COMPUTER GENERATED/SYNTHESIZED IMAGERY (CGSI)

Carl P. Graf
Senior Principal Research Scientist
Honeywell Inc.
Minneapolis, MN
and
Dorothy M. Baldwin
Physicist, Advanced Simulation Concepts Division
Naval Training Equipment Center
Orlando, FL

ABSTRACT

Feasibility has been demonstrated for a hybrid simulation approach which merges two technologies, Computer Generated Imagery (CGI) and Computer Synthesized Imagery (CSI), to form Computer Generated Synthesized Imagery (CGSI). This approach holds promise as a cost-effective, attainable method of providing real-time, high detail imagery for visual and/or other sensors, such as FLIR. A videotape for a nap-of-the-earth flight was generated, demonstrating the fidelity and mobility that can be achieved using this CGSI hybrid approach. A description of the approach will be provided along with selected video frames. Critical features which were demonstrated include: merging of the two technologies, vertical movement for target acquisition and landing, horizontal movement towards and away from objects, dynamic occulting of 3-D objects, dynamic smoke and dust, and color visual and IR imagery.

INTRODUCTION

The advanced development and increased use of such sophisticated sensors as: thermal imaging; low light level TV (LLLTV), and imaging radar have significantly expanded the capabilities of military platforms. Sensor technology developments have contributed both to safer and more effective operations and tactics. These gains in weapons system capability, however, can only be fully realized through effective training programs. This training is more difficult because, perceptually, sensor imagery, such as infrared, is novel to observers and therefore more difficult to use and interpret. This problem is compounded under high data rate/workload conditions such as low-level flight. Training and exercising with the actual sensor systems in the real environment are certainly valuable training and contribute to the effective employment of the sensor system in each mission phase. However, the cost and availability of fuel and other training resources, combined with risk, have resulted in the reduction of actual platform training hours and, consequently, the opportunity to exercise the actual sensor systems. In light of this situation, more attention must be directed to the development and use of simulators. One of the challenges of simulation today is the effective, coordinated simulation of advanced sensor systems for training high data rate/workload conditions such as low-altitude flight missions, nap-of-the-earth (NOE) and shipboard operation. In addition, a highly desirable factor in any sensor simulation approach is the ability to correlate with an out-the-window visual scene.

To achieve the required level of simulation, designers must consider a variety of factors including the aircraft itself, its mission, the sensor system, and the environment in which the system will be employed. In addition, there are other factors related solely to the actual simulation of sensor operation which must be considered.

Hybrid system concepts such as Computer Generated Synthesized Imagery (CGSI) offer near-term solution for simulating sophisticated sensors and correlated visual/sensor imagery. Because of the high potential pay-off from the real-time implementation of CGSI, NAVTRAEQUIPCEN, with support from PM TRADE, has funded Honeywell's Systems and Research Center to provide a non real-time demonstration of CGSI for a NOE flight and a design for real-time implementation of CGSI. This report summarizes the results of the effort.

This section describes the merging of the Computer Generated Imagery (CGI) and Computer Synthesized Imagery (CSI). CGI uses the computer to generate imagery from a data base. CSI uses the computer to insert targets/objects in real-world pictures. CGI provides excellent control of a scene to be constructed and displayed for interaction in a simulation environment. However, the fidelity is low, and as a result, realism in the displayed scene is poor. CSI is just the opposite. Fidelity is high, but control over scene construction is restricted. Figure 1 represents the fidelity/viewpoint flexibility of CGI and CSI and the resultant CGSI (Computer Generated Synthesized Imagery).

Figure 1. A Computer Generated Synthesized Scene

Contributions from Computer Generated Imagery (CGI)

The strength of CGI is its ability to generate surface representations. A real or artificial surface can be measured to get elevations at specified points, usually at intersections of a uniform grid. The surface can be constructed in a computer by connecting the sample elevations.

In addition to realistic surface representations, CGI offers control over the placement of objects on the surface. Since the data of elevations is usually provided with a uniform grid, the placement of other objects can be specified on this same grid. Trees, rocks, shrubs, houses, and roads can all have their positions defined in the data base grid system.

