# Chapter 2. The technology of verification

## *I. Introduction*

The instruments and techniques used for arms control verification are exactly the same as those used for the gathering of military and political intelligence. The great importance of getting as clear and accurate a picture as possible of an adversary's military capabilities and political intentions has meant that enormous resources of money, time and creative talent have been devoted to the task of creating sensitive, precise, reliable and thorough monitoring devices as well as the processing and analytical techniques needed to interpret the data they produce.

Arms control verification, as the junior partner of military intelligence, has been the mostly inadvertent beneficiary of this remarkable technical effort. Very few of the devices described in this chapter were developed primarily for verification purposes, yet now that they exist they have the potential to create the technological base for significant progress towards genuine disarmament. Whether or not they will realize this potential is another matter and this is discussed in later chapters.

The current military intelligence function of the technologies described here, as well as the importance of some level of secrecy and uncertainty to effective verification, require that many of the most interesting technical details of these devices remain classified. Therefore, any attempt to describe their capabilities and limitations must be preceded by the warning that all estimates are tentative and subject to error. No classified data or information have been used in making these descriptions, and the open literature can be contradictory and misleading since it is often based on hearsay or politically inspired leaks.

The best approach in such a situation is to stick as close as possible to the basic physical principles on which each monitoring technology is based and on generally accepted estimates of the state of the technological art, often obtainable from examination of civilian technology. The key assumption in this approach is that where sufficient motivation exists technical capabilities will generally approach theoretical limits reasonably quickly. There can be no doubt that the desire of states, in particular the two leading nuclear powers, to learn as much as possible about the military capabilities of rival states has provided ample motivation, and that the gap between practical and theoretical performance is now quite narrow for many of the monitoring devices used to gather intelligence. Examples of this narrow gap are to be found in seismological detection, satellite photography and communications monitoring. In other areas such as synthetic aperture radar, thermal infra-red imaging, information processing and artificial intelligence the actual capabilities may still be relatively far from their potential, but progress is clearly rapid and can be expected to continue.

In reading the technical descriptions below it may be useful for the reader to visualize the general process of monitoring as made up of a number of components. First, there is an appropriate instrument (a satellite camera, a seismometer, a human inspector); second, there is an appropriate target (a deployed missile, an underground nuclear explosion, an inventory of plutonium); third, there is a means of processing the data (photo interpretation, seismic data analysis, statistical analysis); fourth, there is a set of limitations to accuracy or transparency (clouds or atmospheric turbulence, high seismic noise levels, flow measurement uncertainties or bookkeeping errors); and fifth, there exists a set of evasion or deception techniques capable of spoofing the instrument (camouflage, decoupling, record falsification). The brief descriptions that follow do not allow for detailed examinations of each of these features for every technology, but the interested reader can explore any of them in more detail using the references, which provide a good sample of the important technical literature in each area.

## *II. Visible light photography*

Certainly the most significant technological development in the field of arms control verification has been the photographic reconnaissance satellite. The potential for using satellites to observe the activities of other states was recognized from the earliest days of the effort to launch artificial Earth satellites; it was a natural extension of the already commonplace use of aerial reconnaissance to photograph enemy teritory in wartime. By the early 1950s, well before the capability existed to put objects into orbit, the potential for peace-time aerial and ultimately space reconnaissance over the Soviet Union was being evaluated at the highest levels in the Truman Administration.[1] Current US photographic satellites are direct descendents of the U-2 aircraft and Discoverer satellite programmes of the 1950s.

By 1961 "The ability to carry out satellite observation of large areas of the Soviet Union with sufficient photographic resolution to spot missile silos was available..."[2] The ensuing 24 years have seen a steady and substantial improvement in the technical capabilities of photographic satellites by both the USA and the Soviet Union, as well as several other states. Photography from space has proven useful for many purposes besides military intelligence and arms control, and a number of states have launched satellites for purposes of weather prediction, resource mapping and ocean surveillance.

The key requirements for useful satellite photography are the same as those for good photography on Earth, that is, good light and object contrast, clear air, precise and stable camera optics, and high-quality, high-resolution image recording, whether on film or directly to electrical signals for electronic processing.

Photographic reconnaissance satellites are placed in orbits which bring them as close to the Earth's surface as is consistent with the desired lifetime in orbit. The Earth's atmosphere grows less dense at high altitudes, decreasing roughly by a factor of one-half for each 5 kilometres above the surface.[3] For example, at an altitude of 20 km the atmosphere already has only about one-sixteenth of its density at the surface. However, because of the very high speed of a satellite in orbit (about 7.5 km per second) even this small amount of atmosphere would be sufficient to heat a normal satellite to incandescence. In fact most photo-reconnaissance satellites have been put into orbits in which their point of *closest* approach to the Earth's surface (the 'perigee' of the orbit) is at least 130–140 km.[4]

It is at or near the perigee of the orbit that photographs are taken, since the ground detail (or target detail) of the image is better if the camera is closer to the region being photographed. Even at 150 km altitude there is sufficient atmospheric drag on the satellite to cause it to lose energy rather rapidly and begin to fall towards the Earth. This effect can be reduced by giving the satellite an elliptical orbit which takes it well outside the atmosphere (say to maximum heights—'apogees'—of 300–400 km) when it is not taking pictures. A mission can also be extended by giving the satellite a booster engine which can compensate for the energy losses caused by atmospheric drag. Figure 1 shows the effects of such a booster on the orbit of a photographic satellite.[5]

When satellites are referred to as 'space vehicles' there is a tendency to visualize them as being far away from the Earth. But on the scale of the Earth itself a photographic satellite at an altitude of 200 km is in fact very close to the Earth's surface. Figure 2 illustrates the relationship of a satellite at this altitude to the surface and shows the width of a strip (2 750 km) which is within the line-of-sight of the satellite. The actual width of such a strip is in fact larger (3 200 km) because of the bending of light as it passes through the variable density of the atmosphere. It would of course be foolish to attempt to photograph this entire strip. Not only are the edges about 10 times as far away as the centre, but the light from the edges must pass through much more atmosphere, suffering much greater absorption, scattering and distortion than the light from directly below. A good example of these effects can be seen by observing the very different appearance of the Sun or Moon as it is rising or setting from that when it is nearly overhead.

The actual strip photographed by such a satellite is more likely to have a

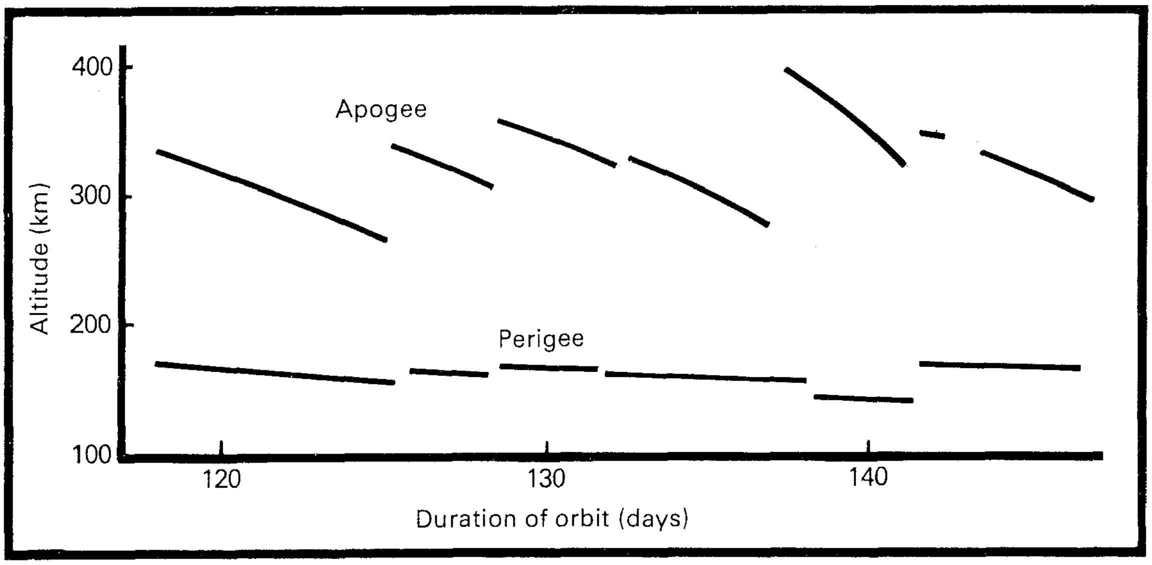


**Figure 1. Effects of air resistance on satellite orbit**
The graph shows the variations in apogee (upper curves) and perigee (lower curves) heights (in km) during the flight of Cosmos 1097. Note the decrease in each parameter as air resistance causes the satellite to lose energy and the sharp increases which result from firings of the booster engine.

*Source*: Jasani, B. (ed.), *Outer Space—A New Dimension of the Arms Race* (Taylor & Francis, London, 1982), p. 142 [a SIPRI book].

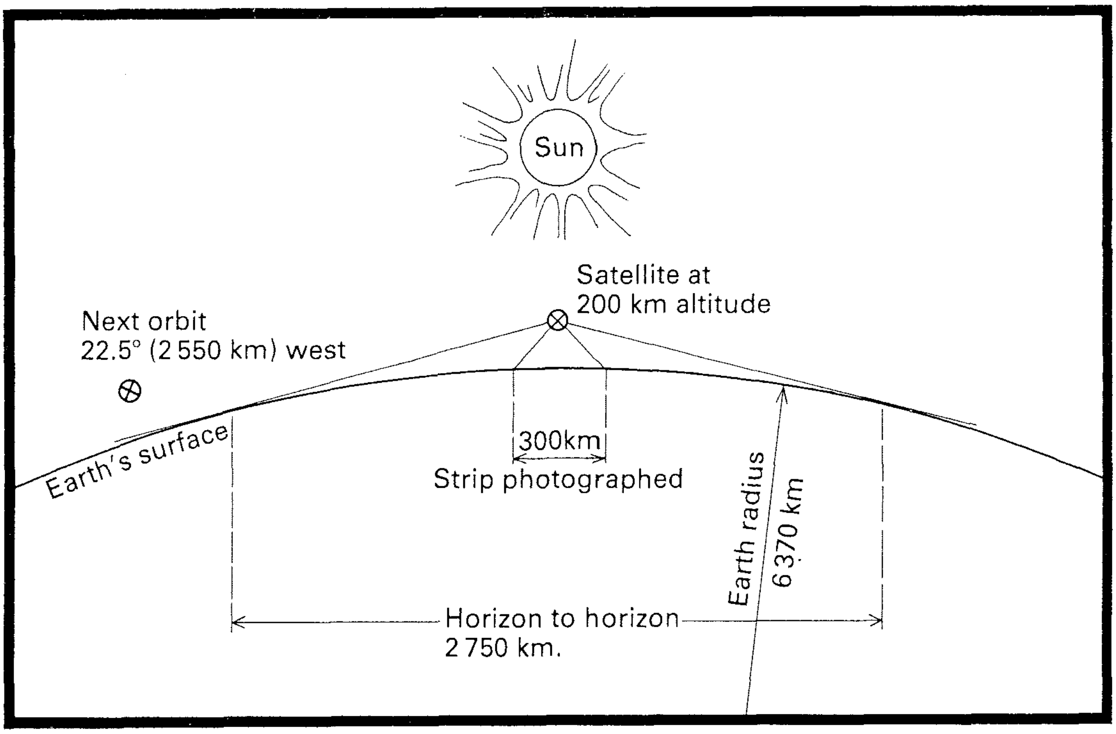


**Figure 2. The spatial relationship between a photo-reconnaissance satellite and the Earth's surface**
The diagram shows the relationship between the height of the satellite, the Earth's radius and the width of a typical strip photograph. The Sun is pictured at local noon as the photographs are being taken.

width of a few hundred kilometres at most, and considerably less if very high resolution is desired. For example, a camera with the high resolution capability of the Space Telescope (see below) at a height of 200 km could photograph an area only 800 m wide on a 23 cm wide piece of film. Such 'close-look' pictures are only taken when there is some reason to believe that they might produce important information. To attempt to survey vast areas at such high resolution is clearly impractical.

