Accuracies of Bomb-Scoring Systems
Based on
Digitized 2- and 3-D TV Images

James L. Rieger, PE/PTBW
Naval Weapons Center, China Lake, California

May 21, 1990

Abstract

Three-dimensional images produced by film or analog television have been used for bomb scoring by triangulation for many years. Use of solid-state imaging devices and digitization of analog camera outputs can improve the accuracy of such measurements, or make accuracy lower or (worst of all) of random accuracy if interpreted incorrectly. This paper examines some of the issues involved, and tabulates the maximum accuracies available for a given system.

1 Introduction

Motion picture film was the original medium of choice for bomb-scoring and miss-distance calculation. Developed film from several cameras is observed one frame at a time, and the position(s) of the object(s) of interest in the frame with respect to one of the corners are recorded, along with the azimuth and elevation angles of the camera pedestal. This data, along with the exact geographical locations (latitude, longitude, and altitude) of each camera allows calculation of object paths and miss distances through trigonometry. Bomb scoring is a more limited application of this technology, since the cameras used are in fixed positions, and need not track the target since the target itself is fixed. In such a system, the horizontal distance (actually an angle) between the target and the impact point as viewed from two or more angles is the only data of interest. Geographical positions need not be known, and the camera positions are expressed in two-dimensional numbers in an arbitrary grid with its origin at the target.

Use of television instead of film has several advantages, not the least of which is that providing such data in “near real time” is a possibility—allowing the test conductor to tell the pilot how to optimize his aim on the next pass. In this paper, we will examine the methods that can be used to make such measurements, and the accuracies that can be expected.

2 Notation and Coördinate System

A bomb-scoring system involves a two-dimensional set of gridwork with a target at the center. Aimed at the target from a ground perspective are two or more cameras, each of which has some indication of the target's location near the center of the horizontal field of view, and the ground at the target near the lower edge of the field. As a bomb falls and contacts the ground within the field of view, the horizontal distance (as a fraction of the full screen width) between the target and the bomb is measured and noted. Distance between the camera and target is noted as D_{r1}, D_{r2} etc., where the number is the camera number; horizontal distance between the target and the bomb is represented as R, with R expressed as a polar or rectangular vector quantity.[1] The angle between the target and impact point for any camera is denoted as 2_{in}, where n is the camera number, with negative values for an impact to the left of the target and positive values to the right.

The units of measure used in this paper are feet and degrees, but the user is at liberty to imagine any units of measurement of angle and distance.

3 Mathematical Basis

There are two arrangements which can be used for bomb scoring as described above. One is based on two or more views that are taken from cameras that are close together and separated horizontally (which would create a 3-D image if viewed properly), and the other more common method is based on two or more views taken from cameras that are widely separated.

Since the grid system for the bombing range is arbitrary, it is assigned such that the y-axis is positive along the flight path. If polar notation is to be used, azimuth is then taken as being in a clockwise direction; for cartesian coördinates, the x-axis is positive to the right of the y-axis. In either case, distance to the target from the camera is known, as is the position of each camera in the system used.

Distance from one horizontal edge of the picture to the other depends upon the distance from the camera—the width of the picture is actually an angle. The horizontal angle received by a camera depends on the focal length of the lens, which is never precisely known, and the extent of the horizontal sweep of the pickup tube (or film, for that matter). Lens focal length and camera position are selected so that the camera can see the entire possible impact area, so these factors vary. Usually cameras are positioned in such a way that they are never along the flight path, just in case a bomb accidentally falls off.

To calibrate each camera system, the angle between the edges of the display must therefore be measured, normally done by measuring the distance between two points at the edges of the picture, with each of those points equidistant from the camera. Actually, any

---

[1] Hence "the bomb impacted 22 feet from the target at an azimuth of 243 degrees", or "the bomb impacted 20 feet to the left and ten feet ahead of the target".

two points will do, but greatest accuracy is afforded by using the edges of the display. This measured distance D_{wn} is then converted to an angle, 2_{wn}, where n is the camera number. In general, the 2_{wn}'s will be different for each camera, even if identical cameras and lenses are used. Knowing the 2_{wn} for each camera and the distance between the target and impact point allows calculation of 2_{vn} by proportionality.[2] A typical range setup for a single camera is shown in Figure 1, below.