Correct illumination and perspective are the final two major contributions from CGI. Correct illumination is achieved by finding the surface normal for each pixel displayed. This normal is used along with line-of-sight and the normal from the illumination source plus amibient intensity and haze factors to compute pixel intensity values. Correct perspective is achieved because the distance is a significant variable in the perspective transformation.

Weakness of CGI

Although the position of an object can be accurately specified, correctly illuminated, and displayed in correct perspective, the fine detail of an object cannot be realistically represented. The current state of the art in CGI object representation is such that objects appear cartoonish. Some scene elements, such as barren terrain, sand, and clouds, can be represented more realistically than highly detailed structured objects like trees and grass.

Contributions from Computer Synthesized Imagery (CSI)

CSI uses photographs of real scenes. The objects in the scenes are not represented individually; nor is the scene modeled by elevation profiles. Usually the scene is held static, while single objects, like aircraft or tanks, move within the scene. The path followed by the moving objects is carefully laid out. In each frame, object occlusion is resolved by deciding which objects are in the foreground and which are in the background. This is a manual procedure.

The strength of CSI lies in its use of photographs. With currently available video equipment, the photographic data can be manipulated in a manner previously considered impossible. Individual photographs can be stored on a video disc, along with hundreds of other photographs. Access is controlled by an index just as with digital data stored on magnetic disc. Thus, a frame can be constructed by calling up individual photographs and merging them in a frame buffer.

Weakness of CSI

The high-fidelity CSI scenes are limited to the view point of the camera. That is, one cannot drive through a scene unless a series of through-the-scene photographs is used. For any reasonable size gaming area, the number of through-the-scene photographs become prohibitive. For limited applications with fixed observers, CSI may be sufficient.

CGSI: The Merge

CGSI combines the best features of both techologies: CGI and CSI. A scene is constructed by placing individual high-fidelity CSI objects on a specified CGI surface. A CGSI scene is constructed much like a CGI scene. The surface elevations and object locations are laid out on a uniform grid. The individual objects used in the scene are transformed for perspective and size. This includes size, position, rotation, warp, and intensity transformations on the image. The surface may be a CGI texture or a series of CSI surface inserts. The scene is constructed by placing the farthest object first and continuing with overlays until the nearest objects have been placed. The CGSI scene may be constructed with imagery from any portion of the spectrum—visual, infrared, millimeter, or radar frequencies.

HOW TO BUILD A CGSI SCENE

This section demonstrates the construction of a typical CGSI scene through eight steps. After the eight steps are completed, the resulting scene is provided to the video monitor. This process is repeated at the 60-Hz field rate.

The construction of a CGSI scene begins with the surfaces, ground, and/or sea and sky. The sequence continues with the addition of objects, both small and large. The objects may be trees, rocks, bushes, houses, roads, lights, vehicles, ships, docks, airplanes, etc. Finally, the special effects are added which include smoke, dust, clouds, and shadows.

Sky (Figure 2)

Sky is added in segments over distant background. By breaking the sky into segments, peaks, and valleys form the skyline shown. In this example, the sky was broken into five segments. In general, the lower edge of the segment does not need to be straight, but may be curved or jagged to simulate rolling or sharp hills or mountains. The individual segments are warped based upon minimum and maximum data base elevations and upon viewpoint.

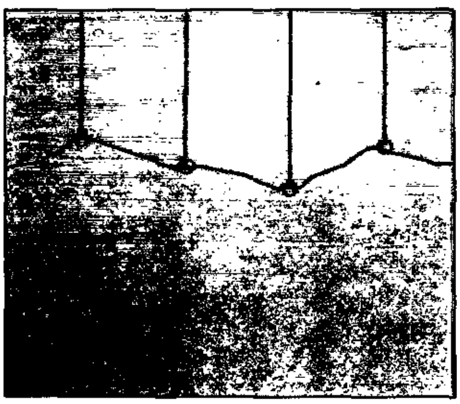

Figure 2. Sky Scene

Textured Surface Scene (Figure 3)

Textured surfaces are also added in segments to form the foreground surface. The untouched region between the sky and foreground forms the distance background. Stored textured surfaces, warped to fit the screen coordinates of the surface polygons, are then added to the screen. The intensity of the surfaces is varied based upon the range or other parameters.

