THE EXPANDING ROLE of SENSOR FUSION

James L. Farrell
Westinghouse Defense & Electronic Systems Ctr.
Baltimore, MD.

ABSTRACT

In marked contrast to the modest data rates involved with nav update, the fusion of image sensor responses requires extraction of information from data streams of several megabytes/sec. The need for this operation arises primarily in situations wherein large numbers of objects (many of which are extraneous) respond to the sensors, and the responses are not readily distinguishable (e.g., no Mode S). Air-to-Ground operations are particularly demanding, due to potentially huge density of sensor responses. An introductory overview is presented here, with no claims of originality.

Usage of multiple sensors for navigation, typified by combining GPS information with Loran or INS data, has been commonplace for years. For tracking multiple objects, however, a restrictive statement of the problem often presupposes that all measurements upon arrival are correctly associated with corresponding estimates. While it enhances manageability to separate various phases of an overall function (e.g., detection, identification, estimation), real-world operation is generally more demanding. A track file is updated after residuals (innovations) are formed from incoming data; but that awaits a decision as to which file should receive which data (the *association problem*). Figure 1 exemplifies a situation wherein some of those decisions are easy (e.g., object ***D*** in the lower right corner of Frame #1 should presumably be associated with object ***d*** in the lower right corner of Frame #2). If all decisions were that simple, that would facilitate the previously mentioned separation of operational phases; typically, however, more than one possible association could be postulated for some objects (such as those in the lower *left* corner of the two frames). Air Traffic Control based on radar skin return (without transponders) is one familiar application which, for high density of objects to be tracked, imposes the need for interdependent decisions.

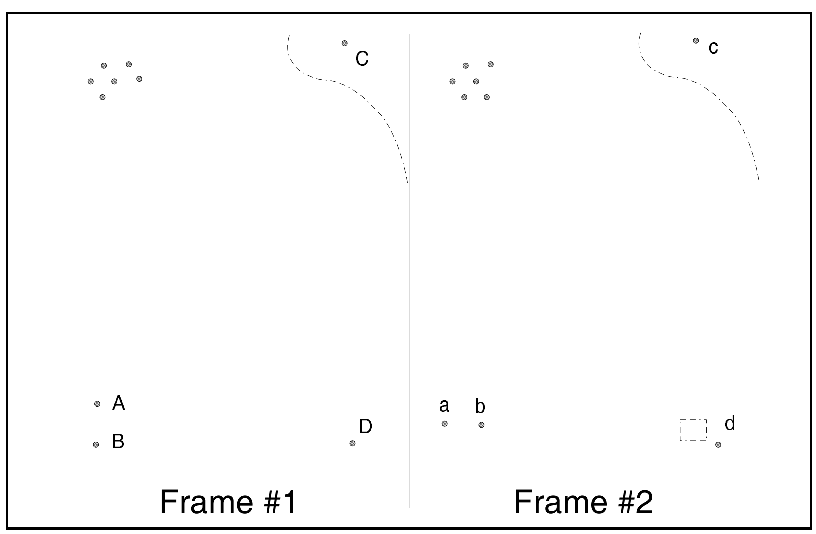


Figure 1 *Multiple object indications on successive frames*

No attempt will be made here to evaluate various approaches to association. For introductory purposes it suffices to note that

> when multiple objects appear in neighboring resolution cells or within one resolution cell (e.g., IR sensor field of view; radar gate & beamwidth), a significant probability of mistaken association exists.
>
> when that occurs, track files can experience major accuracy degradations. For example, a radar echo from one object can be incorrectly matched with a subsequent response from another object, producing a velocity estimate that characterizes *neither* object.

To minimize the problems just identified, multiple sensors can be used together. The desire for synergism generally prompts the usage of devices with different wavelengths (e.g., one with good visibility under adverse conditions, and another with higher resolution). Almost inevitably, the latter translates into some form of optical sensor, which explains the pervasive presence of imaging considerations in much of the pertinent literature. References [1] and [2] typify the kinds of information becoming available.