Another important technique is to produce overlapping photographs of the same area from different angles. This allows the creation of stereoscopic (three-dimensional) images, which can often be extremely helpful in interpretation.[6]

As it takes its pictures the satellite is moving with a velocity of roughly 7.5 km/s relative to the ground. This means that the camera must be designed to focus on objects for at least as long as the shutter stays open. If the exposure time is assumed to be about 0.1 second, then the satellite will move a distance of 750 metres relative to an object on the ground during the exposure. This means that the camera must rotate through an angle of about ¼ degree (15 minutes) in order to stay pointed at the object. This same relative motion could also be achieved by moving the film during the exposure or by the use of rotating elements inside the camera itself. Even if the photograph is somewhat blurred by motion effects it can be improved by image restoration techniques as long as the elements of interest on the target are not smeared together (see below).

After taking its strip photograph the satellite proceeds on its orbit while the Earth rotates from west to east under it. The polar orbits used by most photographic satellites are very nearly stationary in space. In fact, if the orbit is designed carefully it can be made 'Sun-synchronous', which means that the satellite always passes over the light side of the Earth at a given time of day. The time is picked to obtain the best combination of light and shadow length to produce good definition of objects in the photograph. This implies that photographs taken at low latitudes should be taken either in the morning or in the afternoon, while high-latitude pictures are taken near local noon.[7]

Just as local noon moves westwards with the rotation of the Earth, so will the ground track of the satellite's orbit. If the orbit has a period of 90 minutes, each time the satellite reaches its perigee it will be over a point 22.5 degrees west of the previous point. For example, if the first picture strip was taken over Kiev or New York, then the next would be over Frankfurt or Kansas City. If the period were exactly 90 minutes, then after every 16 orbits the satellite would repeat the same pattern of observations. Since this would leave large areas unphotographed, it is generally desirable to have the period differ by some small amount from 90 minutes. In this way the satellite can be made to photograph adjacent strips and, over a period of several days, achieve virtually total coverage of any desired area.

The process can be speeded up if the satellite camera is capable of photographing a wider strip. This can be accomplished by having it pass at a higher altitude, as long as the optical properties of the camera are sufficient to provide adequate resolution at such a distance. The observed gradual increase of perigee heights in both US and Soviet photo-reconnaissance satellites is good evidence of the improvements that have been made in these optical properties. For example, whereas early US 'close-look' satellites had perigees of 140–150 km, the current KH-11 (Keyhole) satellites combine both close-look and area-survey (wide-angle) cameras in the same satellite, whose perigee is now typically at or above 250 km.[8] It is interesting to note that the perigees of even the earliest Soviet photo-reconnaissance satellites were, with few exceptions, very close to 200 km, but that these began to come down to around 175 km for close-look satellites in the early 1970s. The 35 Soviet photo-reconnaissance satellites launched during 1982 had perigees ranging from 170 km all the way to 358 km. These higher altitudes should permit longer orbital lifetimes, but it still seems to be Soviet practice to bring down satellites after two weeks to a month in orbit. This may indicate either a preference or the necessity for carrying and processing smaller quantities of film that is typical for US satellites.

In the early days of satellite photography it was necessary to return exposed film capsules to Earth for developing and processing. More modern photographic satellites develop the film on-board and use optical-electronic scanning devices to convert the image to a digital code and transmit it back to Earth. Image processing can then be done directly on this coded information. The newest satellites, for example the KH-11, reportedly possess the capability for so-called 'real-time' photography and image processing. Images are coded and transmitted instantly to Earth via a geosynchronous relay satellite, enabling photo interpreters and intelligence analysts to monitor crisis situations as they develop.[9]

## Camera optics for satellite photography

The next major consideration in achieving high-resolution pictures is that of the camera optics. These are illustrated in very simplified form in figure 3. The essential element in any satellite camera is a focusing mirror which reflects rays of light coming from an object on the ground to create an image of that object at a focus near the mirror.[10] An example of the truly remarkable quality now achievable in such mirrors is the one being installed in the US Space Telescope scheduled to be launched into orbit in 1985 aboard the space shuttle (see figure 4).[11] There is no reason to doubt that the optical components used in military spacecraft are at least as carefully designed and crafted as this example.

Figure 3 illustrates in a highly schematic way the basic parameters for evaluating the optical properties of a focusing mirror. A distant object of length $L$ reflects light towards the mirror. If the object is hundreds of kilometres away the light reaching the mirror can be described as a bundle of rays parallel

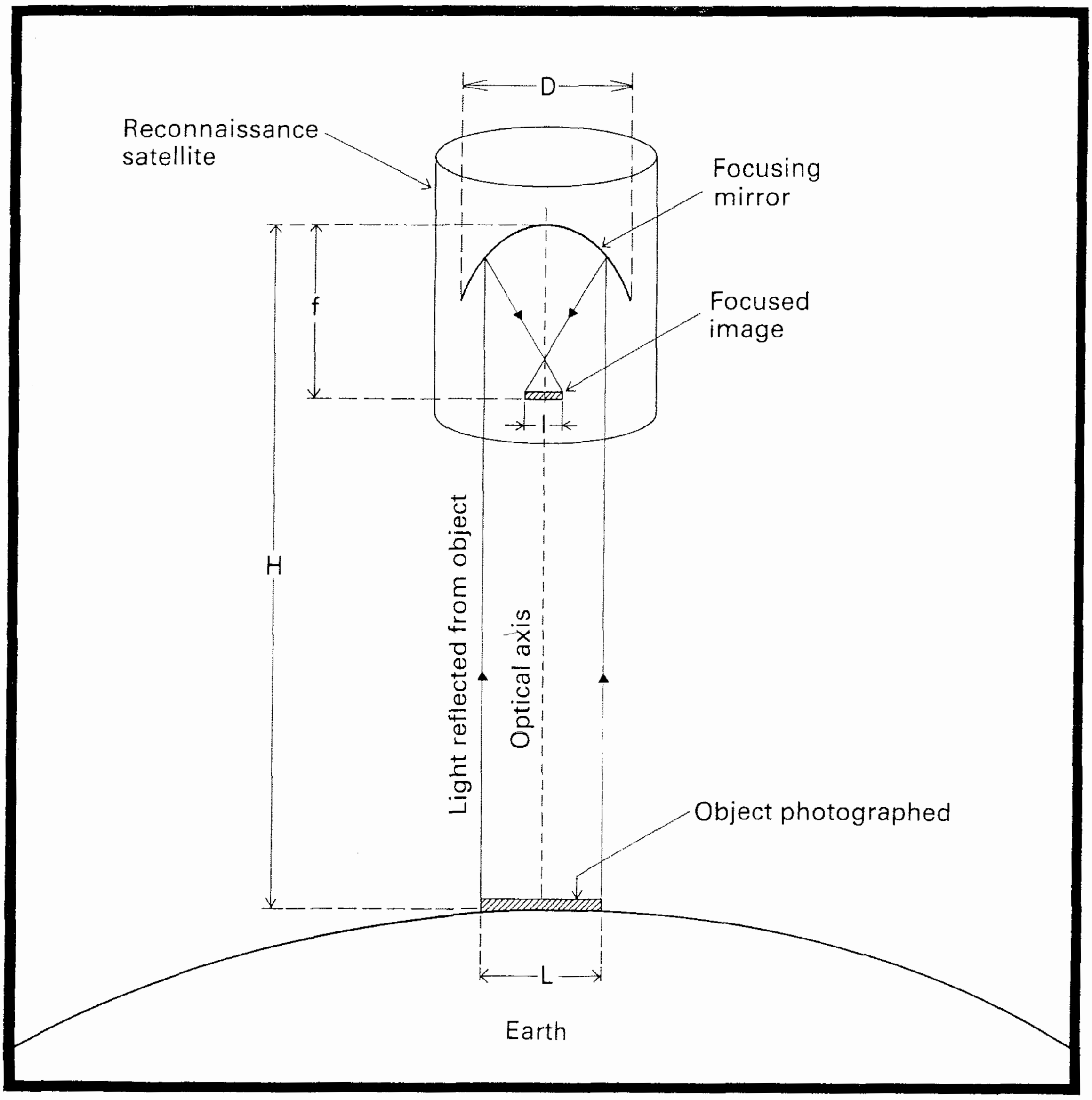


**Figure 3. Camera optics for satellite photography**
A high-resolution satellite camera at altitude *H* utilizes a large diameter (*D*) mirror which produces at a distance *f* from the mirror a real image of an object on the ground. A characteristic dimension of the object is labelled *L* and the corresponding image size is l. Note that the camera is shown pointing straight down for ease of representation. Actual satellite cameras can be oriented at oblique angles if necessary. (The photograph in figure 6 was taken at an oblique angle.)

to the axis of the mirror. This bundle of rays is reflected and brought to a focus at a distance from the mirror known as the focal length (*f*). By changing the shape of the mirror and by introducing other mirrors into the path of the beam the focal length can be made quite long. For example the Space Telescope mirror has been ground to a concave hyperboloid shape, and it will be combined with another much smaller convex mirror (see figure 5) to produce a focal length of 57.6 m in a telescope whose overall length is only 12.8 m. This technique of packing a long focal length into a much shorter distance is called 'folding' the optics.