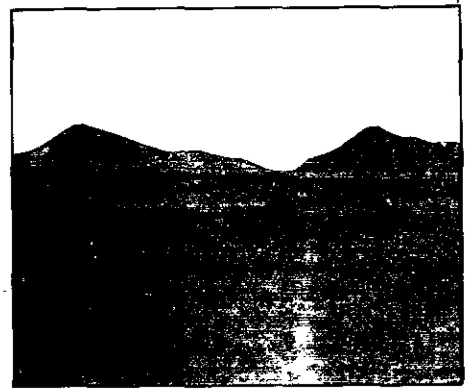

Figure 3. Textured Surface Scene

Road Scene (Figure 4)

The surface library may also contain special surfaces such as roads, streams, and ponds. In this example, roads in the data base library were warped to fit the screen coordinates. This scene also contains four textured surfaces.

Figure 4. Road Scene

Small 2D Objects (Figure 5)

Small two-dimensional (2D) objects, less than 1/16 of the scene's area, are added. Most natural objects such as trees, bushes, and rocks may be represented from one side. These are called 2D objects. Objects which cannot be represented from one side such as houses, tanks, ships, etc. are referred to as three-dimensional objects (3D). Small objects are processed by less expensive processing hardware/software than large objects and surfaces. During the flight through a scene, the 2D object will be handed off to a large 2D processor when it occupies more than 1/16 of the area of the scene.

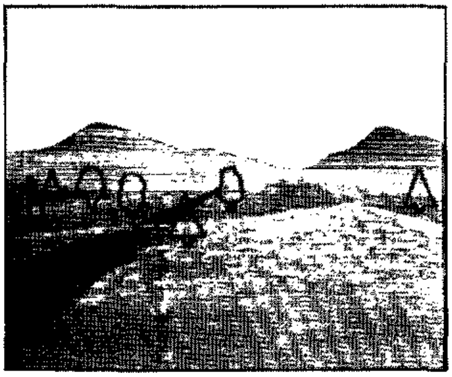

Figure 5. Small 2D Objects

Multi-View Object Scene (Figure 6)

The tank is an excellent example of a multi-view object. Multi-views of the tank are stored and the correct view, based upon the tank path, elevation, and observer's viewpoint, is used in constructing the scene. The tank may be moving and may be very large.

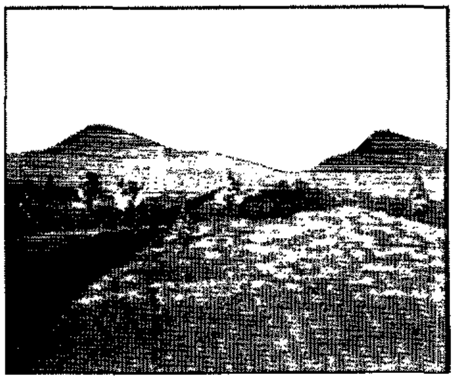

Figure 6. Multi View Object Scene

Multi-Surface Object Scene (Figure 7)

The house is an excellent example of a multi-surface object. The house is separated into three surfaces, two roof segments (one if both sides are identical), two ends, and two sides. The individual surfaces of the house are warped from a normalized view to the perspective dictated by the screen coordinates and then joined together. Images of this type may be constructed for enemy objects if intelligence information is presented.

