a baseline test to determine the range at which mis-correlation is noticeable. Although not widely useful, the data indicates that if the detection of defilade at a given range is an important parameter in a training exercise, that the task will be very sensitive to terrain mis-correlation

3.1.2 Network Battle results. Eight subjects were used in an experiment involving a combat gunnery style exercise between M1 and T62 tank entities over various ranges in the presence of mis-correlated terrain where the mis-correlation consisted of a constant altitude offset between the two simulator's virtual environment. Ranges were classified as close (less than 700 meters) and far (greater than 750 meters). The exercise platforms consisted of one Silicon Graphics Indigo Extreme and a PC-based CGF that communicated via the DIS paradigm. Terrain varied between the two platforms by an offset of up to ±3 meters. The results of this particular experiment (Figure 3.1-1), indicate that a vertical (altitude) error in the terrain of -1.0 meters or more between platforms impacts the ability of a subject to hit a target as evidenced a significant change in the hit-to-miss ratio. These results appear to be relatively insensitive of target range.

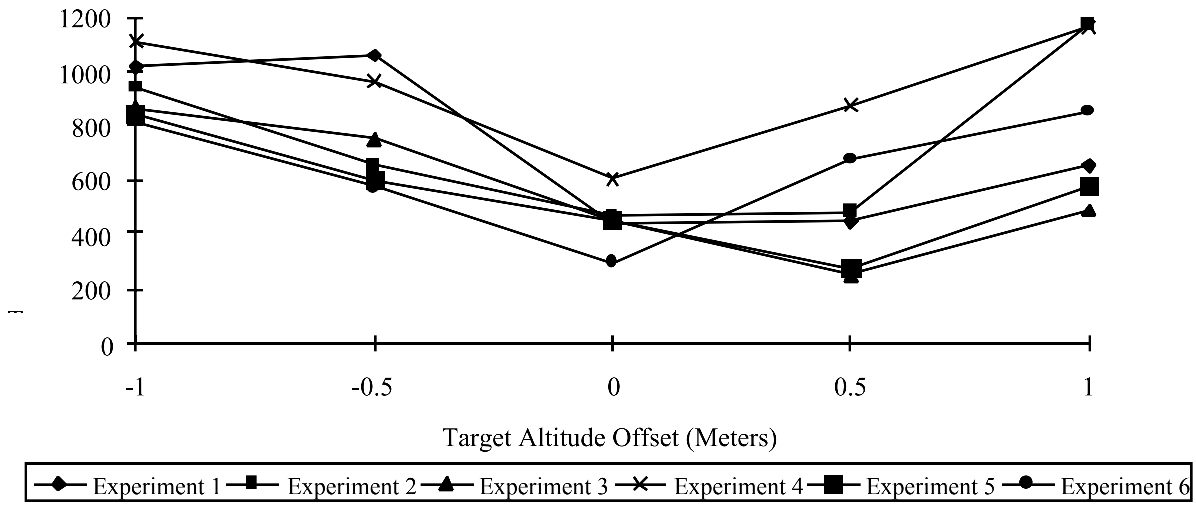


Figure 3.1.1-1 Target vertical position detection versus detection range.