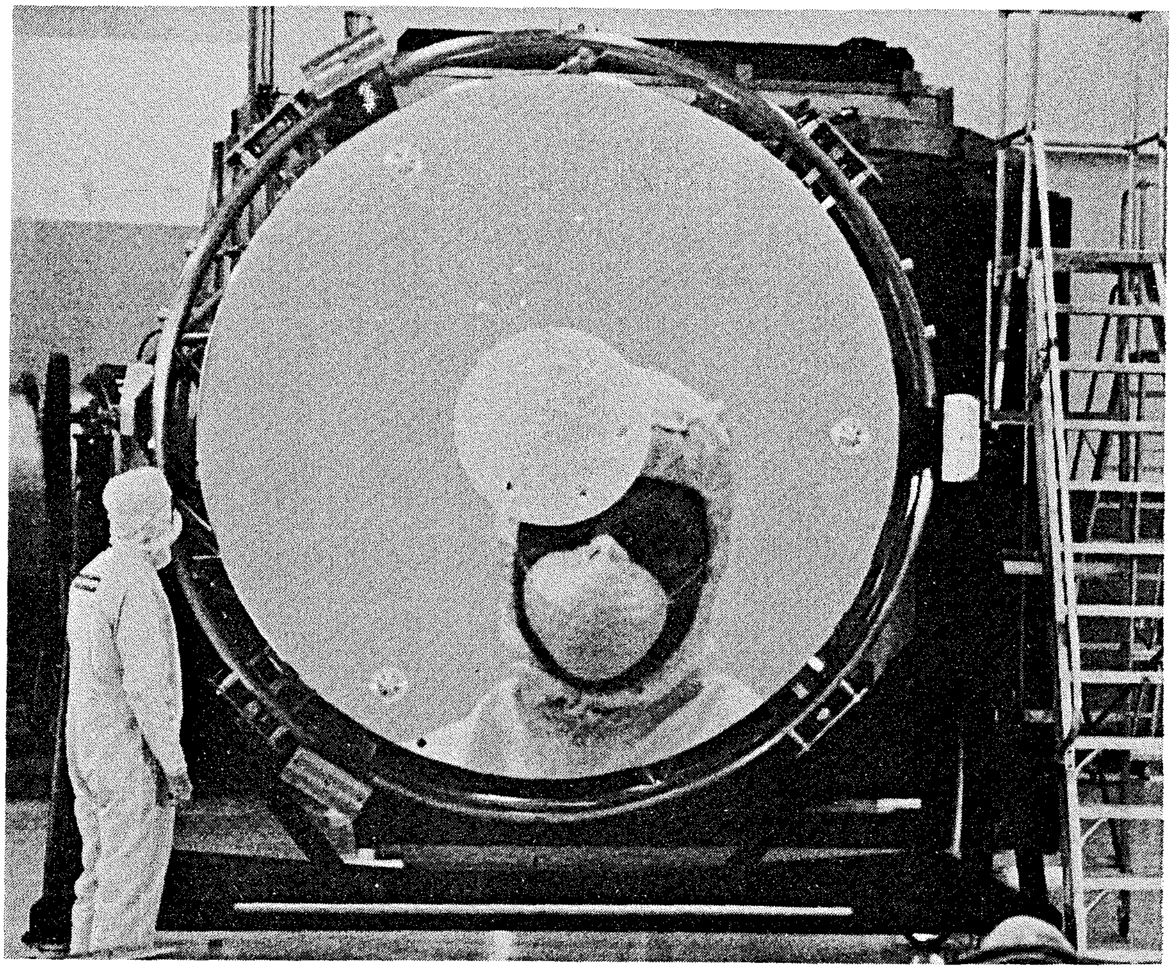


**Figure 4. Primary mirror for the US Space Telescope**
The primary mirror for the Space Telescope was photographed at the Wilton, CT, plant of the Perkin-Elmer Corporation just after its front surface had been coated with a reflective film of aluminium 0.076 μm thick, followed by a protective layer of magnesium fluoride 0.025 μm thick. The mirror, which is made of fused silica glass with an extremely low coefficient of thermal expansion, is 2.4 m in diameter and weighs about 818 kg. It consists of a lightweight cellular core approximately 25.4 cm thick sandwiched between two endplates, each about 2.5 cm thick. Some 91 kg of material were removed from the front plate in the course of the 28 months of grinding and polishing required to give the surface its proper figure, which is that of a concave hyperboloid. The masked man seen enlarged in reflection is standing next to the photographer some 18.3 m from the mirror. Another man, also wearing a mask and a special suit to maintain the cleanliness of the mirror's surface, is at the left. A metal plate temporarily covers the hole in the centre of the mirror through which light from the telescope's secondary mirror will pass.

*Source*: Photo courtesy of the Perkin-Elmer Corp., Norwalk, CT, USA.

Even when the optical path is folded, there is a simple proportional relationship connecting the sizes and positions of the object and image. If the object has length $L$ and is a distance $H$ from the mirror, and if focal length is $f$ and the image length $l$, then the relationship is as follows:

$$\frac{l}{L}=\frac{f}{H} \tag{1}$$

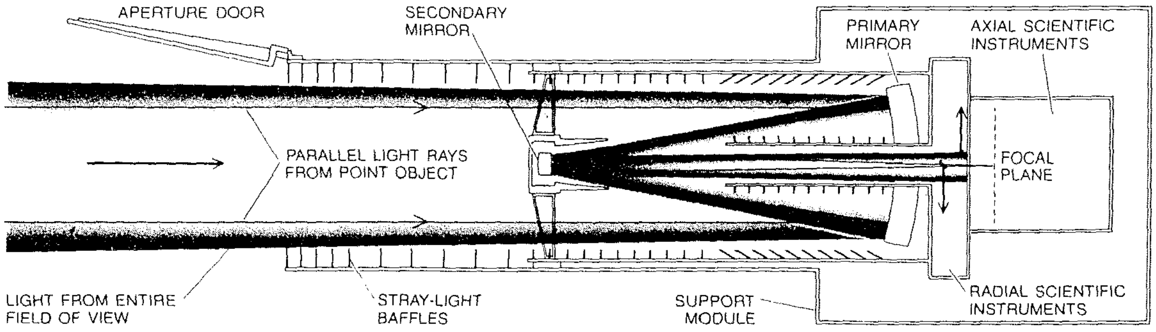


**Figure 5. The optical path in the Space Telescope**
The optical path in the Space Telescope is said to be folded: light from the concave primary mirror is reflected from the convex secondary mirror and passes through a hole in the centre of the primary before coming to a focus at the image plane in the instrument section several feet behind the primary. Technically the telescope is described as a Ritchey-Chrétien type of Cassegrain optical system.

Some representative dimensions are the diameter of the primary mirror (2.4 m), the diameter of the secondary mirror (0.3 m), and the distance behind the primary mirror of the focal plane (1.52 m).[12]

*Source*: Bahcall J. N. and Spitzer, L., Jr, 'The Space Telescope', *Scientific American*, Vol. 247, No. 1, July 1982, p. 39. Copyright 1982 by Scientific American, Inc. All rights reserved.

The ratio $l/L$ is called the magnification, and the formula shows that the longer the focal length the greater the magnification of an object at a given distance. Since the focal length is always much smaller than the altitude of the camera, the 'magnification' is really a small fractional number, and the image is a tiny replica of the object.

A simple application of the above formula would be to imagine taking a picture with a typical personal camera from a satellite 200 km above the Earth. Such a camera has a focal length of about 5 cm which implies a magnification of $0.05/2 \times 10^5$ or 1 : 4 000 000. In order to appear 1 mm long on the film a feature on the Earth's surface would have to be 4 km long. But this same 1 millimetre of film when exposed in the 57.6 metre focal length Space Telescope would record an object only 3.5 m long, roughly the length of an average motor car.

If the capabilities of the film are now taken into account it is possible to calculate how much detail could be recorded within this millimetre of film. The resolving power of photographic film is usually expressed in terms of lines per millimetre, that is, the number of distinguishable parallel line pairs that can be squeezed into 1 millimetre of film. Typical resolutions for commercial film are around 100 lines/mm, but for films used in military surveillance activities resolutions of up to 900 lines/mm have been reported.[13]

Using such a film in the Space Telescope would result in a resolution on the ground of one nine-hundredth of 3.5 metres, or 4 millimetres. Such a photograph would certainly enable one to read the licence number of a motor car from an altitude of 200 km—indeed, one could probably recognize the driver!

It may be calculations such as these that lead some writers on verification to assert that satellite cameras can read car licence numbers.[14] But in practice the use of such high-resolution film is almost never warranted. In the first place high resolution demands excessively long exposure times or very good lighting, neither of which may be available in satellite photography. Second, there are a group of other resolution-degrading effects which render such high-resolution film superfluous. These effects are vibrations and instabilities in the camera itself, diffraction effects, and the presence of turbulence and density variations in the atmosphere, even on the clearest of days.

### *Vibrations and instabilities*

Even though the satellite is in almost empty space and is therefore not buffeted by winds or air drag, it still contains moving parts such as film drives, pointing motors, rotating or oscillating mirrors, and so on. It also passes periodically in and out of direct sunlight, which means that its temperature will fluctuate. These effects produce vibrations and distortions which must be stablilized to a very high degree. That this is feasible is shown again by the capabilities of the Space Telescope which can hold its optical axis steady to within 0.01 arc seconds for as long as 10 hours.[15] A deviation of 0.01 arc seconds at 200 km altitude corresponds to a pointing error on the ground of 1 cm. This can be taken as a reasonable estimate of the optical stability of a sophisticated reconnaissance satellite.

### *Diffraction*

This phenomenon results from the fact that the camera mirror has a finite diameter and can therefore capture only a fraction of the light reflected by the object. The result of this limitation is that the image of a geometrical point (that is, a point with diameter equal to zero) on the ground is spread out into a spot on the film whose diameter gets larger for smaller diameters of the light-gathering mirror. The angular width of the diffracted light beam is given by

$$\Delta = \frac{w}{D} \tag{2}$$

where $w$ is the wavelength of the light being focused, and $D$ is the diameter of the telescope mirror (see figure 3). The diameter of the spot on the film is then computed by multiplying $\Delta$ by the focal length

$$d = \Delta f \tag{3}$$

Wavelengths of visible light vary from 400 nanometres (nm) for violet to 700 nm for red, with the brightest part of the spectrum in the yellow-green at about 500 nm. If this last value is taken for $w$, $D$ is taken to be 2.4 m and $f$ is 57.6 m, it can be shown that a pure geometrical point on the ground will be recorded as a spot on the film with diameter 0.013 mm. But the previous formula shows that 0.013 mm on the film corresponds to a distance of 4.6 cm on the ground. Therefore two point sources of light separated by only 4.6 cm on the ground would produce heavily overlapping spots on the film and be indistinguishable as individual sources. This is the so-called 'diffraction limit' on ground resolution. One authoritative forecast of technological developments predicts that by 1990 available telescope diameters will be 3 m, with 3.5 m a possibility.[16] Such diameters would reduce the diffraction limit on resolution from a height of 200 km to 3.3 cm, or possibly 2.8 cm.

### *Atmospheric turbulence*

Light on its way from an object on the ground to a camera in space must pass through air whose density varies from place to place and fluctuates in time. There is first of all the overall variation of density with altitude which causes light rays to bend. Then there are local and essentially random fluctuations caused by winds and temperature variations. These latter density variations cause slight random bending of the light rays as they pass through the atmosphere, and the result at the camera is an image which tends to wander and flicker over a small region of the focal plane. This effect is the precise analogue of the 'twinkling' of stars on a clear night.[17] There is no easy method of estimating the effect on resolution of this twinkling, but various attempts have produced values between 5 and 10 cm.

Taking together the uncertainties from pointing error, diffraction and atmospheric turbulence it can be estimated that if the Space Telescope were directed at the Earth's surface from an altitude of 200 km (there is, of course, no intention of actually doing this; the Space Telescope is designed for astronomical research and will be pointed *away* from the Earth) a ground resolution of something like 10–15 cm could be achieved on clear, cloudless days. (Some techniques which might enable these limits to be improved by manipulating the developed photographic image or by computerized operations on electronic data from the focal plane sensors are discussed in section V on image processing, pp. 51–54.) There is good reason to assume that this also gives a reasonable estimate of the capabilities of existing military reconnaissance satellites. For example, the Big Bird satellite is reported to be about 14 m long and 3 m in diameter.[18] This compares quite well with the 12.8 m length and 4.3 m diameter of the Space Telescope.[19] It would also not be surprising to learn that much of the technical know-how needed to construct and operate the Space Telescope was first developed in the military space programme.