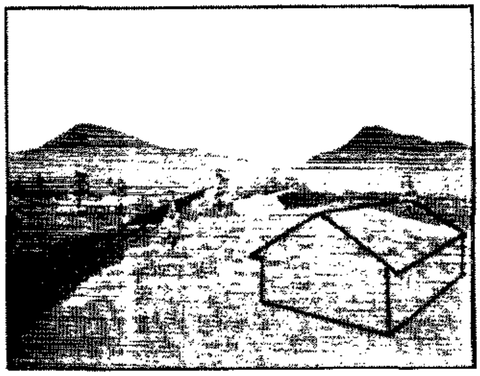

Figure 7. Multi Surface Object Scene

Large 2D Object Scene (Figure 8)

Large 2D objects typically occupy more than 1/16 of the scene's area. These objects may be expanded so that an object will cover more than the entire surface of the screen.

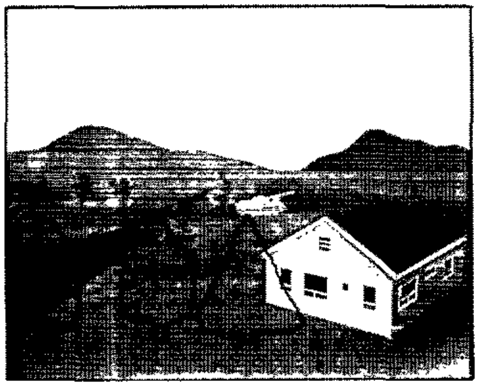

Figure 8. Large 2D Object Scene

Special Effects Scene (Figure 9)

Special effects are used for translucent surfaces which include clouds, dust, smoke, and shadows. A mask controls the transmission functions and a second input word controls the intensity and color.

Figure 9. Special Effects

Complete Scene (Figure 10)

This completes a CGSI scene.

Figure 10. Complete Scene

CGSI SYSTEM OVERVIEW

Figure 11 is a functional overview of a real-time CGSI system.

Data Base Construction

The data base consists of two very different types of data: the object library and the gaming area.

The object library contains images of objects and surfaces, and transmissivity masks of special effects from one to many bands of the spectrum. This allows the simulation of not only the visual domain, but also infrared, millimeter and radar frequency sensors. The object library may also contain a mixture of 2D and 3D images. The images may contain a variety of day/night and diurnal conditions. The object library consists of many actual high-fidelity images taken in the real world from sensors and stored in a photographic manner. In constructing a high-fidelity object library, images from individual real-world elements, highly accurate models, artist drawings, photographs of enemy devices, etc., are restored (correct edge roll off) to form "near-perfect" images. This is achieved by operator and machine controlled thresholding (dropping out the background) intensity corrections, realistic color control and perspective normalizing. Ground contact and height reference points are also added. The "near-perfect" objects, surfaces, and special effects are stored on a rapid access and high-speed data rate media.

The gaming area data base provides the information necessary for placing the contents of the object library, objects, surfaces, and special effects on a grid or gaming area. The objects may be placed by an operator or in a random manner by the computer. The objects in the library may be either stationary or capable of movement.

The output of this function determines ***"What is in the scene."***

Vehicle Simulation Computations

The vehicle simulation computations, based upon the vehicle math model and control inputs, determines the locations and viewing direction of the visual or sensor system for the primary vehicle. In addition, computation may be performed on secondary vehicles based upon vehicle models and selected paths. The output of this determines, ***"Where am I?"***

Systems Interface

The I/O of the vehicle simulation system and I/O of the CGSI system must interface in an efficient manner. The communication subsystem is a bi-directional link and buffer interfacing the two systems. This function is the handshake and data flow between the systems, ***"Communications."***

Field of View (FOV) and Coordinate Transform Computations

The FOV processor determines the presence of objects, surfaces, and special effects in the scene under construction. The output of a transformation matrix (V) converts real-world coordinates to screen coordinates. This data from the transformation matrix permits rapid testing and determines if all or any portion of the objects, surfaces and special effects are present in the scene. To avoid testing for the presence of all the objects in the data base, a "smart" algorithm tests only those objects or surfaces which are in the proximity of the scene. The FOV processor maintains a list of objects in the FOV and their object, surface or special-effect channel assignment. The function of the FOV computer is to determine, ***"What can I see?"***