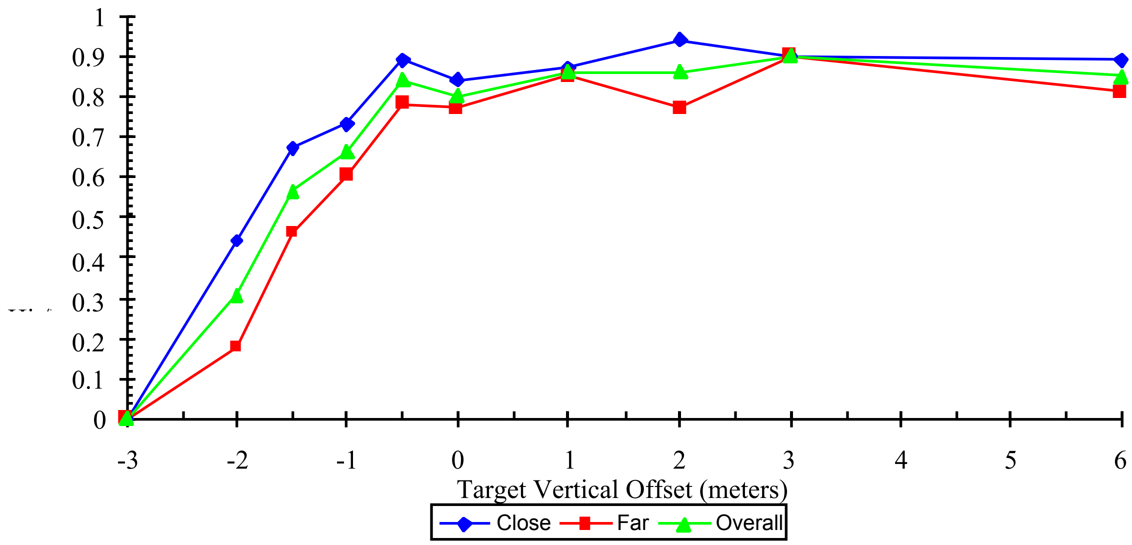


Figure 3.1.2-1 Effect of terrain offset-type miscorrelation on a tank combat engagement (shooting hit-to-mis ratio) at near and far proximity.

The utility of such results is straightforward: Terrain mis-matches of greater than 1.0 meter are not acceptable for training exercises that involve gun engagements if the target is instantiated at the one meter mis-matched location. Another possible interpretation is that when 30% of a target is hidden due to terrain mis-correlation, training is adversely affected.

Generalizations and the utilization of these results should be held in abeyance until the investigation is complete. Constant bias errors are rare between databases. The next stage of investigation includes studies of errors between databases that exhibit variations that are quantifiable using statistical techniques. The errors may be random, induced from the data base source, or through coordinate conversion, data base processing, or graphics pipeline processes. The variations, though, appear random. This is the typical situation found in networked simulators. The IST previously developed the tools for the measurement of such mis-correlations and is presently working of ways to mitigate them. These tools are being supplemented by tools for determining image shifts, rotations, dilations, and skewness. When the investigation is complete, the results can be directly applicable to the interoperability certification process.

3.2 Complex scene experiment.

The complex scene experiment is complete and documented in an internal IST report[11].

3.2.1 Relative noticeability. A histogram illustrating the relative noticeability of parameters is given in Figure 3.2.1-1. For the suite of change classes and magnitudes utilized (i.e., the magnitude of changes at or about the threshold of detectability) in this study, the results indicate size is the most important factor. Size changes were noticed more often than any other change class as the first five things that were noticed by the subject population. The second most important parameter appears to be object position. The category of 'other' consists of identification of differences that were not actually present. Following the 'other' category color changes were most noticeable.

3.2.2 Luminance change noticeability. The results of the test for luminance change noticeability are given in Figure 3.2.2-1. Test results indicate that there is a threshold of change detection for color luminance of about 12 percent from the baseline value. Six items' colors were modified: The garage door on each of the three buildings and three rocks in the foreground

3.2.3 Size change noticeability. The results of the test for size change noticeability are given in Figure 3.2.3-1. Test results indicate that there is a threshold for change detection of size when the change encompasses about 0.2% of the display.

Nine items' sizes were modified: The three buildings, three large trees and three clouds in the sky

3.2.4 Position change noticeability. The results garnered relevant to positional shifts are the most intriguing. Although they were the second most likely change class to be noticed (Figure 3.2.1-1), there appears to be little or no correlation between positional change detection and any other parameter. Correlation of the positional shift detection was attempted with size of the change, sized of the object, the number of other proximate objects that could be used as a horizontal reference and their mean distance. No strong correlation was found. No explanation is given for this phenomenon; it is counter-intuitive. This is an area where further study is indicated. Figure 3.2.4-1 is a plot of the magnitude of the positional shift versus the detection rate

3.2.5 Texture change noticeability. Three objects had texture changes that consisted of magnification of the texture. These were the roofs of the three buildings in the scene. These textures were magnified by factors of 15, 25 and 30 percent respectively. One subject noticed the change at both 15 and 30 percent. Three rocks also had a texture applied, the hue of which was modified by 25, 30 and 40 percent respectively (HLS color system). None of these changes were detected. It is concluded that for the size of objects in this scene that had texture modifications, that the change magnitude was below the threshold of detection. Were this experiment to be repeated, larger shifts would be employed to determine the threshold of detection.