In conclusion, while it may not be possible to read motor car licence plates from an altitude of 200 km it is probably possible to distinguish different makes of car.[20] Or as William Colby, former Director of the US Central Intelligence Agency, has put it

> You can see the tanks, you see the artillery, but you may not quite see the insignia on the fellow's uniform.[21]

A more systematic assessment of the capabilities of photographic satellites can be made by referring to table 2.[22] For example, from the entries for 'Missile sites (SSM/SAM)' it can be seen that the resolution required for 'detection' of such a site is 3 m (Soviet ICBM silos are 5–6 m in diameter),

**Table 2. Resolution (in metres) required for interpretation tasks**

| Target | Detection[a] | General identification[b] | Precise identification[c] | Description[d] | Analysis |
|---|---|---|---|---|---|
| Bridges | 6 | 4.6 | 1.5 | 0.9 | 0.3 |
| Communications | | | | | |
| Radar | 3 | 0.9 | 0.3 | 0.15 | 0.04 |
| Radio | 3 | 1.5 | 0.3 | 0.15 | 0.15 |
| Supply dump | 1.5 | 0.6 | 0.3 | 0.03 | 0.03 |
| Troop units | 6 | 2 | 1.2 | 0.3 | 0.08 |
| Airfield facilities | 6 | 4.6 | 3 | 0.3 | 0.15 |
| Rockets and artillery | 0.9 | 0.6 | 0.15 | 0.05 | 0.01 |
| Aircraft | 4.6 | 1.5 | 0.9 | 0.15 | 0.03 |
| Command and control headquarters | 3 | 1.5 | 0.9 | 0.15 | 0.03 |
| Missile sites (SSM/SAM)[e] | 3 | 1.5 | 0.6 | 0.3 | 0.08 |
| Surface ships | 7.6 | 4.6 | 0.6 | 0.3 | 0.08 |
| Nuclear weapon components | 2.4 | 1.5 | 0.3 | 0.03 | 0.01 |
| Vehicles | 1.5 | 0.6 | 0.3 | 0.05 | 0.03 |
| Land minefields | 9 | 6 | 0.9 | 0.03 | – |
| Ports and harbours | 30.5 | 15 | 6 | 3 | 0.3 |
| Coasts and landing beaches | 30.5 | 4.6 | 3 | 1.5 | 0.08 |
| Railway yards and shops | 30.5 | 15 | 6 | 1.5 | 0.6 |
| Roads | 9 | 6 | 1.8 | 0.6 | 0.15 |
| Urban areas | 61 | 30.5 | 3 | 3 | 0.3 |
| Terrain | – | 91 | 4.6 | 1.5 | 0.15 |
| Surfaced submarines | 30.5 | 6 | 1.5 | 0.9 | 0.03 |

[a] Requires location of a class of units, object or activity of military interest.

[b] Requires determination of general target type.

[c] Requires discrimination within target types of known types.

[d] Requires size/dimension, configuration/layout, components construction, count of equipment, etc.

[e] SSM and SAM refer to surface-to-surface missiles (i.e., intercontinental or intermediate range missiles) and surface-to-air (i.e., anti-aircraft) missiles respectively.

*Source: Reconnaissance Handy Book* (McDonnell Douglas Corp., USA, p. 125.

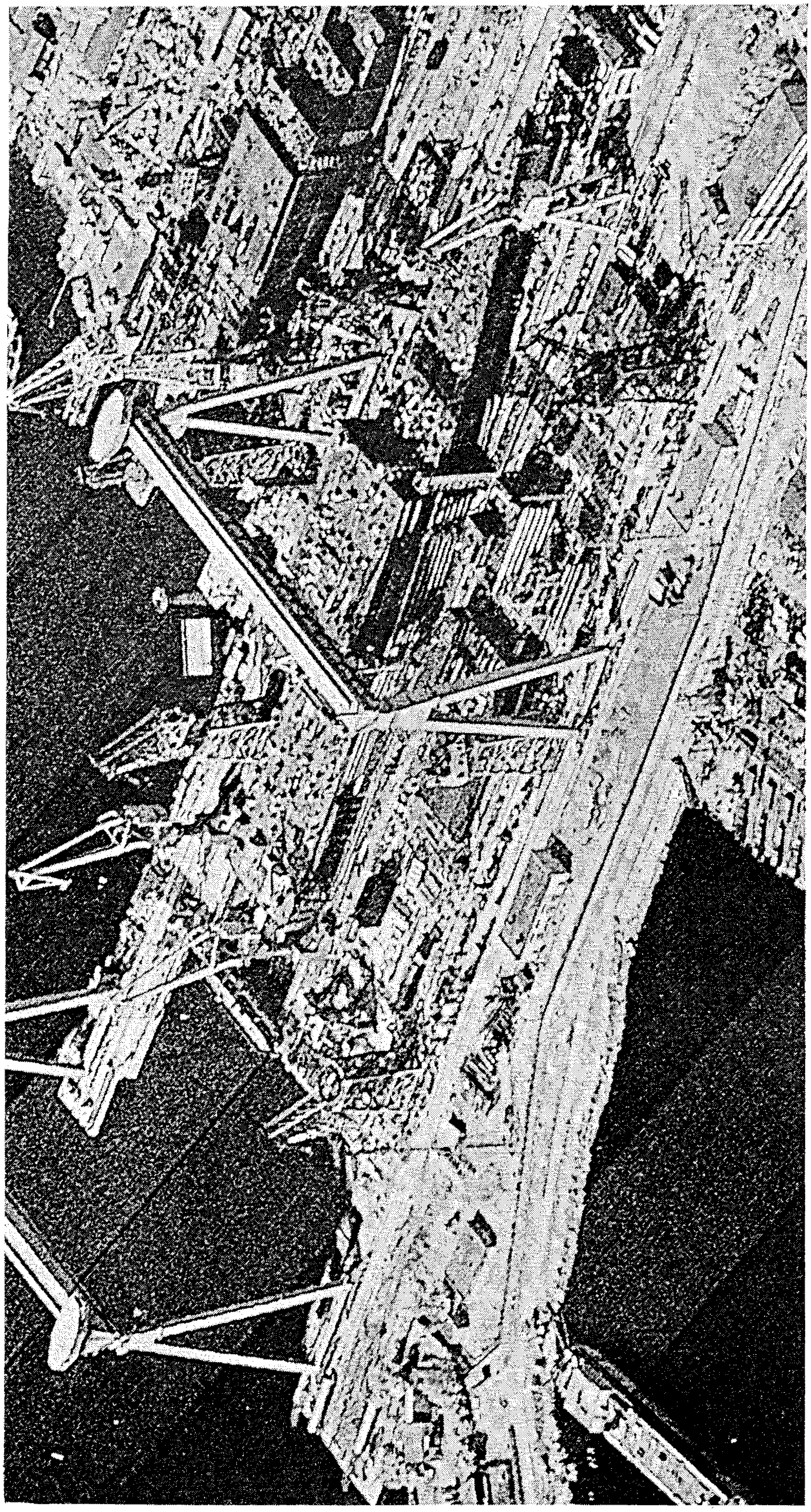


**Figure 6. Computer-enhanced satellite photograph**
A Soviet aircraft-carrier under construction at the Nikolaiev shipyard on the Black Sea as photographed by a US photo-reconnaissance satellite. The resolution of this photograph appears to be somewhat better than 1 m, probably not the best that can be achieved with current technology.

*Source: Jane's Defence Weekly.*

with 1.5 m resolution required for 'general identification'. Presumably this was the range of resolution available to the first US photographic satellites which in 1961 were able to 'spot' Soviet missile silos.[23]

In 1974, during the SALT II negotiations, it was pointed out that it would be possible to verify the proposed limitation on changes in silo diameter to a maximum of 10–15 per cent.[24] This implies that a ground resolution of at least 0.5–0.75 m was available in 1974.

Current resolution capabilities of about 0.1 m could detect even smaller changes in silo design as well as a great many other details of missile site layout; equipment such as radars, communications facilities, vehicles, storage buildings, and so on, can now also be seen and described in considerable detail.

Figure 6 shows a Soviet aircraft carrier under construction at a shipyard on the Black Sea.[25] The photo was taken from a US satellite and processed by one or more of the computerized techniques described below in the image processing section. Its resolution appears to be in the neighbourhood of 1 m, suggesting that even sharper satellite photographs are possible.

A brief study of table 2 will show that current satellite ground resolutions are sufficient to allow 'precise identification' of every item listed as well as 'description' of all but five. This adds up to a very impressive list of capabilities for satellite photo-reconnaissance.

While these capabilities are impressive and extremely valuable for verification purposes, it must be kept in mind that they represent the upper limits of achievable resolution. Such high-quality photography depends on good light, which in some important areas at high latitudes is not available for substantial portions of the year. Other areas suffer from unusually frequent cloud cover, making it impossible to photograph them for long periods of time. This limitation can be mitigated somewhat by manoeuvring the satellite to take advantage of fortuitous breaks in cloud cover. There has been a considerable effort applied over many years to accumulate accurate cloud cover statistics to be used in optimizing satellite orbits.[26] Objects which are underground, inside buildings, camouflaged or underwater cannot be photographed with visible light. However, as will be shown below, there are other techniques which can compensate to some degree for these limitations.

## III. Infra-red detection and imaging

The cameras described in the previous section use *visible* light to produce their images. Visible light has wavelengths in the interval between 0.4 and 0.7 micrometres ( μm), the same interval within which the Sun emits light with the greatest intensity. It is of course no coincidence that the human eye has evolved to take full advantage of the light emitted by the Sun.

Every object emits radiation with a spectrum of wavelength characteristic of its temperature. There are two important general laws which govern this phenomenon: one (called the Stefan-Boltzmann law) states that the total amount of radiation emitted by an object is proportional to the fourth power of its absolute temperature. The second law (called the Wien displacement law) states that the wavelength at which maximum intensity is emitted is inversely proportional to the absolute temperature.

Table 3 illustrates these laws by showing the relative brightnesses and dominant wavelengths of the same object at a number of different temperatures. Notice how strongly the brightness of an object depends on its temperature. An object at 84°C (still below the temperature of boiling water) is already emitting twice as much infra-red radiation as a body at room temperature, and by the time the object becomes just barely visible in a dark room (500°C) it is emitting 48 times the room temperature value. The same object raised to the surface temperature of the Sun would be 165 000 times brighter.

As the brightness of the object increases rapidly the dominant wavelength of the emitted radiation falls more slowly. The light from a room-temperature object is centred near 10 μm while the light from the Sun is centred near 0.5 μm, close to the centre of the visible portion of the electromagnetic spectrum.

The wavelength spectrum of any object extends well out on both sides of the maximum, although the extension to longer wavelengths is considerably greater. So, for example, the Sun emits considerable amounts of light in both the ultraviolet (less than 0.4 μm) and infra-red (greater than 0.7 μm) portions of the spectrum. Although most of the ultraviolet light is filtered out by the ozone layer, most of the solar infra-red light reaches the Earth's surface. A number of constituents of the atmosphere, especially water vapour and carbon dioxide, strongly absorb certain wavelengths of infra-red light, so the atmosphere is transparent only in certain ranges of wavelengths, called infra-red windows. The most important of these windows for reconnaissance purposes are from 0.7 to 1.0 μm (just above the visible spectrum), from 3 to 5 μm and from 8 to 14 μm.[27]

**Table 3. Relative brightness and dominant wavelength of an object at different temperatures**

| | Temperature | | | | |
|---|---|---|---|---|---|
| | (°C) | (K) | Relative brightness | Dominant wavelength (μm) | |
| Room temperature | 20 | 293 | 1 | 9.87 | far infra-red |
| Sauna | 84 | 357 | 2 | 8.13 | |
| Just visible | 500 | 773 | 48 | 3.75 | |
| ICBM plume[a] | 1 727 | 2 000 | 2 170 | 1.45 | near infra-red |
| Sun surface | 5 630 | 5 903 | 165 000 | 0.49 | visible light |
| Nuclear fireball | $10^7$ | $10^7$ | $1.4 \times 10^{18}$ | $2.9 \times 10^{-4}$ | X-rays |

[a] See Hudson, R. D. and Hudson, J. W., 'The military applications of remote sensing by infra-red', *Proceedings of the IEEE*, Vol. 63, No. 1, January 1975, p. 123.