Controllers—Object/Surface/Special Effects

The controllers fan out and process the control functions generated during the FOV computation. The processed control functions are passed to the object/surface/special effects processing channels. The main functions performed by the controller are the transform of gaming area corrdinates to screen coordinates, range from the vehicle to each object in FOV, intensity of each object based upon range, object identification, and commands to the object library base for the retrieval of the correct image data. The function of the controller is to ***"Fan out FOV data and generate precise control data."***

Object/Surface/Special Effects Library

The library stores the images used to construct a scene. The controllers command the selected images

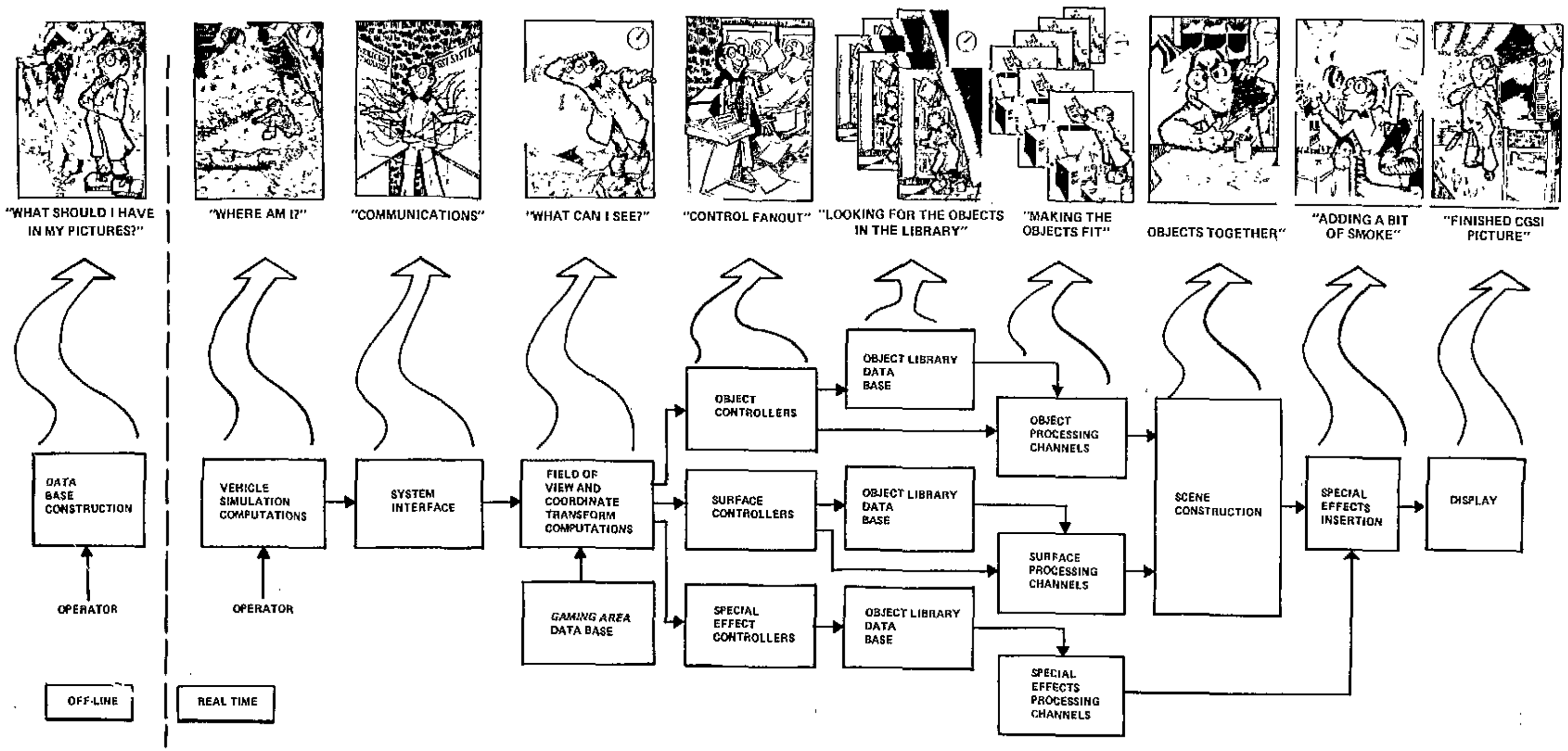

Figure 11. CGSI Functional Overview

which are passed to the processing channels. The function of the library is to *"Provide the correct image upon command."*