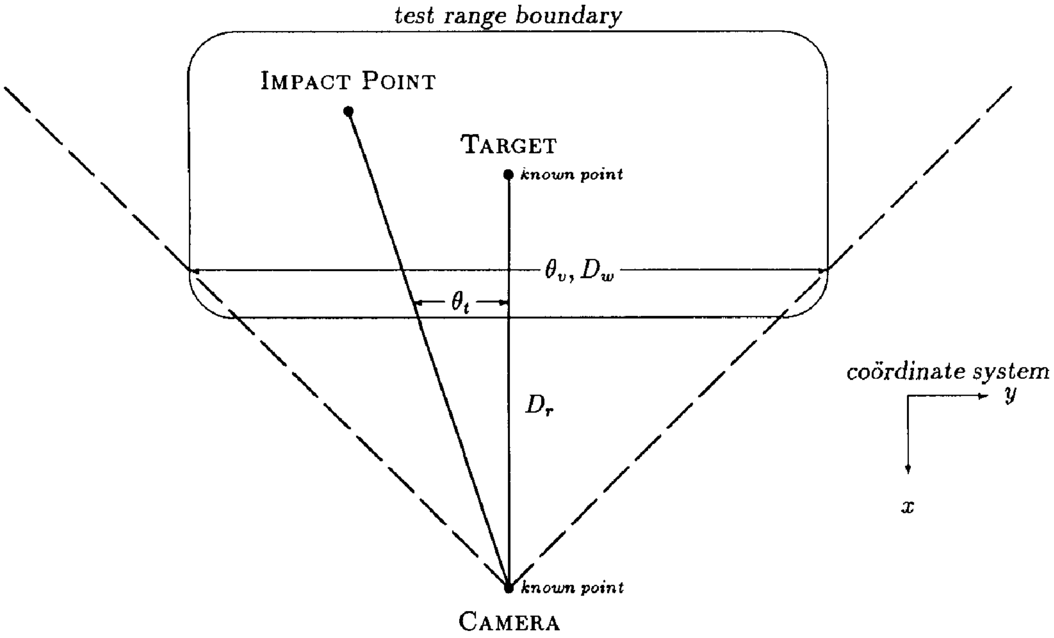

Figure 1: System View, Single Camera Perspective

While a single camera produces a single angle, two such measurements must be made to locate the impact point in space. In practice, more than two cameras are used, with the two values of 2_{tn} which have the largest absolute values used for calculation. Positioning of the cameras is such that the entire range (drop zone) is covered by at least two cameras at all points, and (unless 3-D display is desired), the cameras are as far apart as possible.[3]

To solve for the unknown target position, the angles between the target and impact point for two cameras are used. If more than two cameras are used, the most accurate results can be derived from those two cameras with which the largest angle is produced. In the case of the example shown in Figure 2, this would be the angles produced by cameras one and three.

---

[2] Note that the target need not be at the center of the screen. In general, the edges of the picture are set at the edges of the bombing range, and the target is near, but rarely at, the center of the screen for any camera.

[3] In general practice, the cameras are never distributed over an arc greater than 180° because of the disorientation to a viewer observing the same event from both sides; this is the same arrangement used in television coverage of football, for example.