The issue of conducting the requisite procedures from a moving platform needs to be addressed at the outset. Since these considerations are not restricted to any one class of operations (e.g., whether the tracking sensors and/or the tracked objects are stationary or moving), it might seem upon initial reflection as an unnecessary complication to consider airborne sensors in a brief expository text. That point of view is not followed here, however, for two reasons. First, INS information is used in *motion compensation* to "nip in the bud" the effects of sensor motion on the time history of data received; properly done with high-quality instruments, it is as easy to operate from a high-performance aircraft as from a stationary position. Secondly, motion of the sensor itself is *essential* for some operations [e.g., synthetic aperture radar (SAR) imaging of stationary[1] scenes]. Due to the importance of imaging in this context, further discussion here will concentrate on the Air-to-Surface mode (in which case "tracking" becomes a navigation function for stationary objects).

The concept emerging here, then, involves a pair of airborne sensors in different spectral regions, obtaining different responses from various objects. For a long sequence of image frames with high density of trackable objects, the number of plausible detection/file association *combinations* can easily grow beyond the best computers' capabilities. When many of the objects being illuminated are of no interest, a premium is thus placed on discarding extraneous track files as early as possible. Ref. [3] cites the example of a radar-IR combination looking down on a field; the IR sensor ignores certain extraneous objects (e.g., rocks) while the radar ignores others (e.g., *livestock*) in the field.

---

[1] There is an *inverse* SAR *(ISAR)* operation, wherein the aspect change that enables doppler separation (and therefore imaging) comes from motion of the tracked object, rather than motion of the radar. *ISAR*, which places heavy demands on allocation of sensing resources, is beyond the scope here.

Immediately the prospect just identified draws attention to a central fusion issue: Multiple mistaken associations are far more readily averted when the decision module has direct