Object/Surface/Special Effect Processing Channels

The individual processing channels pipeline processors process one object, surface or special-effect per channel. All the processing channels operate in an identical manner; large object, small object, surface or special effects. Each processing channel modifies objects, surfaces, or special effects from the object library by the transformation specified by the control functions. That is, the object, surface, special-effects processing channels change a stored image (normal perspective) to scene conditions (screen coordinates) by changing image, position, size, rotation, and warp. Image intensity is modified based upon range and object type. The function of the parallel pipeline processing channels is to *"Modify each object, surface, and special effect used in the scene."*

Scene Construction

The scene construction module takes the individual image from each processing channel, separates the image from the background, and assembles the scene based upon range. Near objects occlude more distant objects. The high-frequency edges generated by assembling a scene from individual images are smoothed by a Gaussian function. This operation matches edge and internal frequencies.

This function receives range information from the two object controllers. This range is used to determine whether or not a particular object is in front of, or behind other objects in the scene. If the particular object pixel is the closest occupied pixel in the scene, then it will be the pixel displayed. This may be termed a *"nearest"* algorithm.

The scene construction function accepts video inputs from each video channel, and from a background level source defined by the FOV computer. This function outputs real-time video signals to the special effects.

The digital scene construction function contains the following subfunctions: 1) object channel combination, 2) scene intensity adjustment to accommodate scene-wide intensity corrections, and 3) smoothing to compensate for object-to-object and object-to-background boundaries. The function of the scene construction module is to *"Assemble the picture."*

Special Effects

The translucent special effects are added after the generation of the scene. The special effects module adds the special effects based upon range. Special effects, such as smoke, or dust, may occur ahead of, or behind images in the scene. The intensity masks stored in the object library and processed in the special effects processing channel control the transmissivity of the special effects. The intensity value input controls the intensity/color of the special effects: the black smoke and white clouds. The function of the special effects module is to *"Add dynamic or static translucent effects."*

TECHNOLOGY OVERVIEW

Figure 12 presents a hardware/software overview. It shows that:

1. Most of the hardware exists as a complete subsystem or at the card level.
2. The system does not contain large amounts of software.
3. The system has common components; the same hardware and software is used for controlling surfaces, objects, and special effects.
4. The system is modular. The technology can be configured for many different applications by adding or subtracting the modular components.
5. The input images are real.
6. The real-time requirements of 30 frames/second and 100 millisecond transport delay are achievable.
7. The key components in the system are the object/surface/special effects processing channels.

Object/Surface/Special Effect (OSSE) Processing Channel Hardware

Because of the extreme importance of the OSSE processing channel hardware it is discussed in greater depth.

To obtain the correct intensity, color, image, size, location, rotation, and perspective the following functions are performed on the library data:

a. A high-speed (approximately 100 nsec samples) analog-to-digital converter converts the object image to a digital format. The digital format has 512 pixels per line, 480 active lines (525 total) and eight bits per pixel (256 gray shades).

b. A high-speed memory card accepts the digital data in either the X or Y axis. The axis and direction of loading is dependent on the rotation of the image. The data is loaded to minimize pixel compression during the processing passes. Rather than rotate an image 60 degrees which may result in some image loss, the data is loaded in the perpendicular axis (at 90 degrees) and rotated 30 degrees. This memory card also holds the object image for processing when the optical disc controller is selecting a new track (image). This card may be omitted if the objects are stored on the disc in 90 degree increments or if the rotations are less than ±45 degrees.