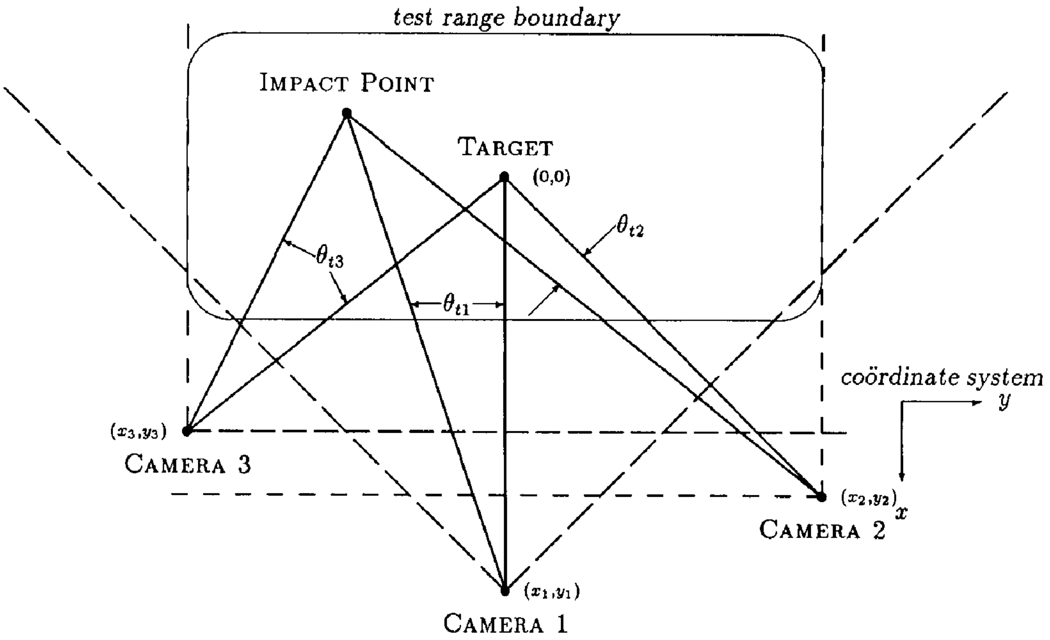

Figure 2: System View, Multi-Camera

As seen in Figure 3, two large triangles can be drawn, both with a line connecting cameras one and three, and with two vertices—one with the target, and one with the impact point.

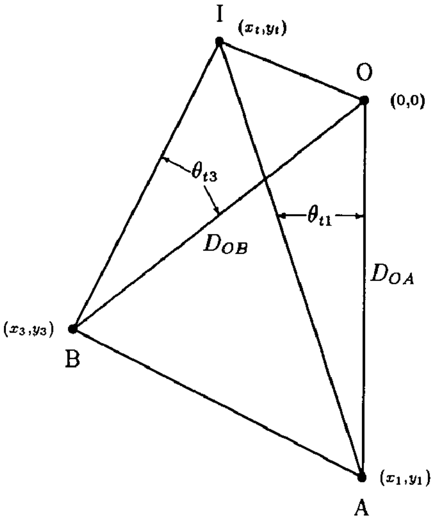

Figure 3: Triangulation for Solution for Impact Point

To determine the lengths of lines OA (called D_{r1}) and OB (called D_{r3}), the coordinates of the camera positions are used:

$$D_{OA} = \sqrt{x_1^2 + y_1^2} \tag{1}$$

and

$$D_{OB} = \sqrt{x_3^2 + y_3^2} \tag{2}$$

Similarly, the distance AB is given by

$$D_{AB} = \sqrt{(x_3 - x_1)^2 + (y_3 - y_1)^2} \tag{3}$$

and the angles of triangle OAB can then be solved for by the law of cosines[4] as

$$\angle A_{OAB} = cos^{-1}\left(\frac{D_{AB}^2 + D_{OA}^2 - D_{OB}^2}{2D_{AB}D_{OA}}\right) \tag{4}$$

$$\angle B_{OAB} = cos^{-1}\left(\frac{D_{AB}^2 + D_{OB}^2 - D_{OA}^2}{2D_{AB}D_{OB}}\right) \tag{5}$$

The angle made by the line OA in the defined coordinate system is

$$\angle OA = tan^{-1}\left(\frac{y_1}{x_1}\right) \tag{6}$$

and the angle may be negative. The angle to at least one camera for each solution must be known.

All these values can be calculated and stored beforehand. When an impact occurs at point I, the triangle IAB is formed. Length AB is known from before, and angles 2_{t1}, and 2_{t3} are determined from the data. In the example shown, both 2_{t1} and 2_{t3} are negative, decreasing the angle at point A and increasing the angle at B due to geometry. Knowing the values of one side and two angles, the law of sines[5] can be used to calculate the length of line AI or BI (only one need be calculated). Finally, knowing length OA, and having determined angle 2_{t1} from data, the law of cosines is again applied to determine line length IO as

$$D_{OI} = D_{OA}^2 + D_{AI}^2 - 2D_{OA}D_{AI}\cos\theta_{t1} \tag{7}$$

where the distance D_{OI} is the magnitude of the miss distance.

---

[4] The law of cosines is $a^2 = b^2 + c^2 - 2bc \cos A$, where a, b, c are sides opposite angles A, B, C.