3.2.6 Foreground versus background change noticeability. The scene was arbitrarily divided into a foreground and background with the vertical half-way point on the screen as the dividing line. This places the leftmost and rightmost structures in the foreground and the middle structure in the background. The number of items that were modified in each area was determined along with the number of instances of detections of those changes. The purpose of this was to determine if there was any significant difference in detection rates as a function of position on the screen. The data was normalized to give the detection rate per unit object present in each given area. Nine objects in the foreground were modified. These were detected a total of 40 times (out of a possible 180 times--20 subjects times 9 objects). This gives a foreground detection rate of 4.4 detections per unit changed object-observer. Fifteen background objects were modified and were detected 32 times (out of a possible 300) giving a normalized foreground detection rate of 2.1 detections per unit changed object-observer. One could draw the conclusion that foreground changes were detected at a rate 2.1 times higher than the background items. However, the original thesis of the investigation was closely coupled parameter relationships. Given a different distribution of the types of changes between foreground and background, the results might vary significantly. This is only one data point, but in the absence of other data it can be used as a guide as to where care should be placed to achieve scene fidelity.

3.2.7 Regular versus chaotic change noticeability. Within the scene there were 9 artificial objects whose geometry can be defined as symmetrical. They were detected a total of 40 times (out of a possible 180) to give a detection rate of 4.4 detections per unit symmetrical object-observer. There were also a total of 15 modified chaotic (natural) objects such as trees, rocks and bushes that were detected a total of 32 times (out of a possible 300). This gives a detection rate of 2.13 detections per unit chaotic object-observer. One possible conclusion is that a difference in a symmetrical object is twice as noticeable as a difference in a chaotic object. Again, the original thesis of the investigation was closely coupled parameter relationships. Given a different distribution of the change classes between chaotic and symmetric objects, the results might differ.

If no further tests were to occur, the following conclusions could be drawn from the existing data:

a. The order of change noticeability (beginning with the most prominent) is as follows: Size, position, luminance and texture. With the exception of position, this comparison is relevant only for differences in the neighborhood of the threshold of detection. (See Figure 2.2.1-1).

b. The threshold for size change detection occurs when the change constitutes 0.2% of the display area.

c. The threshold for luminance change detection occurs at 12% relative to the baseline luminance.

d. There does not appear to be a well-defined threshold for position change detection. In spite of this, it is the second-most important parameter after size insofar as change noticeability is concerned (Figure 4.0-3).

e. Foreground fidelity is more important than background fidelity.

f. Symmetrical object fidelity is more important than chaotic object fidelity.

3.2.8 Utility of results. The following recommendations are derived from this study to improve the realism of terrain databases.

a. High priority should be placed on matching object size with real-world counterparts. Objects should match the size of their real-world counterpart (or with other image generators to within 0.2% of the area of the display or better.

b. The difference in luminance of objects between image generator scenes or between the scene and its real-world counterpart should not exceed 12 percent.

c. Proper positioning of objects is exceeded in its importance only by proper object size, even though no error threshold has been discovered.

d. Greater emphasis should be placed on the fidelity of objects in the foreground since it appears that differences are more noticeable. Foreground is defined as the bottom half of a CRT display.

e. Greater emphasis should be placed on the fidelity of symmetrical objects (e.g., man-made objects) over chaotic objects (trees, rocks bushes & etc.); differences in symmetrical objects appear to be more noticeable.

These results are directly useable by vendors in the creation of terrain databases and by the government in defining acceptable quality levels of the same. These criteria could also be used as the definition of acceptable levels of correlation for the parameters investigated.