### Photographic infra-red

Infra-red light with wavelengths between 0.7 and 1.0 μm is generally called photographic infra-red because it interacts with certain photographic films in exactly the same way as visible light, making it possible to take photographs using a broader spectrum of wavelengths. This has a number of advantages. First, at longer wavelengths the radiation is less scattered by small haze particles, so infra-red photographs taken on a hazy day will show distant objects with more clarity and contrast than visible-light photographs.[28] Other advantages derive from the high infra-red reflectance of vegetation and the greater contrast in reflectance between land and water. These can improve photographic contrasts and, most importantly, can often detect attempts at camouflage. While green paint, dying vegetation and living vegetation all look the same on an ordinary photograph, they look very different on an infra-red photograph.[29]

Film sensitive in the infra-red can be used in combination with other film to produce so-called 'false colour' images of areas on the Earth's surface.[30] The resolution of such photographs can be comparable to that of good quality black and white photographs using visible light, and 'multi-spectral' cameras are generally assumed to be part of the equipment of modern reconnaissance satellites.[31]

The use of photographic infra-red light faces problems similar to the use of visible light. Because the technique relies on reflected sunlight it is only usable in the daytime on relatively clear days. While the use of infra-red has some haze-penetrating capabilities this should not be overstated, and fog and cloud cover remain serious obstacles to satellite photography.[32]

### Thermal infra-red

The two atmospheric windows at longer wavelengths are used to observe infra-red light emitted (as opposed to reflected) from hot or warm objects.These windows lie in what is called the thermal infra-red region, generally taken to range between 3 and 14 μm in wavelength.[33]

Photographic film cannot be used to detect light at these longer wavelengths since film sensitivity falls sharply beyond 1.1 μm. But there are many other materials which are sensitive to infra-red light at longer wavelengths. Semiconductor compounds such as silicon, lead sulphide, indium antimonide and lead tin telluride can absorb infra-red light and convert the energy into a detectable voltage or current. By this principle photoelectric cells can convert solar radiation directly into electricity.

Infra-red detectors can be made both extremely small and highly sensitive. They also have some important advantages over film in that they have a linear response and a much broader dynamic range. 'Linear response' means that the electrical output signal is directly proportional to the intensity of the light that falls on the detector. Film does not respond linearly. A large dynamic range allows for much greater sensitivity to contrast variations.

On the other hand thermal infra-red imagery cannot approach photographic imagery in resolution because of two important limitations. First, the much longer wavelength of thermal infra-red radiation leads to much larger diffraction effects (see above, equation 2). In order to achieve the same 3–4 cm diffraction limit on ground resolution obtained with visible light (see above, p. 25), an infra-red telescope would have to have a diameter about 20 times as large as the Space Telescope, that is, about 50 m. Second, there are limits to the density with which infra-red sensors can be packed in an array. Each individual sensor produces an electrical signal, and a single image might consist of more than one million such signals (see below, p. 36). Any attempt to further increase resolution causes an even more rapid increase in the rate at which information must be transmitted to produce images, and any attempt to use an array of detectors with the same density as the tiny silver halide grains on photographic film would require astronomically high data transmission rates.

As a result of these two limitations the best thermal infra-red imagery will generally have resolutions about 100 times poorer than the best visible light photographs,[34] that is, at best 10 m from an altitude of 200 km. In 1972 a US Air Force meteorological satellite was reported to have a ground resolution of 600 m from an orbital height of 830 km,[35] which becomes a resolution of 150 m at 200 m altitude. By 1982 it was reported that an infra-red telescope carried by the US space shuttle would provide better than 0.1 milliradian angular resolution, which corresponds to a 20 m ground resolution from an altitude of 200 km.[36] This particular telescope is not designed for ground surveillance, but it suggests that technological developments in optics, sensor arrays and information processing may be bringing thermal infra-red imaging close to its theoretical limits.

Resolutions of 20 m or so will never produce sharp pictures of warm objects on the ground but are useful for locating and measuring the temperatures of such objects. For example, a sensor with a 20 m resolution could easily locate nuclear power or other industrial facilities that generate heat. It could also make thermal maps of areas to display subtle temperature variations which might be created by underground objects or an underground nuclear test. Such thermal mapping would also be useful in monitoring an agreement to shut down plutonium production facilities, which require either cooling towers or a river to carry away waste heat. The thermal plume from the Savannah River plutonium reactors in the USA would be readily visible from a satellite.[37]

Sensitivities of thermal infra-red detectors to temperature differences are very great, so even a slight warming or cooling of the Earth in a localized region can be detected. For example, it was claimed as long ago as 1967 that airborne infra-red sensors designed to search for submarines could, under optimum conditions, detect temperature differences of only 0.005°C.[38] This would allow the detection of submarines at depths substantially greater than 40 m, the depth at which a submarine raises surface temperatures by 10–100 times this amount. If it were in fact possible to detect and track submarines at depths of a few hundred metres this would have serious implications for the vulnerability of nuclear missile submarines, which depend for their survival on an ability to hide in deep water.

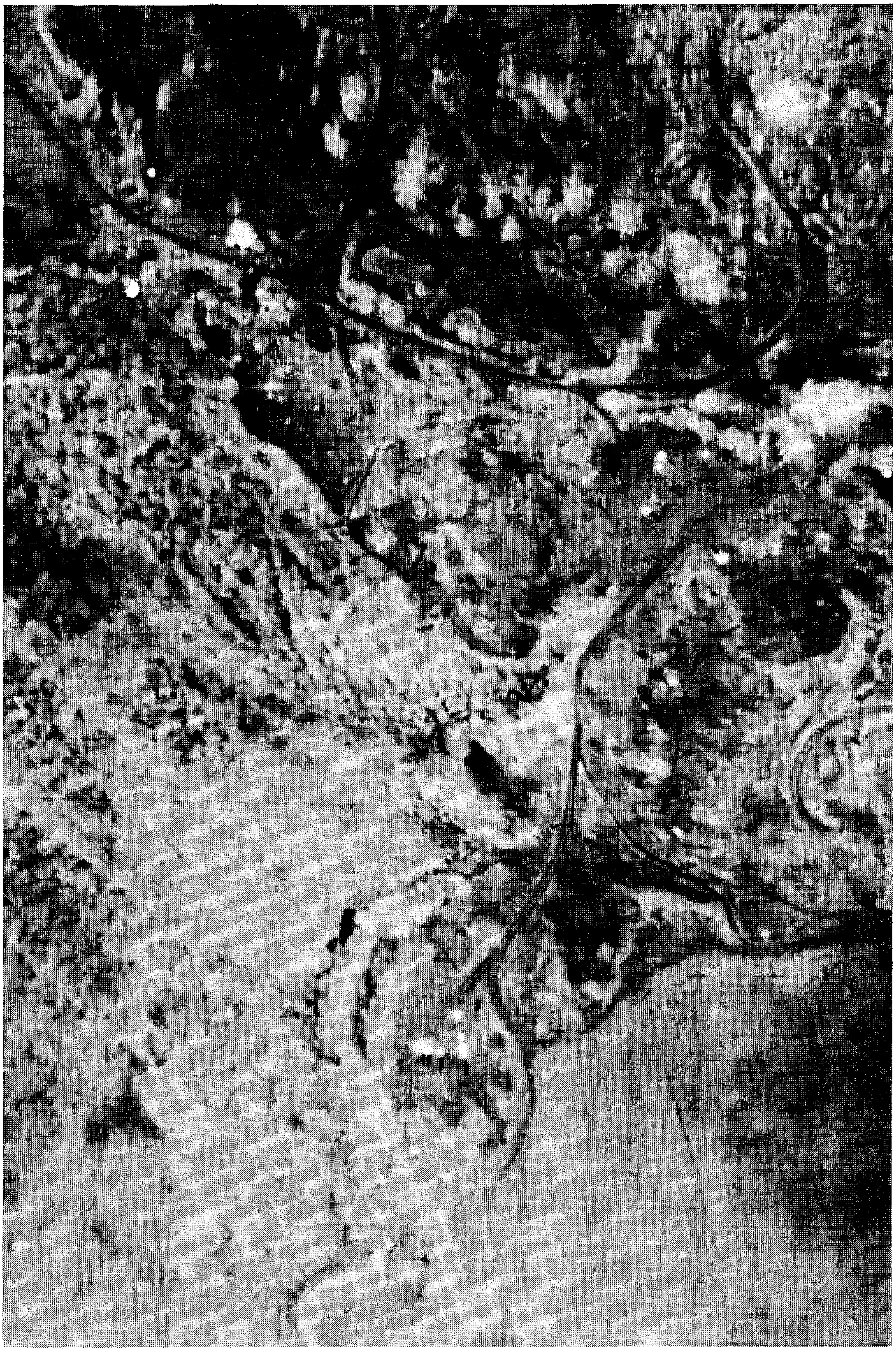
Thermal infra-red images are generally taken at night to avoid interference from reflected solar infra-red light and the elevated temperature of the illuminated background. Night photography using thermal infra-red is an excellent reconnaissance and surveillance technique and could serve many functions in verification, for example in aerial monitoring of a military disengagement zone for illegal activities. Figure 7 shows a night infra-red image taken from an aircraft at an altitude of 300 m. Clearly visible on the image are a camp-fire near a road junction (careful examination of the image shows people near the camp-fire), vehicles whose engines are still warm from recent running, and a set of aluminium foil 'resolution targets'. The very low emissivity of the aluminium makes the strips appear black, and for contrast each one has been placed next to a small pit containing three or four hot charcoal briquets, the bright spot adjacent to the black strip.

## Imaging systems

Infra-red imaging devices come in two varieties: those that operate in a *scanning* mode and those that employ a *staring* mode. In the scanning mode a single detector (or if multi-spectral detection is desired a few detectors with appropriate filters) is used in conjunction with a rotating or oscillating mirror to scan an area (see figure 8). In this way the radiation from adjacent patches of the area is focused sequentially on the detector and the current or voltage produced is monitored electronically and either stored on tape or transmitted directly to receivers on Earth, where the signal can be transformed back into an image of the scene. The image will show variations in temperature, with warmer areas appearing brighter than cooler ones. Such scanning imagers are

**Figure 7. Night infra-red image**
This image was produced from the digitized record of a thermal infra-red scanner in an aircraft at an altitude of 300 m. The large white spot at the upper left is an open campfire around which can be seen several people, recorded as small white spots. The bright spots adjacent to black strips just above right centre are vehicles whose engines had been warmed up and then turned off shortly before the image was recorded. The bright segment is the part of the vehicle which contains the engine, while the dark segment reveals a cold metal surface. The V-shaped set of images at the lower centre is a resolution target consisting of strips of aluminium foil (dark) placed next to small, 10 cm deep pits containing three or four hot charcoal briquets (bright). Notice the very different infra-red brightness of various types of vegetation and surface features (e.g., roads).

*Source*: Image courtesy of Daedalus Enterprises, Inc., Ann Arbor, MI, USA.