[5] The law of $\frac{a}{\sin A} = \frac{b}{\sin B} = \frac{c}{\sin C}$ sines:

Then, since D_{OA}, D_{OI}, and 2_{t1} are known, angle AOI can be calculated using the law of sines and the fact that a triangle's interior angles add to 180°:

$$\angle AOI = 180^\circ - \theta_{t1} - \sin^{-1}\left(\frac{D_{OA}}{D_{OI}}\sin\theta_{t1}\right) \tag{8}$$

so the angle between the target and the impact point, 2_{miss}, is the algebraic sum of ∠OA and ∠AOI. The space position of the impact point in cartesian form may then be solved for, using

$$x_t = D_{OI}\cos\theta_{miss} \tag{9}$$

$$y_t = D_{OI}\sin\theta_{miss} \tag{10}$$

which are two equivalent forms of the desired solution. While many of the above steps can be combined, there is no particular advantage in having a computer perform one complex calculation or a number of simpler ones, and the intermediate results obtained may be of some use.

4 Data Transfer

Whether the film is observed on a small rear-projection workstation, or the television picture observed on a monitor, an operator examines the frame which shows the bomb's initial impact and measures and enters the position of the impact point and perhaps the target into a computer (perhaps one punched card is produced for each frame). In electronic analysis, the horizontal line or lines of interest are observed for level changes. Since the objects observed have finite widths, distance to the left side, right side, and center differ, and the exact position of either edge is more readily observed than that of the center, so the left edge is commonly used. Resolution of this edge is imperfect in any system as well, for a number of reasons.

5 Television vs. Film

Several differences exist between film and television, whether the television signal is "live" or on tape. First, the possibility of a "live" television observation of an event is possible, while film produces a picture only after the film is retrieved and taken to the drugstore for developing. Analysis of film requires projection on a screen and measurements taken of the distances between the target and dropped bomb, and transcription of that data; a television picture requires a monitor to allow viewing, but could at least in theory use some electronic reader to perform the desired measurement. Film frame rates can be anything, however, and television rates are pretty much fixed at

59.94 pictures[6] per second for color systems and 60 pictures per second with black-and-white. If a geophone were placed near the target, the moment of ground contact could be made to mark the soundtrack of either medium; a television signal can be electronically analyzed to determine a change in the scene, and the picture automatically tagged. Since television pictures occur once every 16 2/3 millisecond, the likelihood of one picture showing the bomb before impact is slight and the likelihood of two or more showing the bomb in flight is even lower. If this is a consideration, the electronic examination of the picture can take into account only those lines associated with the ground.

Assuming that optics are used which are better than the resolution of the film or television pickup, accuracy of measurements is limited by the film or television medium itself. While film has no theoretical resolution limit, the limits of the television medium are fairly easily categorized.

6 Limitation of Television Signals

While a television display is two-dimensional (height and width), only the resolution in the horizontal direction has any significance in the type of bomb scoring described here. The nature of those limitations, and ways to mitigate them are discussed below.

6.1 Sweep Linearity

A television picture is generated by either a tube surface onto which the scene is focused and read by an electron beam, or an integrated circuit matrix of light-sensitive diodes which are polled by a pointer duplicating the motion of the swept electron beam in the tube. While ideally the horizontal and vertical velocities of the scanning beam in the tube are everywhere uniform, they are not. As a result, even if displayed on a perfect display, objects which were evenly spaced in the original scene will appear unevenly spaced on the display. The nature of this sweep nonlinearity may be described by a simple curve or something more complex, and may vary with the tube's temperature and age, and with interference from things like power supply hum. A solid-state imager's linearity is controlled by the accuracy with which the mosaic was etched (normally high accurate enough to be considered perfect), and by the stability of the oscillator with which the picture information is stepped, use of a crystal-controlled clock can reduce this effect to negligability. This effect can be seen with film as well, and pertains to curling of the film in the recording and playback mechanisms, and elasticity of the film base; the effects are far more random and less predictable than with television signals.

---

[6] The term "picture" refers to a single field of a television signal, from an interlaced or a noninterlaced original. The 2:1 reduction in vertical resolution is of no consequence in this system.