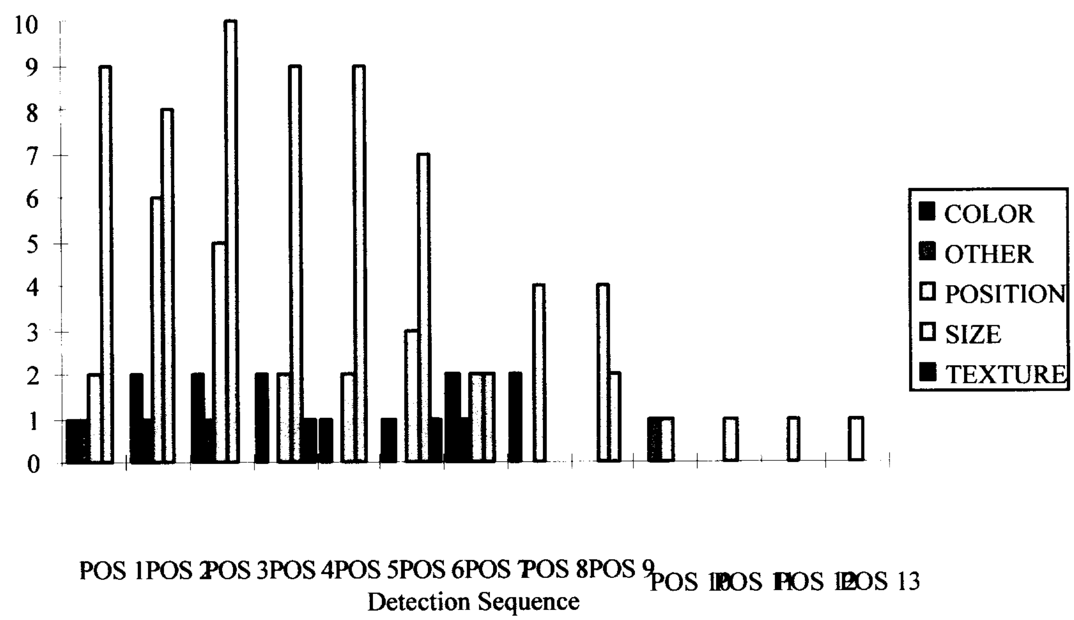


Figure 3.2.1-1 Relative noticeability of difference classes. Pos (position) indicates the sequence in which a difference was noted (e.g., Pos 1 differences were noticed first by a given subject)

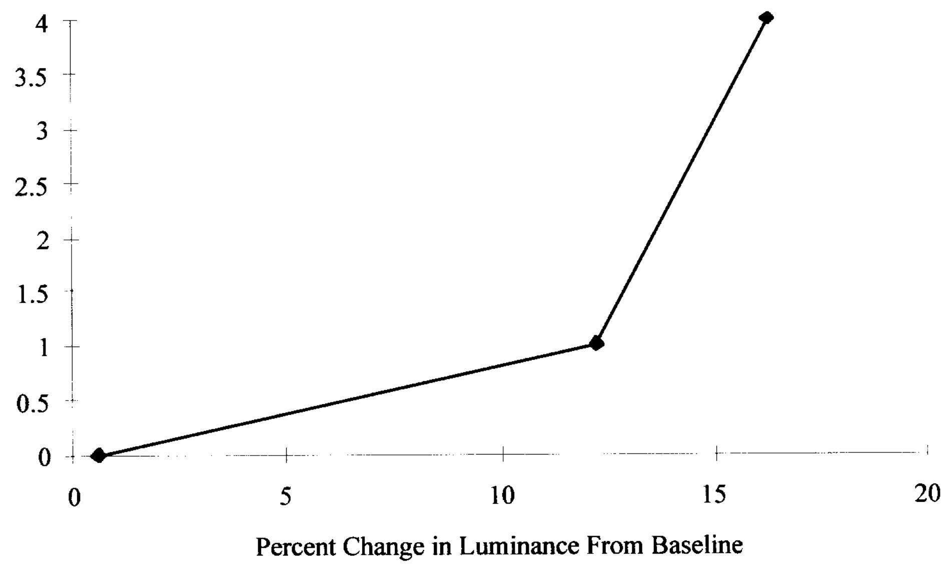


Figure 3.2.2-1 Noticeability of luminance changes. The y-axis is the number of subjects that noticed the difference for any given difference magnitude. A total of 20 detections was possible for each change level noted.