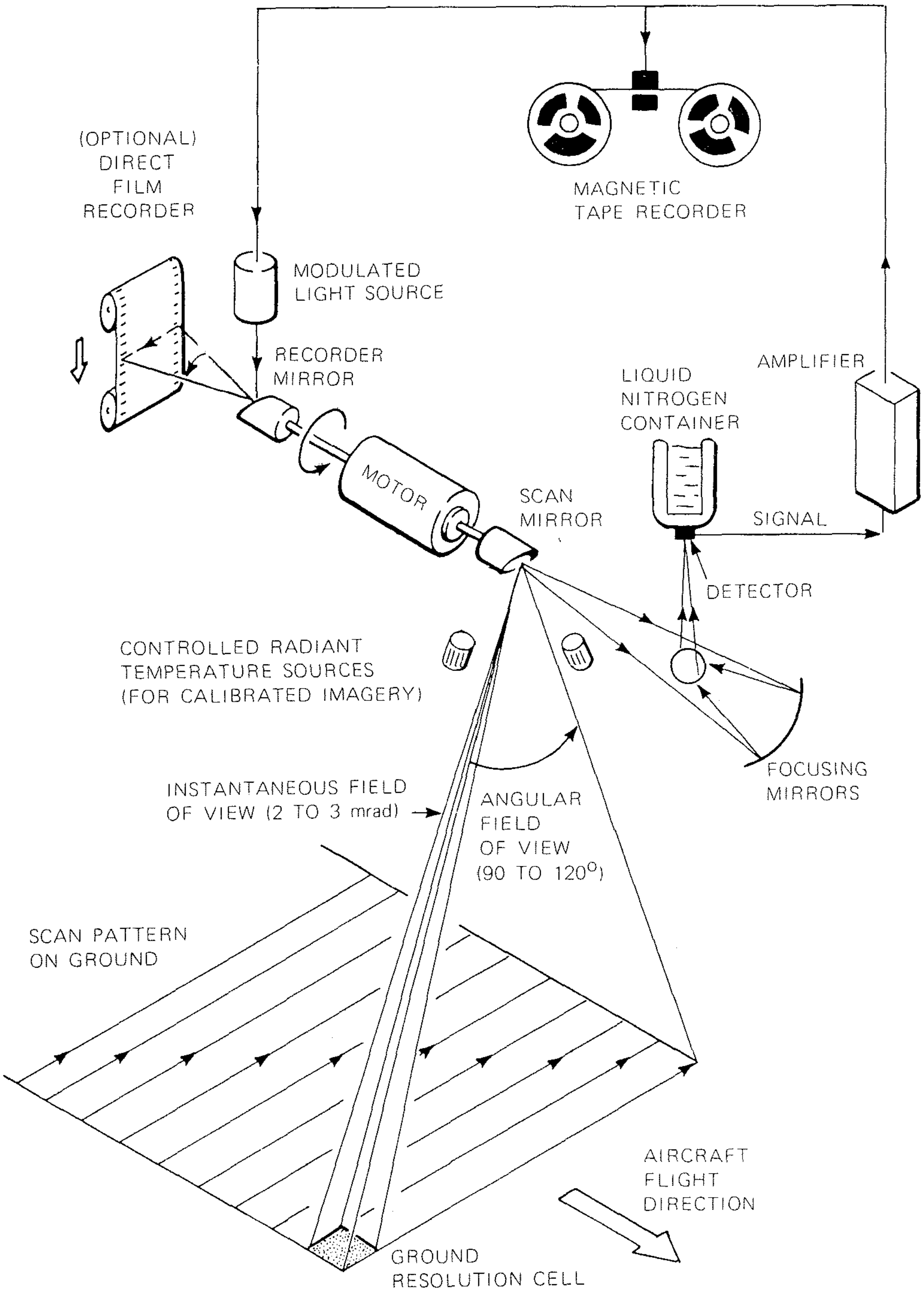


**Figure 8. Thermal infra-red scanner system**

As the scan mirror rotates it reflects infra-red radiation from a strip on the ground to a focusing mirror and then to a cooled detector. The signal from the detector is amplified and recorded. The recorded signal can then be used, either immediately or at some later time, to produce a photographic image by modulating a beam of light directed at photographic film. The calibration sources are used to provide brightness standards so that the temperatures of objects on the ground can be determined from the brightness of their images. Note that an instantaneous field of view of 2–3 mrad corresponds to a ground resolution of 400–600 m from a satellite at an altitude of 200 km or 60–90 cm from an aircraft at 300 m.

*Source*: Sabins, F. F., Jr, *Remote Sensing: Principles and Interpretation* (W. H. Freeman, San Francisco, 1978), p. 131, figure 5.9.

the most commonly used for thermal mapping and night surveillance and reconnaissance missions.

This ideal picture is greatly complicated in practical infra-red imagery by interference from other infra-red emitters. Radiation at certain wavelengths may come from objects of interest on the ground, but at other wavelengths it can be coming from some layer of the atmosphere. At still other wavelengths the atmosphere both absorbs radiation from objects on the ground and emits some of its own radiation, partially obscuring the object of interest.[39]

The radiation from an object depends not only on its temperature, but also on the nature of its emitting surface. The total radiation emitted at a given temperature depends directly on the size (i.e., surface area) of the object as well as on the radiation efficiency (emissivity) of the surface. Emissivity varies for different wavelengths, and different materials and surface textures have different emissivity functions. Therefore, if several infra-red frequencies are observed it is often possible to distinguish one type of hot object from another by comparing the infra-red 'signatures' of the two objects. For example, such signatures are associated with missile re-entry vehicles as they pass through the Earth's atmosphere. Friction with the air causes them to become very hot and to radiate intensely in the infra-red. Detection and spectral analysis of this radiation provides information on the size and shape of re-entry vehicles.[40] This technique was, of course, not created for the purpose of arms control verification, but for research and development on anti-ballistic missile systems. Nevertheless it has applications to verification, since measurements on re-entering warheads can help determine the throw-weight of a MIRVed ICBM, a parameter controlled by the SALT II Treaty.

The other form of infra-red imager is the staring type, which consists of a mosaic, or two-dimensional array, of small detectors placed in the focal plane of a telescope. Instead of scanning the field of view the imager 'stares' at it, just as an ordinary camera would do, except that the staring is continuous for the infra-red imager and not controlled by a shutter as in a camera. Such imagers can be made extremely sensitive, as illustrated by the ability of early-warning satellites stationed in geosynchronous orbits (36 000 km above the Earth's surface) to detect the exhaust plumes of missiles launched from the ground. These detectors are more than 100 times as far from their target as are the visible light cameras described in the previous section, yet geosynchronous satellites equipped with staring infra-red imagers can detect any launch of an intermediate or long-range ballistic missile anywhere within their field of view which, because of their long distance from the Earth, encompasses virtually an entire hemisphere.

While it is useful to detect objects like missile plumes and jet aircraft exhausts, it is even more useful to be able to track such objects. By tracking the exhaust plume of an ICBM during the powered segment of its flight an accurate prediction of the impact point of the warhead can be made. As usual such a capability has both military and arms control applications, and the military applications have apparently been important enough to provide the incentive to develop this capability to a high level of sophistication. US early-warning satellites can both detect and track Soviet ICBMs using infra-red radiation from the exhaust and can predict the impact point within one minute of the initial detection.[41] This precise tracking ability is useful in verifying restrictions on launch and throw-weights, but in doing so it also provides militarily important collateral data from which the accuracy of the missile can be estimated.

The exhaust plume of a missile is made up almost entirely of the products of fuel combustion, mainly carbon dioxide and water vapour. The molecules of these substances radiate energy strongly in the same spectral region as they absorb energy, at a wavelength of about 2.7 μm. Generally the detectors used to detect missile launches are designed to be most sensitive at this wavelength, which means that they cannot see the exhaust plume until it rises above any clouds, which, because they are saturated with water vapour, are opaque at 2.7 μm. However, even with this limitation the satellite is capable of tracking the missile through most of its powered flight. Another problem arises in tracking missiles launched at sea. Reflections from ocean waves cause a flickering noise background called 'ocean glitter' which is difficult to filter out of the signal received by the imager.[42]

Recent progress in micro-electronics has allowed the construction of three-dimensional mosaics in which the imaging sensors are deposited on top of sophisticated signal processing chips which convert the image directly into digital data for real-time transmission and display. Such a mosaic might consist of more than one million individual detector elements packed at a density of about 150 per square millimetre, equivalent to a resolution of 12 lines per millimetre.[43] When light falls on a detector it creates a small 'bunch' of electrons, whose size is proportional to the intensity of the light. The electrons then pass directly into a so-called 'charge-coupled device' which is capable of converting the array of one million individual electron bunches into a stream of digitized data ready for transmission to Earth.[44] Modern microprocessors allow extremely rapid and elaborate signal processing techniques to be applied to these data permitting, for example, the discrimination of the desired target from background clutter.[45] Mosaic arrays of this type can also track several objects simultaneously and, if they are made from detectors sensitive to longer wavelengths, can even be used to track relatively 'cold' objects such as satellites or re-entry vehicles in space.[46]

An example of the use of such mosaic staring detectors is the so-called 'Teal Ruby' system being prepared for the US Air Force (see figure 9).[47] This device is supposed to detect and track flying aircraft from an orbital height of 650 km. Its sensor uses 'thousands' of detector chips and is operated at a very low temperature (probably liquid helium temperature—about 4 K) to increase its sensitivity to long-wavelength infra-red radiation (up to about 16 μm).

The necessity for keeping an infra-red detector at a temperature much lower

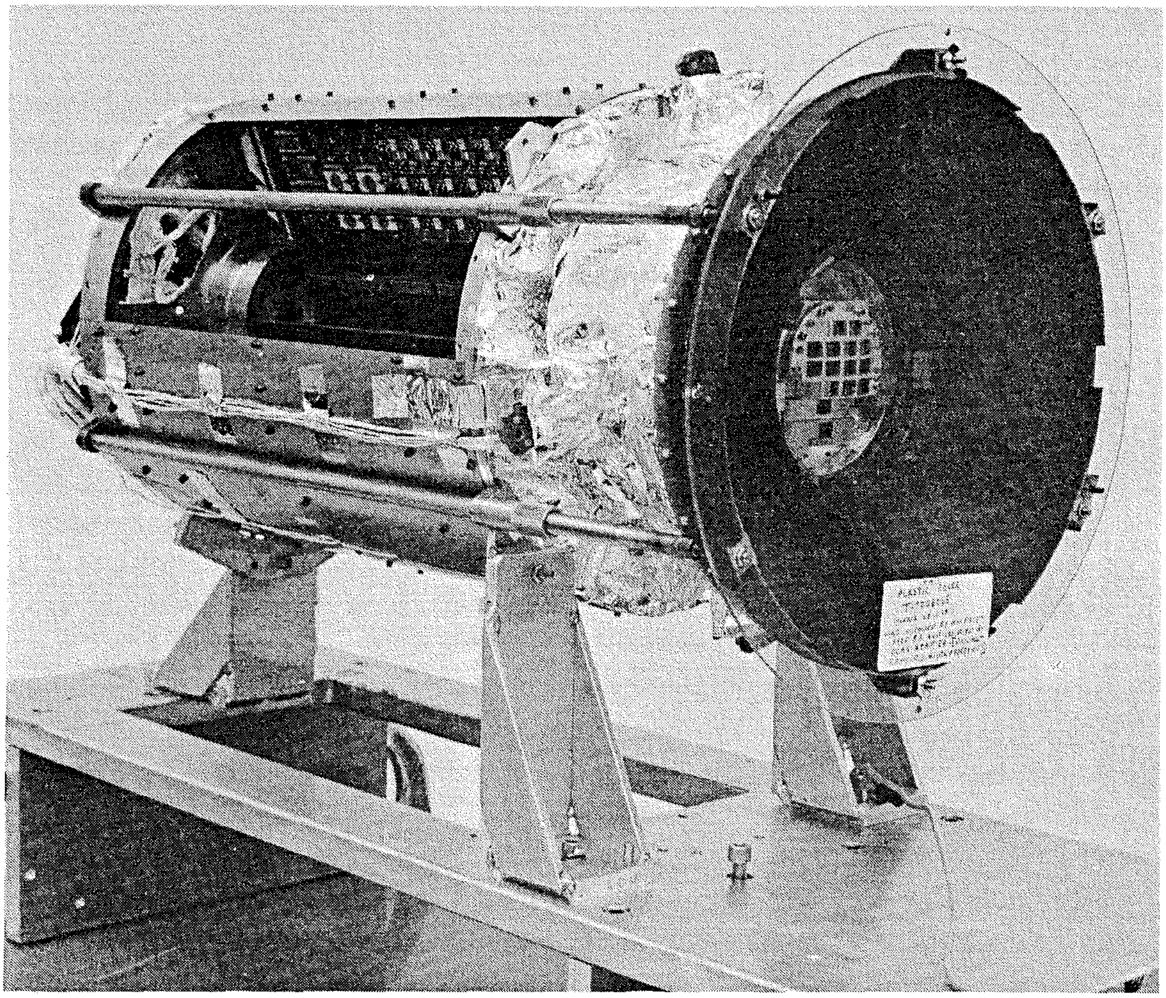


**Figure 9. Teal Ruby mosaic staring detector system**
The focal plane of the sensor is visible at the centre of the dark, circular structure, which interfaces with sensor optics. Electronic components for the infra-red detectors in the focal plane are located in the cylindrical structure.