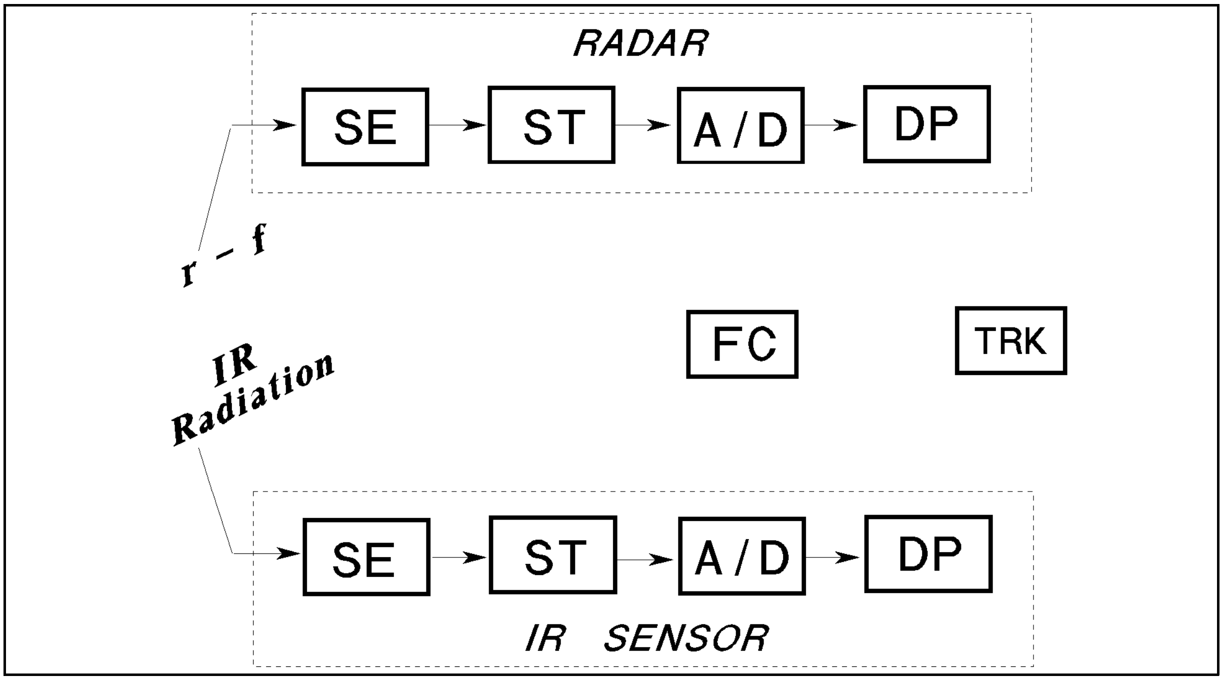


Figure 2 *Elements involved in fusion process*

access to information from both sensors. Figure 2 characterizes a pair of sensors with generic operations symbolized as signal extraction ***SE,*** signal transformation ***ST,*** analog-to-digital conversion ***A/D,*** and data processing ***DP.*** If all those functions are performed before the fusion center ***FC*** accesses the results, there will be comparatively little protection against mistaken associations; the time-shared track filter ***TRK*** can then produce far more data files than necessary, and with substantial degradations as previously discussed. The way to alleviate these problems is to connect the ***FC*** ahead of the association decisions. How far ahead? That could depend on computational capability. As an extreme example imagine the fastest known RISC (reduced instruction set) chips with massively parallel processing, able to form candidate track files for every conceivable combination from long frame sequences. Even with only a dozen frames each containing a dozen objects there would be an astronomical number of candidate track files. Most of these would be highly unlikely (e.g., any combination allowing correspondence of the leftmost object on one frame to the rightmost object on the next); therefore most of the computed track files would contain unrealistic dynamics. Typical algorithmic approaches thus impose "correlation windows" at the outset, limiting possible excursions between frames[4] to regions of limited size such as the dotted square in Fig. 1. The sequences ***Aa, Bb, Cc, Ab,*** and ***Ba,*** for example, would seem plausible; ***Ac, Bc, Ca,*** and ***Cb*** would not.

Obviously the use of correlation windows can eliminate much unnecessary computation. With hundreds of frames and/or objects, however, the number of plausible tracks remaining can still be prohibitive. Another way of disallowing many of those would thus be highly desirable. Herein lies the opportunity: In the Air-to-Ground mode of primary interest here, it is not at all unusual for the vast majority of signal-producing objects to be extraneous. Suppose that objects ***D*** and ***d*** in Fig. 1 produced both radar and IR responses but the dotted window area contained several objects yielding detections in IR only. The fusion center ***FC*** of Fig. 2, if given data access ahead of the track function ***TRK***, could preclude further consideration of objects in the dotted region (deemed nonmetallic); possibility of their association with ***D*** need not be considered.

The example just cited hardly scratches the surface but, for present purposes, it suffices to introduce a basic point: sensor response data should be available to the ***FC***. Given this requirement, system designers familiar with specific sensor configurations will pursue the issue further – at *what* point in the chain of internal sensor operations should the response data be seen by the ***FC***? There are too many individual sensor characteristics for a general prescription, but a few examples can provide some guidelines:

- The radar ***ST*** in Fig. 2 contains various mixer stages wherein the information appears in the form of analog modulation on an *rf* signal. This is too early a stage for transfer of information to the ***FC***.
- The point at which analog-to digital (***A/D***) conversion occurs might also be too early, since some of the intermediate frequency (*i-f)* operations are performed digitally in many modern receivers.
- IR sensors typically contain "on-the-fly" calibration operations; nothing ahead of those stages would be useful to the ***FC***.

Other design decisions along these lines may be more subtle and/or subject to change with state of the art in digital processing and in algorithms for fusion itself.