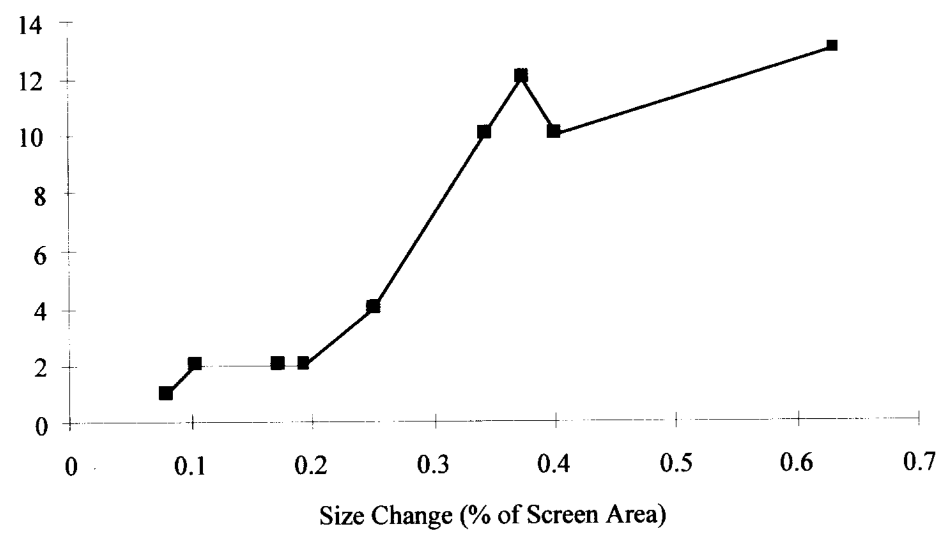


Figure 3.2.3-1 Noticeability of size changes. The y-axis is the number of subjects that noticed the difference for any given difference magnitude. Twenty detections were possible for each change level noted.

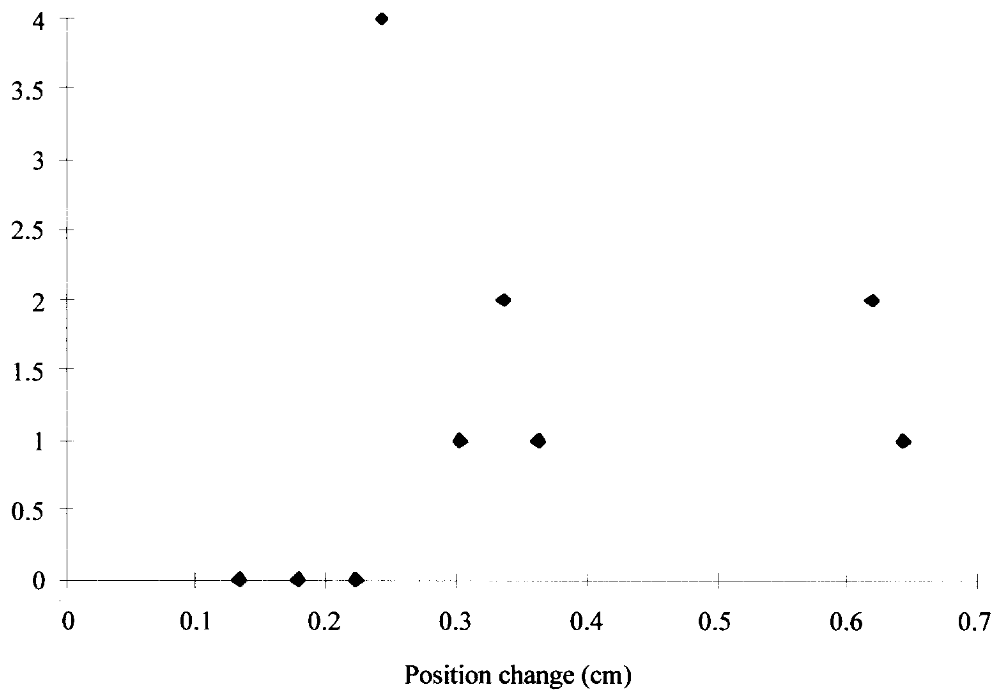

Figure 3.2.4-1 Noticeability of position changes scatter plot. The y-axis is the number of subjects that noticed the difference for any given difference magnitude. The x-axis is the magnitude of the position change.