*Source*: Photo courtesy of US Air Force.

than that of the object it is looking at can be understood when it is recognized that the detector itself is a source of infra-red radiation whose spectrum and intensity depends on its own temperature. Attempting to detect an object at 20°C with an infra-red sensor at the same temperature would be equivalent to attempting to take a photograph using film that gave off its own light. The need for long-lasting, reliable cooling of infra-red sensors adds considerably to their cost.

The Teal Ruby system has encountered a number of developmental delays and cost overrun problems[48] suggesting that it is pushing at the current technological limits. If it is successful it will demonstrate the capability of monitoring from space the flights of aircraft or cruise missiles, a capability which could be very important in verifying a ban on testing or deployment of such weapons.

## *IV. Radar*

The photographic and sensor systems described so far detect radiation of short wavelength (0.4–14 μm) either reflected or emitted by objects. These systems are generally referred to as 'passive' since they depend on radiation from other sources for their detection capability. In contrast an 'active' surveillance system generates its own radiation and then detects it after reflection from objects of interest. A common example of such an active system is a camera with a flash attachment.

Radar is an active system which employs electromagnetic radiation of much longer wavelengths than light—generally in the range 3–50 cm. A typical radar consists of a signal generator which produces a pulse of electromagnetic radiation; an antenna which sends this pulse off in a well-defined direction and then remains quiet in order to detect the return ('echo') from objects which reflect the radiation; a collection of electronic devices which process the return signal; and some form of visual display or recording device to enable the radar operator to 'see' the detected objects.

Radar surveillance has both advantages and disadvantages when compared with optical or infra-red techniques. First, because radar wavelengths are so much longer than those of light, radar waves do not interact strongly with small particles such as water droplets or suspended dust or aerosols. This means that radar has no trouble in penetrating any thickness of fog, cloud or other material opaque to short wavelength radiation. Second, because radar is an active system it can be used at any time of day or night. It has been called an "all-time, all weather sensor. . . not limited by any environmental factor".[49]

The disadvantages of radar have to do with its need for an accompanying power source, its relatively long wavelength, which means that small objects cannot be resolved and identified, and some peculiarities of image formation which make the job of image processing and interpretation more difficult. The latter will be discussed further in the section on image processing.

Despite their need for a power source radars can be quite portable. They are widely used on ships, aircraft and, more recently, on satellites, as well as at permanent ground stations. Three types of radar are most relevant to the issue of verification: large ground- or ship-based phased-array radars (PARs), used for early warning of attack, missile test monitoring, and space object tracking; over-the-horizon (OTH) radars, used to observe distant objects which are hidden from line-of-sight radars by the Earth's curvature; and synthetic aperture radars (SARs), used to produce high-resolution images of objects on the ground either from aircraft or satellites.

### Phased-array radars

The purpose of large phased-array radars is to detect and track with high accuracy a large number of objects moving at high speeds, for example the many re-entry vehicles which would be approaching a country during a massive nuclear attack. Both the United States and the Soviet Union have a number of such radars, with the US versions having such exotic names as Cobra Dane (figure 10) and Pave Paws (figure 11).[50]

In order to resolve closely spaced objects at long distances the radar beam must have a very narrow spread in angle—it must have high angular resolution. The spreading of the beam is caused by the same diffraction effect discussed above in connection with an optical camera (see pp. 24–25) and the angular spread of the beam is given by the ratio of the wavelength of the radiation to the width of the antenna. For example, the Cobra Dane radar beam has a wavelength of 24 cm (called 'L-Band') and its aperture has a diameter of 29 m. This gives an angular spread of just about 0.01 radian or 0.57 degrees. Such a beam would be able to resolve two objects 10 km apart at a distance of 1 000 km.

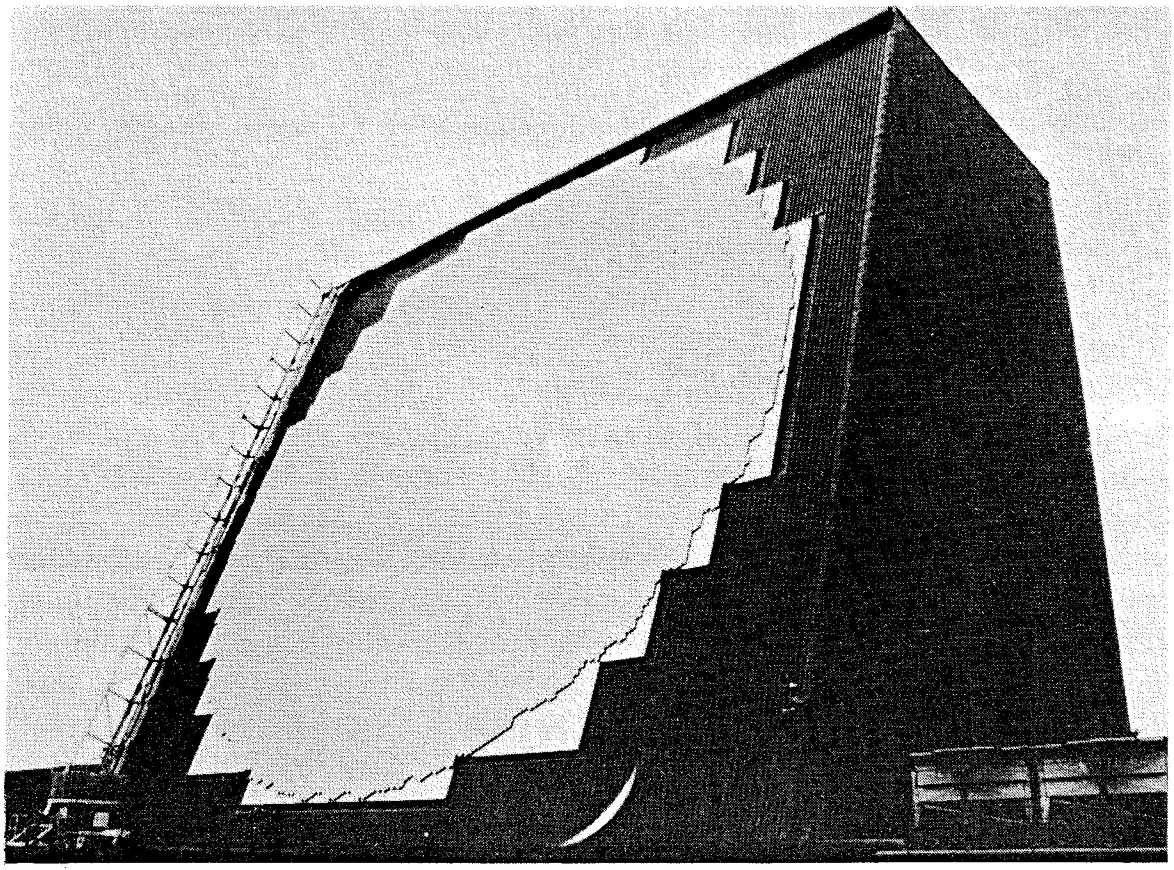


**Figure 10. The Cobra Dane radar**
This phased-array radar based on Shemya Island in the Aleutians has a diameter of 29 m and consists of 34 769 individual elements of which 15 360 are active and 19 409 are 'dummy' elements. The latter could be activated at some future time if greater sensitivity were desired.

*Source*: Brookner, E., 'A review of array radars', *Microwave Journal*, Vol. 24, No. 10, October 1981, p. 25. Photo courtesy of Eli Brookner, Raytheon Co.

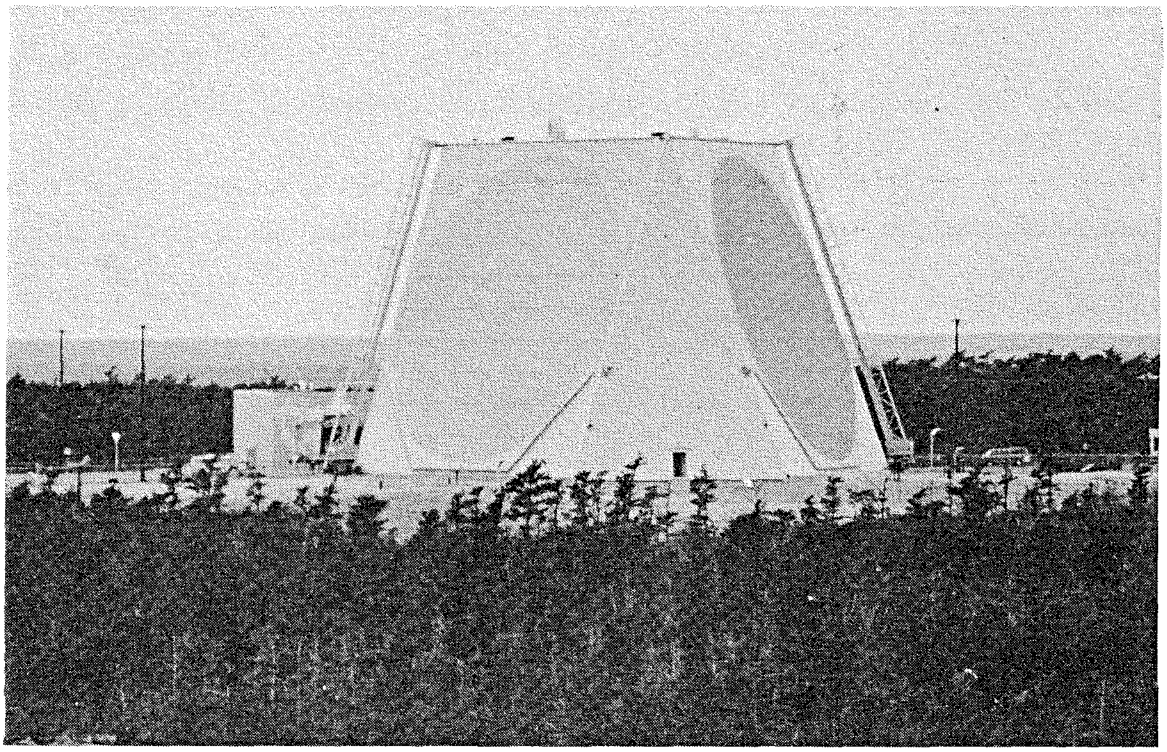


**Figure 11. The Pave Paws radar**
This radar is 22 m in diameter and uses 1 792 active elements and 885 dummy elements, a total of 2677.

*Source*: Brookner, E., 'A review of array radars', *Microwave Journal*, Vol. 24, No. 10, October 1981, p. 26. Photo courtesy of Raytheon Co.