Extant systems suffer from state-of-the-art limitations (e.g., 1024 x 1024 resolution @ 1 byte/pixel -> 30 Mbyte/sec) but capabilities continue to race ahead. With only moderate parallelism, transputer boards[5] can already achieve 1000:1 speed improvement. As these trends continue, some of the increased capability will be used to process more data, some to enable better fusion algorithms, and some to enhance coordinatizing the sensor outputs. This last item draws attention to another key operation pertinent to each sensor: responses appear naturally in cells formed by locus line intersections (e.g., azimuth-elevation for IR; range-doppler for SAR). In the absence of detailed elevation profile information, those locus lines are formed as intersections with the ground plane. Isorange lines, for example, are intersections of the ground plane with spherical shells centered about the radar antenna; isodops are intersections of that plane with cones whose axes coincide with the velocity vector. Ref. [3] contains mathematical formulations and illustrative plots.

Placement of each sensor response into the proper cell, and establishment of its location relative to that of responses from *another* sensor, is of utmost importance for fusion. Since the cells are formed using a ground plane, the placement accuracy can be compromised by departures from that assumed plane (such as nonuniform height of tracked objects or unknown terrain slope). Unfortunately these effects differ for different sensors. One basic example is illustrated in Fig. 3, wherein two objects are separated by a fixed distance L and the distance from sensor to the closer object is R_1. It takes very little analysis[2] to show that, for a terrain sloped as in Fig. 3(b), distance D is greater than R_2. Thus $D - R_1 > R_2 - R_1$, so that a mapping algorithm based on an assumed level terrain would produce an apparent separation greater than L. The same slope has an opposite effect on inclination angle, however; in that case a mapping algorithm based on an assumed level terrain would produce an apparent *shortening*, because the angle difference $(\beta-\phi_1)$ is *smaller* than $(\phi_2-\phi_1)$.

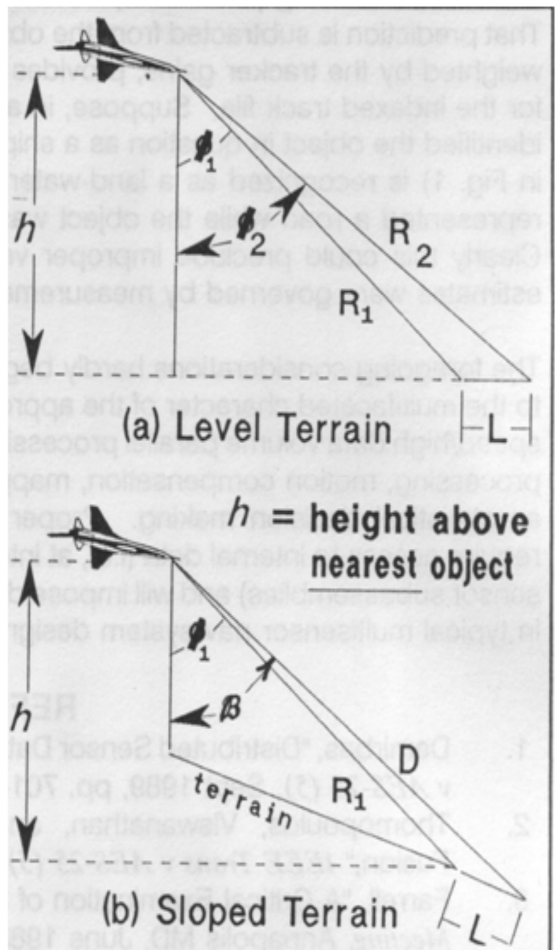


Figure 3 *Effect of terrain slope*

Difficulty in associating sensor responses can sometimes be partially alleviated by using other signal characteristics, but no panacea exists in that effort. Normalized amplitude is sometimes used but, even with the normalization, scintillation can cause radar signal strength to change by tens of dB within milliseconds. A host of image processing techniques (correction, transformation, restoration, enhancement, segmentation, feature extraction, pattern recognition, categorization, registration, photogrammetry, etc.) could be invoked, depending on the specific application.[6] Another procedure used for high-density regions (e.g., top left of Fig. 1) is to establish only a centroid until the cluster disperses. Even after every available measure is taken, there can still be

- more than one possible track file that could be updated with a given detection,
- more than one possible detection that could be used to update a given track file,

as well as new isolated detections and false alarms but, by exploiting fusion techniques just described, designers can vastly reduce their vulnerability to incorrect associations.