## 4.0 FUTURE EFFORTS AND CHALLENGES

In order for achieve the solution to the interoperability problem proposed by this paper there are several activities that need to be completed:

Data acquisition from operational simulators. Access to the existing simulator facilities and hardware is essential in order to acquire data for the empirical determination of the limits of interoperability parameters. Access must be granted to a large enough number of simulators to make the data statistically significant.

Research extensions. There are many simulators that neither use nor depend upon image generation of a synthetic environment. Computer generated forces and radar simulators are examples of this class. These systems need to be studied and interoperability parameters identified and bounded to ensure that they are compatible with other types of simulators.

Tool development. Tools need to be developed to support several portions of this investigation: (1) To either enable or facilitate the data acquisition and analysis process, (2) To enable performance of the tests methods in a timely manner and (3) To provide economical means of problem mitigation or solution once the interoperability problems have been identified.

## 5.0 CONCLUSION

The IST believes that heterogeneous simulator interoperability is achievable in a practical sense by decomposition of the problem to an atomic or parametric level. At that level the problem becomes quantitative and hence subject to measurement and verification. The task that IST is engaged in is the identification of the salient parameters and the determination of appropriate limits of the same. The approach taken is two fold: Empirical acquisition of data from existing simulator networks and the satisfaction of the user community with the same and a laboratory-based experimental approach. A test procedure prototype has been developed and executed which will be used in the empirical data acquisition portion of the task. This procedure has been executed in-house and has proven to be viable. Several experiments have also been performed in house that have yielded data that is immediately useful to the user community to facilitate interoperability. Both approaches are an on-going effort.

## ACKNOWLEDGMENTS

This paper and the research associated with it were done under funding of contract N61339-93-C-0076 by the U.S. Army Simulation, Training and Instrumentation Command (STRICOM) whose support is gratefully acknowledged. The data and positions advocated by this paper, however are those of the authors and do not necessarily reflect the position of STRICOM on the subject matter.

## REFERENCES

1. Spuhl, Karl A., "Correlation Considerations in the Simulation Environment", Position paper presented to the 11th DIS Workshop, Orlando, September 1994.

2. Goldiez, Brian, Russell S. Nelson, "Terrain Database Correlation", International Training Equipment Conference and Exhibition Proceedings, The Hague, The Netherlands, April 1994.

3. Schiavone, Guy, Russell S. Nelson, Brian Goldiez, "Statistical Certification of Terrain Databases", Interservice/Industry Training Systems and Education Conference Proceedings, November 1994

4. "A Report on the 1994 I/ITSEC Distributed Interactive Simulation Demonstration" Institute for Simulation and Training, 10 May 1995.

5. Purdy, Richard, "A PDU Solution to the Inter-Visibility Problem in Distributed Interactive Simulation due to Miscorrelated Terrain", 12th Workshop on Standards for the Interoperability of Distributed Simulations, March 1995

6. "Demonstration of Dissimilar Simulator Networking' Institute for Simulation and Training, 14 July 1994.

7. "Interoperability Functional Specification, Version 1.3 (draft)" Loral Systems Company, 30 October 1994.

8. "Interoperability Functional Specification, Version 1.3 (draft)" U.S. Army Simulation, Training and Instrumentation Command, 30 October, 1994.

9. Purdy, Richard, Brian Goldiez, "A Comprehensive Methodology for Attacking Simulator Interoperability (draft)" Institute for Simulation and Training document CR-95-21, 30 May 1995.

10. Moskal, Patrick J., Teresa R. Johnson, Guy Schiavone, Pamela Woodard, "Target and Error Position Identification Under Variable Viewing Conditions", International Training Equipment Conference and Exhibition Proceedings, The Hague, The Netherlands, April 1995

11. Purdy, Richard , "A Report on the Complex Scene Vignette Investigation" Institute for Simulation and Training report IST-CR-95-22, 25 May, 1995