6.1.1 Display and Channel Nonlinearities

Even with a perfectly linear sweep, of course, sweep nonlinearities are produced by most displays,[7] and by the record/reproduce process. Velocity-error compensation devices introduced to stabilize the color subcarrier often exacerbate horizontal linearity problems.

6.1.2 Fiducial Marking

If a pickup device is hit by too much light under some circumstances, the picture generated by the pickup will retain the overload marks thereafter. This is normally considered a disadvantage, but the same mechanism can be used to advantage in bomb-scoring systems and the like, where measurement is of primary importance and aesthetics secondary. Intentionally burning calibration marks into such a pickup will thus produce a ruler with divisions much smaller than one screen width, between which shorter distances can be interpolated with considerable accuracy even in the presence of nonlinearities at any step or steps in the chain. [1] These marks, called fiducials can be placed in vertical and horizontal positions on the screen where their interference with picture content is unlikely. The width and number of the fiducials is governed by the resolution available in the system and the severity of the expected sweep nonlinearities.

6.2 Grayscale Nonlinearities

Grayscale nonlinearities are of no consequence in a system such as that described here. Systems which limit the number of grayscale steps must have the steps close enough to identify the target and impact points, and the step size should not interact with the background in such a way as to cause parts of the picture to flicker by toggling between two nearby steps. This is seldom a problem in practice.

6.3 Bandwidth

The bandwidth of a “standard” NTSC television picture for broadcast transmission is limited to about 4.2 MHZ, although there is no bandwidth limitation for any other use. Since the aspect ratio of a television picture is 4:3, the bandwidth should be about 6 MHZ so that the smallest displayable picture element is a dot rather than a dash. Many closed-circuit instrumentation systems have bandwidths of 10 MHZ or more, and horizontal resolution is consequently ultimately limited by the pickup and display devices. Most videotape recording methods produce bandwidths far lower (on the order of 1.5-3 MHZ). Increasing the number of “lines” in a picture for a given bandwidth actually lowers the resolution in the horizontal direction, so use of such systems is contrary to what is

---

[7] Liquid-crystal displays are essentially linear, but are seldom used.

needed for highest accuracy in bomb scoring. For a tube-type system limited solely by bandwidth, a brightness transition can take place at any point, with a risetime equal to $1/2\pi f$. The "active" display time of a horizontal line is nominally[8] 51.6 μsec, hence in a 4.2 MHZ system, a risetime of 5 x 37.9 nsec takes place over 0.37% of the sweep; this is a fair measure of the maximum resolution attainable for a scanning beam width infinitesimal in size. For a scanning beam with width on the order of the height of a line, such a transition will take 1/640 of the line duration, or 0.15% of the sweep. As a consequence, increasing the system bandwidth beyond about 9 MHZ will not improve resolution unless the horizontal dimension of the beam is decreased.

6.4 Slew Rate

A "real" video system is limited not only by bandwidth, but by the rate at which the scanning beam can change brightness from black to white or white to black.[9] One measure of system performance is "lines of resolution" which is affected both by the system bandwidth and slew-rate limits—such limiting resolution is the number of brightness changes perceptible in the horizontal direction which can just barely be perceived. Operation near this limit presents a problem to a human or a machine, because some threshold needs to be established beyond which a data point may be assumed. This is further complicated by searching for such a variation in a background that may not be entirely "flat".

6.5 Noise

Noise introduced at any point affects picture brightness in such a way as to make some point on the display brighter or darker when scanned at one time than when scanned another time. The nature of most types of noise is gaussian, meaning that the magnitude of the noise (but not its polarity) has an average "expected" value, although this value is exceeded at least some of the time. To avoid the effects of noise, the threshold or change for any point on the picture that will be used to indicate that data is present should be at least three times this average noise value. A human can often read data in the presence of extreme amounts of noise, but to do the same thing with a machine is difficult.

---

[8] NTSC standards trace back to vacuum-tube circuits circa 1939, when pulse durations were timed with one-shots rather than by counting down from the master clock, and thus have fairly wide tolerances.

[9] To further complicate the issue, these two rates can differ.