---

[2] The line of length R_1 is more nearly collinear with the segment of length L in Fig. 3(b) than in 3(a), thus producing a larger resultant vector sum. Ref. [3] includes comparative IR and SAR locus plots for level and sloped terrain.

If the data available for fusion can be expanded to include topographic information, and the aforementioned categorization of objects carried out, a further interaction is feasible, between processing of sensor responses and the track filter, as follows: Conventionally the time-shared multitarget tracking filter (**TRK** in Fig. 2)[3] receives sensor responses that are time-tagged and, correctly or incorrectly, indexed for association with a specific track file. The *a priori* estimated dynamics for that track file, extrapolated to the time tag for the measurement being processed, will form the basis for predicting the measured value. That prediction is subtracted from the observed value to form the residual; each residual, weighted by the tracker gains, provides estimated adjustments (e.g., position, velocity) for the indexed track file. Suppose, in addition, the aforementioned categorization had identified the object in question as a ship while a nearby boundary (e.g., the curved line in Fig. 1) is recognized as a land-water separation at a harbor; or suppose the curve represented a road while the object was identified as a vehicle constrained to a road. Clearly this could preclude improper vehicle placement that might otherwise occur if estimates were governed by measurement residuals alone.

The foregoing considerations hardly begin to address sensor fusion, but draw attention to the multifaceted character of the approach. Elements are drawn from extremely high-speed/high data volume parallel processing, data base techniques, image sensing/image processing, motion compensation, mapping and coordinatization of detections, as well as statistical decision making. Proper coordination of these functions will generally require access to internal data (i.e., at intermediate stages of cascaded operations *within* sensor subassemblies) and will impose demands significantly beyond those encountered in typical multisensor nav system design.