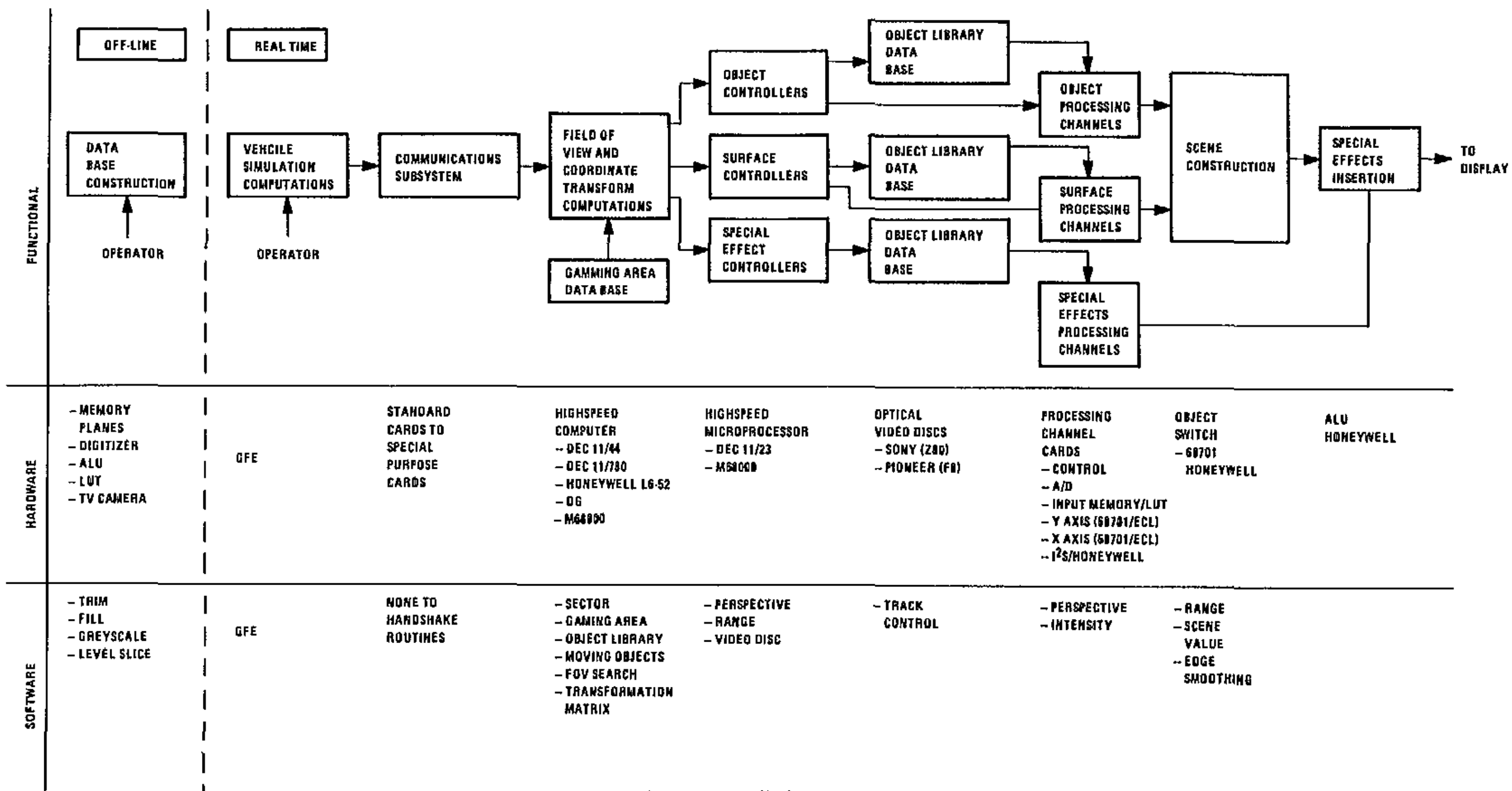

Figure 12. Hardware Software Overview

c. A lookup table modifies the intensity values of images (for range and contrast effects). This operation requires only a delay of several pixels.