6.6 Illuminant Changes

The light seen by a camera (be it film or television) is the light reflected by, passing through, or emanating from objects in the scene. Any rational arrangement for bomb scoring will consist, at least principally, of light reflected from most of the scene. If the source of illumination changes in brightness, evenness, or color between pictures, the resulting changes may be interpreted as data when data isn't there. While the changes due to sun angle and cloud cover are gradual, the shadow cast by the aircraft dropping the bomb, or the aircraft itself passing through the scene are not. For these reasons, use of some external marking on the audio track (voice annotation, geophone, etc.) is necessary in automated systems to prevent false triggering.[10] The geophone arrangement has the additional advantage of marking when impact has occurred, thus eliminating from consideration any pictures where the bomb is visible but has not yet made ground contact.[11]

6.7 Pixillation

Television as originally invented involved analog circuits coerced into doing digital things with the minimum number of tubes. While transistorized systems and digital technology are compatible with the analog standards, there are differences. While the brightness voltage as rendered by a tube-type pickup can change anywhere along the line, limited only by bandwidth, slew rate, and scanning dot size, a solid-state pickup consists of a countable number of "pixels", and brightness changes may only take place between them. Assuming (as is usually the case) that the frequency response and slew rate limits can be reduced to insignificance, the minimum resolvable distance between two objects on a horizontal line is the reciprocal of the number of pixels on the line. In a system with, say, 640 pixels per line (which would generate a square pixel in an interlaced NTSC system)[12], hence a change occurring between two pixels will be resolved in I/640 of the sweep and any change that occurs wholly within a pixel will be resolved entirely over twice that distance and observable as a partial change in the pixel within which it occurs, and completely resolved on either side. As a consequence of this, the value of the middle pixel

---

[10] No matter how accurate an automated system is, false triggering will reduce confidence in the system when it reports an impact when none has occurred.

[11] As an interesting aside, geophones used in groups do not make particularly good scoring systems because different underground masses have different sound speeds and echoes occur at points where masses change.

[12] A standard NTSC receiver is incapable of resolving pixels this small and hence gives the appearance of being continuous. A VGA computer has 480 x 640 pixel resolution, for comparison purposes.

can be used to interpolate the location of the change to I/10 pixel or so, depending on system noise. As a result, a square 640 pixels-per-line system has about the same resolution as a 9 MHZ continuous system.

Pixillation can also occur if a continuous source is time sampled for the purpose of processing in such things as time-base correctors or digital recorders. Even if the signal is continuous, comparison to an electronically-generated marker which has finite-sized steps produces much the same effect. Again, interpolating between steps of such crosshairs can increase the resolution available.

Nothing prevents use of systems which have pixels narrower than their height, or use of systems with square pixels that use two or more pickups with their pixels interleaved in the horizontal direction and the necessary logic to raise the resolution accordingly. Such systems are ultimately limited by the bandwidth and slew rate of the circuits between the pickup and the display.

6.8 Pixel Stagger

If the pixillation of the image is done by the encoder or other processing equipment rather than by the pickup itself, the size and location of the pixels are arbitrary. If the number of pixels on a line is fixed by bandwidth or channel capacity considerations, the positional resolution of any object at least two lines tall can be increased by staggering the pixels like bricks, as shown in Figure 4.

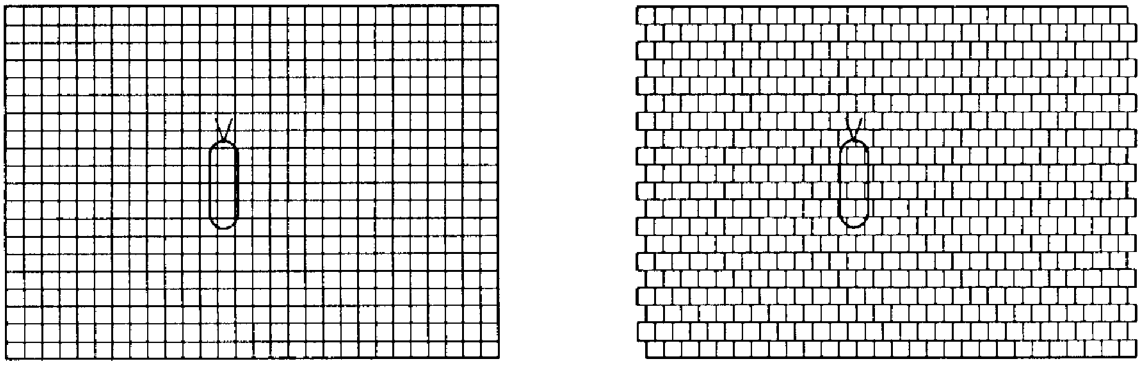

Figure 4: Effect of Pixel Staggering

The alternating lines which have the offset pixels can either contain one less pixel than the lines which start at the “normal” unblanking time, or as shown here the blanking time for staggered lines can be delayed by half a pixel. In either case, the likelihood of a transition occurring between the pixels on one of the lines is doubled, increasing positional resolution without resorting to gray-scale evaluation.