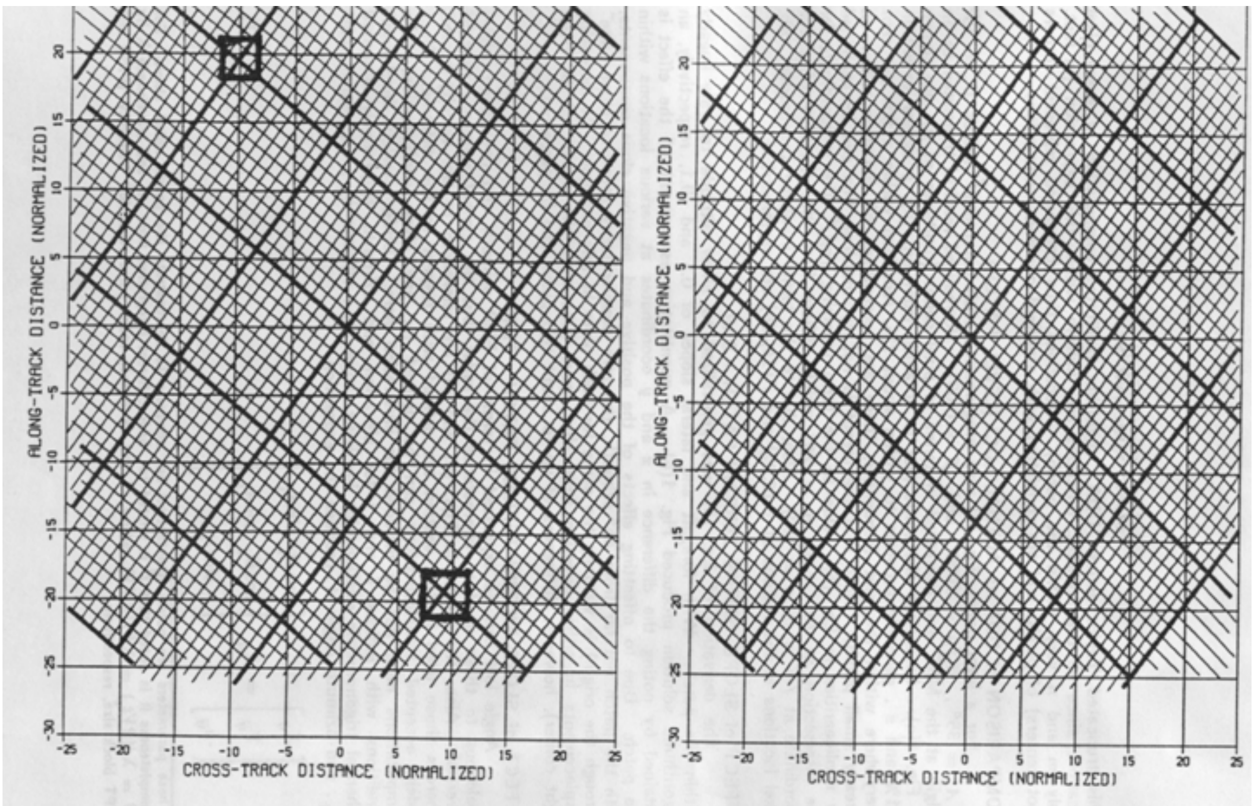
REFERENCES

1. Demirbas, "Distributed Sensor Data Fusion with Binary Decision Trees," *IEEE Trans v AES-25 (5)*, Sept 1989, pp. 701-709.
2. Thomopoulos, Viswanathan, and Bougoulias, "Optimal Distributed Decision Fusion," *IEEE Trans v AES-25 (5)*, Sept 1989, pp. 761-765.
3. Farrell, "A Critical Examination of Sensor Fusion," *Proceedings of ION 44th Annual Meeting,* Annapolis MD, June 1988.
4. Zhou and Kumar, "An Investigation of Correlation Region in Maneuvering Multi-Target Tracking," *AIAA J Guidance v 8 (2)*, Mar-Apr 1985, pp. 249-254.
5. *Transputer Handbook*, INMOS Division, SGS-Thomson Microelectronics, Oct 1989.
6. Gonzalez and Wintz, *Digital Image Processing* (2nd ed), Addison Wesley, 1987.
7. Farrell and Quesinberry, "Track Mechanization Alternatives," *NAECON Proceedings,* Dayton Ohio, 1981.
8. Farrell, Tom, and Nemec, "Air-to-Air Designate/Track with Time-Sharing," *NAECON Proceedings,* Dayton Ohio, 1978.

[3] Little attention is devoted herein to the track filter itself (e.g., Kalman, alpha-beta, alpha-beta-gamma, etc). These are adequately described elsewhere, such as in Ref. [7] for individual targets and in Ref. [8] for the mundane bookkeeping tasks involved in the multitarget case.