d. A warp card transforms the image in the Y axis on a line-by-line basis. The starting point (offset) and magnification factors shift and compress or expand the pixels of each line. This operation delays the flow of pixels by one line.

e. A second identical high-speed read/write X and Y axis memory card accepts and stores the transformed Y data for an odd and even field to form a frame. After the Y axis field is loaded in the Y axis, the X axis data is read out by line, and even and odd fields. This buffer operation requires one video frame.

f. A second warp card identical to Y processes X axis data by shifts and expands or compresses lines. Again, this operation delays the image by approximately one video line.

Figure 13 shows a possible hardware implementation of a processing channel.

CGSI System Configuration

CGSI uses a modular set of building blocks which may be configured to meet specific training and simulation requirements. A system may be configured for a wide FOV visual color display or a narrow FOV sensor display. In addition, CGSI may be added to existing CGI systems as a product improvement to improve realism and fidelity.

CGSI STATUS

The feasibility of CGSI in non-real time has been demonstrated using a video tape showing a nap-of-the-earth helicopter flight. The CGSI demonstration video tape shown at the conference was generated in non-real time. All the algorithms, edges, perspective, dust, smoke, shadows, etc., were tested in non-real time. A real-time CGSI system has been designed for NAVTRAEQUIPCEN/PM TRADE under the contract ***"Extend CGSI Concept."***

Because of the high potential payoff from the development of this hybrid approach, the Navy and the Army have current plans aimed at demonstrating feasibility of this CGSI technology in real-time. CGSI pipeline processors will be interfaced to NAVTRAEQUIPCEN's *Visual Technology Research Simulator* (VTRS) facility for evaluating CGSI concepts and for demonstrating combined multisensor displays. Figure 14 demonstrates the high fidelity and realism of CGSI.

ABOUT THE AUTHORS

Mr. Carl P. Graf has degrees in Psychology, Math, and Engineering. His 20 years of experience at Honeywell in the man-machine interface area include: Apollo and LBM manual controllers, eye tracking, eye switching, passive and active camouflage, maintenance trainers, image processing, multisensor imagery and displays, dual resolution displays, and the generation of high fidelity imagery.

Ms. Dorothy M. Baldwin obtained her M.A. in Physics from Kent State University in 1968 and her B.A. in Physics from Hartwick College in 1965. Her 13 years of professional experience have been in government and academia. Her 5 years at Naval Training Equipment Center includes work on the 360° laser scan display system, the helmet mounted display system, and the annular projection system. Her current assignments include: Principal Investigator on the Multi-Spectral Image Simulation Project and Advanced Sensor Simulation Project.

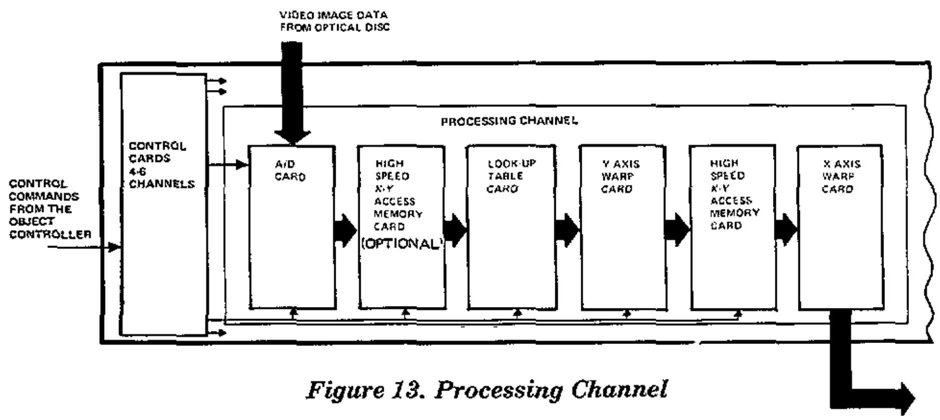

Figure 13. Processing Channel

Figure 14. CGSI