6.9 Double Pixillation

When an image contains elements which repeat at a constant rate which is near the pixel rate of a pixillated source or a continuous source which is pixillated for processing, a moiré pattern is generated whose intensity may be greater than the intensity of objects within the picture. A similar effect is caused by taking a television picture, especially one with lots of activity in the horizontal direction, and feeding that picture as an analog signal into a system which pixillates the incoming signal at a different rate or in a different phase than the original was generated. This double-pixillation effect can be removed (or minimized) by lowpass filtering of the signal between the camera and the processor, or by analog trapping of the offending frequencies.[13] Both of these actions are anathema to the desire for maximum resolution. Consequently, pixillated originals should be treated as such on a 1:1 basis by the digitizing gear.

7 Three-Dimensional Images

Color images differ from three-dimensional black-and-white images in several ways. Among these are

1. Color signals require <u>three</u> views of the same scene, each as seen through a color separation filter; offsets between object positions in the three separations are unintentional. Three-dimensional signals consist of <u>two</u> (usually)[14] views of slightly different scenes; offsets between the objects are the desired information.

2. While the color differences (and thus the "chrominance") of a scene is generally quite small, it doesn't have to be and sometimes isn't. To make any sense, the shapes of objects between the scenes in a three-dimensional image and their offsets should be small or the picture formed won't make sense.

3. Color differences are not perceived by the eye with much acuity, i.e., color "blobs" suffice. Edge differences in three-dimensional images must be as sharp as the images are to convey any information.

As a consequence of this, use of, say, the red and green channels of a color TV channel for sending two views intended to be a 3-D image is suitable for entertainment purposes,[15] but a different scheme is required for instrumentation and metric video.

---

[13] A similar situation occurs when a color signal is fed to any system which expects a black-and-white signal, since the color subcarrier is not in the same location from line to line.

[14] If more than two views are involved, they are still displayed two at a time.

[15] "It came from Outer Space" is available for home viewing in that format.

Because the difference between the two views is “sparse” when the cameras are aligned properly (identical views at infinity), the adaptive/predictive delta code generates one bit per pixel (for “don’t change”) for much of the scene. In the HORACE digital protocol[2], the number of pixels per line can be higher in the difference channel than in the sum channel to increase accuracy. HORACE video decoders not equipped with the 3-D decoder display a single picture made from the sum of the two camera signals, hence even such a decoder can be used for setup of the 3-D system, even though it cannot display it in 3-D.

The decoded 3-D signal consists of two views intended for display on two monitors or on two colors of in RGB monitor for viewing with color separation glasses.[16] For electronic analysis, the separation signal may be used directly, or the two television signals compared.

8 Conclusions

While use of television systems, especially digital systems with limited resolution compared to traditional film methods, may appear to be less accurate and thus not an improvement, but the advantages of rapid—essentially real-time—calculation of miss distances, especially by automated means, appear to outweigh any disadvantages. Digital television interfaces can be used to provide security for such data and pictures without substantial loss. Lower raw material costs and operator time per event make such systems more cost effective, and the possibility of creating “hard copy” still exists. Consequently, television systems have replaced many existing film-based scoring systems already and the trend is likely to continue.

References

[1] Keller, Patrick; Franck, J. B.; Swing, Rictor; and Silberberg, George G.: A Process for Producing Laser-Formed Video Calibration Markers, 15th International Conference on High Speed Photography and Photonics, San Diego, CA, August 1982

[2] Rieger, James and Gattis, Sherri: Draft Standard for Digital Transmission of Television Images, RCC/TCG-209, NWC Technical Publication TP-7025, China Lake, CA, July 1989

---

[16] This two-color mode of display is not limited by bandwidths as is the NTSC-encoded version discussed above.