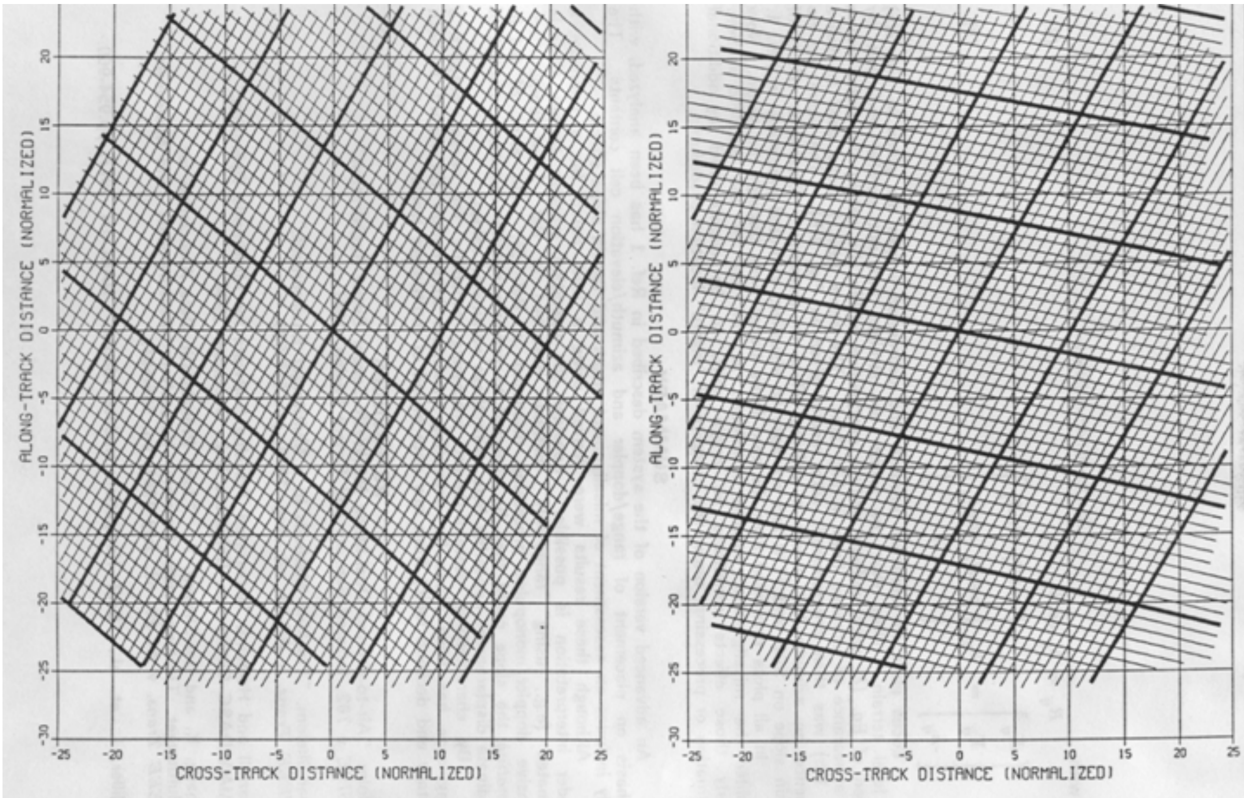
ADDENDUM: EXCERPT FROM Ref. [3]

To illustrate effects of unknown terrain slope, plot figures from Ref. [3] are inserted here. Fore/aft and lateral slopes of 0.1 and -0.1 caused the over-terrain altitude to vary linearly within an image patch, affecting coordinates within the image on the right below (the plot on the left is for level terrain). Due to offsetting effects of the positive and negative slope components, with the patch not far from 45^0 azimuth, the zero-doppler locus (positive slope "line" through the origin) is nearly the same for both plots. At increasing displacements in either direction from that locus (near the upper left or lower right), however, differences of more than one pixel can be seen.



Range-Doppler (SAR) Map Coordinate Loci

The effect of terrain slope on SAR imaging was minor in comparison to what follows here. Azimuth and elevation loci were computed under simplified flight conditions (straight-level, with no angle-of-attack nor sideslip (e.g., velocity vector coincident with airframe roll axis). The azimuth loci, straight but not parallel, were unafffected by the $\pm$ 0.1 slope – elevation, however, showed a dramatic effect. Upon first glance I initially thought a programming error caused such a large change. There wasn't an error; terrain slope combines directly with depression angle in a FLIR image. When the grazing angle is shallow (as it was in the case shown here), considerable distortion of an image can occur with terrain that has either undulations or unknown slope.



Azimuth-Elevation (FLIR) Map Coordinate Loci