The range resolution of a radar is defined as its ability to resolve two objects at different distances within the same beam angle. The distance to an object is determined by the time it takes the radar pulse travelling at the speed of light to go out to the object and return to the antenna. A second object slightly further away would be seen as a reflected pulse returning slightly after the first one. In order to separate these two pulses the duration in time of the pulse itself must be shorter than the time between the returning pulses. For example, the Cobra Dane radar is said to be able to resolve two objects whose distance from the radar differs by only 75 cm.[51] The radar pulse from the farther object must travel an extra 150 cm in its round trip, and since radar waves move at $3 \times 10^8$ m/s this adds only 5 nanoseconds ($5 \times 10^{-9}$ s) to the total travel time. Therefore the pulse duration transmitted by the radar must be shorter than 5 nanoseconds. (The actual radiated pulse has a longer duration than this in order to allow sufficient energy to be put into it. But by a process called 'pulse compression', which involves frequency or phase modulation, the effective duration of the pulse can be shortened to the required value.[52])

In addition to being able to locate an object in distance and angle a radar can also determine its velocity towards or away from the radar. When the pulse is reflected off the object, the frequency of the reflected radiation is changed in proportion to the speed of the moving reflector. This is called the Doppler effect, and radars can be designed to detect these 'Doppler shifts' and indicate the rate at which the distance to the object is increasing or decreasing. Velocity in the cross-beam direction cannot be measured this way and must be determined by 'tracking' the object with the beam.

In most common radars the beam is 'steered' by rotating the antenna. But this mechanical motion is too slow to allow the tracking of fast-moving objects, so in a phased-array radar the beam is steered electronically. The antenna is constructed as an array of many thousands of identical small antennas, each of which can be driven independently (see figure 12).[53] The resultant radar beam is the sum of all the individual beams, and by electronically varying the timing (phase) relationships among the many sub-beams, the full beam can be steered very rapidly. Rotations through large angles can be accomplished in millionths of a second.[54]

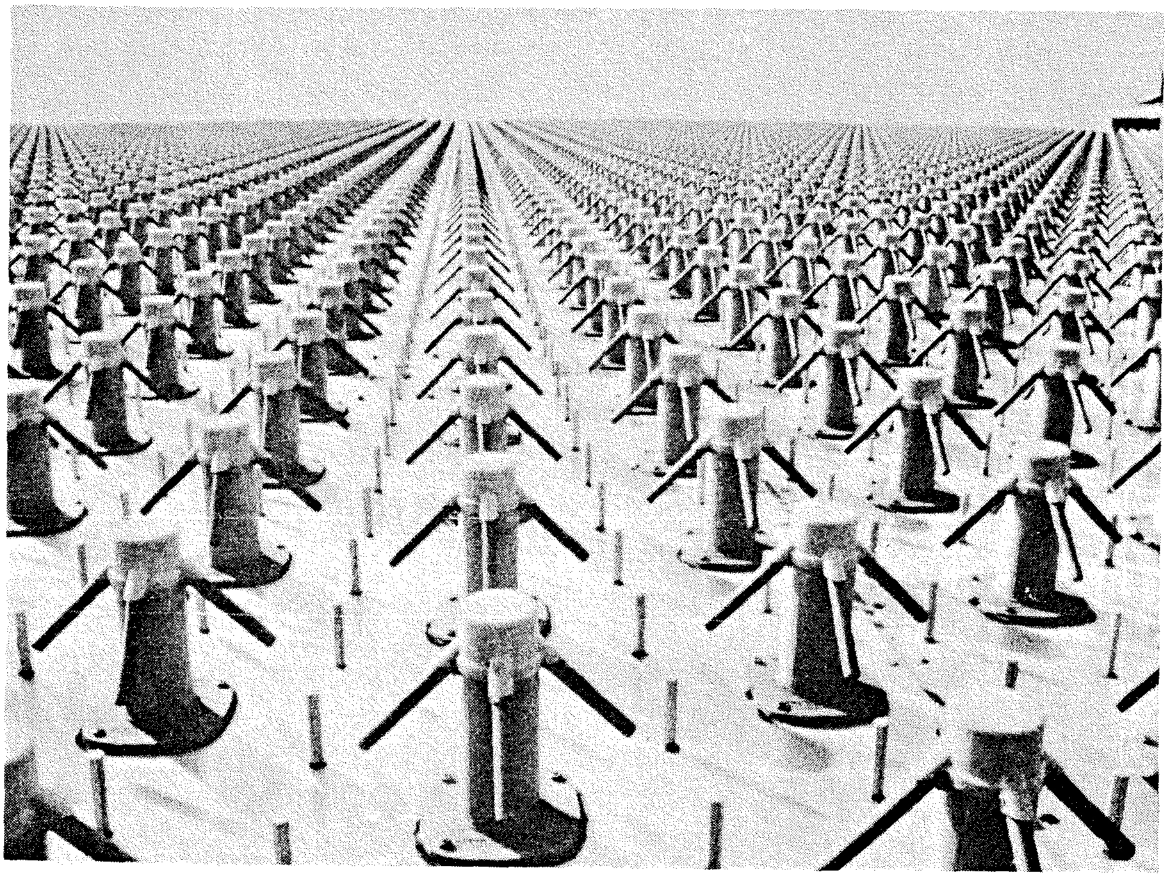


**Figure 12. Individual phased-array elements**
A close-up of the Pave Paws radar of figure 11. Each active element is a radiator of radar waves and a detector as well. Each element occupies an area of 0.14 m$^2$, which corresponds to a square with sides of about 38 cm. The elements themselves appear to be about 50 cm high.

*Source*: Photo courtesy of Eli Brookner, Raytheon Co.

This feature permits the Cobra Dane radar to track as many as 100 re-entry vehicles simultaneously at a distance of 2 000 km and make accurate measurements of their speed and trajectories.[55] Located on Shemya Island in the Aleutians, and accompanied by a smaller but similar ship-borne PAR called Cobra Judy, the Cobra Dane is well placed to monitor Soviet intercontinental ballistic missile tests. A number of qualitative features of Soviet ICBMs, such as throw-weight and accuracy, can be determined in this way.[56]

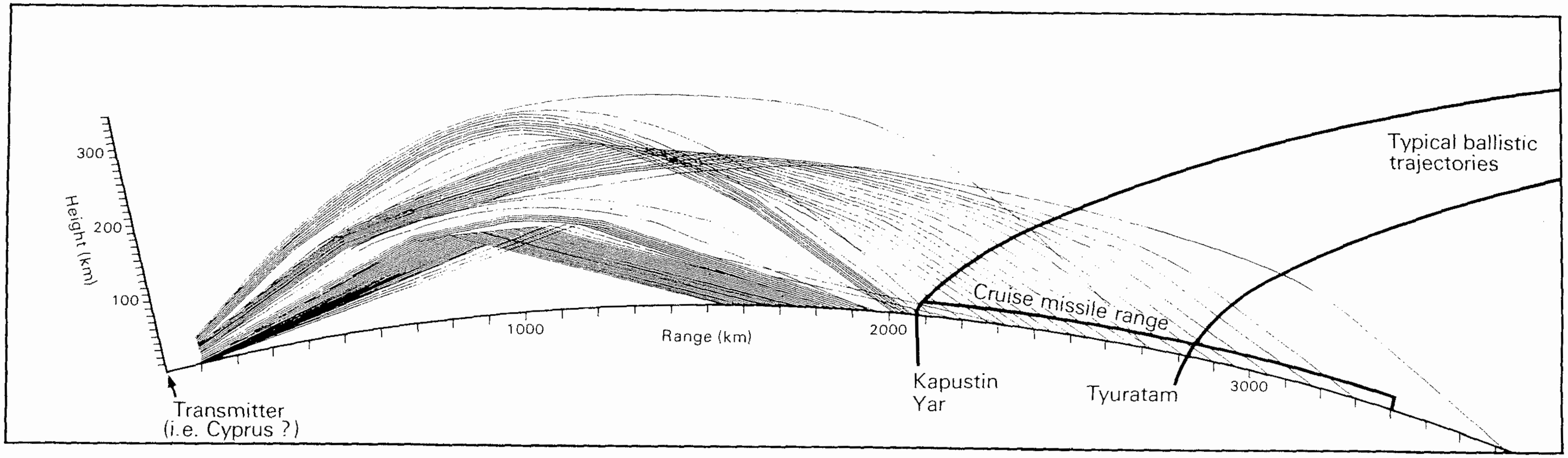
Phased-array radars are essential to any attempt to create an effective anti-ballistic missile (ABM) system, an application recently highlighted by the US discovery of such a radar under construction by the Soviet Union at Krasnoyarsk in Siberia. The USA has accused the Soviet Union of a violation of the ABM Treaty on the basis of this discovery. In reply the Soviet Union has made similar charges against the Shemya Island Cobra Dane radar. Such problems of interpretation are not surprising since phased-array radars can serve a wide variety of functions in which the simultaneous tracking of a number of flying objects is necessary. While the design of a given radar may be optimized for a specific purpose, for example to monitor tests of ICBMs, the performance characteristics are virtually indistinguishable from those needed to support an ABM system. A phased-array radar can therefore be at the same time a national technical means of verification and an apparent violation of a treaty. Such ambiguities are extremely difficult to resolve in a technical way. They are the stuff of political compromise.

### Over-the-horizon radar

Normal radars are limited in useful range because the beam they produce travels in a straight line, while the Earth's surface is curved. This means that a beam emitted parallel to the Earth's surface at one point will pass other points on the surface at progressively higher altitudes, making it impossible to detect low flying aircraft at large distances, even though the beam still has sufficient power. For example, an aircraft flying at an altitude of 5 000 m cannot be seen by ground-based, line-of-sight radar at distances greater than 250 km.

The possibility of using radar at considerably greater distances arises from the reflection of radar waves by the Earth's ionosphere, a layer of electrically charged gases at an altitude of from 80 to a few hundred kilometres. When they encounter the free electrical charges which constitute the ionosphere, radar waves are partially reflected and can return towards the Earth's surface at distances of from 1 000 to 3 000 km from their point of origin (see figure 13).[57] If they are reflected from an object during this downward portion of their path, the waves can return along roughly the same path and be detected near the original antenna location.

It is not surprising to learn that an OTH radar antenna must be both large and powerful if it is to accomplish such a task. One such radar under develop-



LiberKartor, Stockholm

**Figure 13. Backscatter over-the-horizon (OTH) radar**
Radio beams from the transmitter are reflected back down to Earth by the ionosphere. Scattered radio energy is reflected back to the transmitter along the same paths. Objects such as missiles reflect a signal detectable at the transmitter site. The figure shows trajectories of missiles from Kapustin Yar and Tyuratam superimposed on the ray paths of a typical OTH radar.

*Source*: SIPRI, *World Armaments and Disarmament, SIPRI Yearbook 1980* (Taylor & Francis, London, 1980), p. 298.

ment by the US Air Force and the General Electric Corporation consists of an antenna which is 690 m long located on the eastern coast of Maine in the USA (see figure 14).[58] Plans are eventually to expand this antenna to a length of 1 100 m and to have it employ a variety of wavelengths between 11 and 60 m. At the 11 m wavelength the beam should have an angular width of 0.01 radian, giving it a target resolution of 20 km at a distance of 2 000 km (see equation 2, p. 24). The beam can be steered electronically to track moving objects in a manner similar to that of a phased-array radar.[59] Such a broad beam would not be able to resolve and identify aircraft by their size or shape, but it could detect and track objects, such as ballistic or cruise missiles, which would be difficult to observe reliably in any other way. The greatest advantage of a stationary OTH radar over satellite-based sensors is its ability to maintain more-or-less continuous surveillance of relatively small areas, such as missile test ranges.

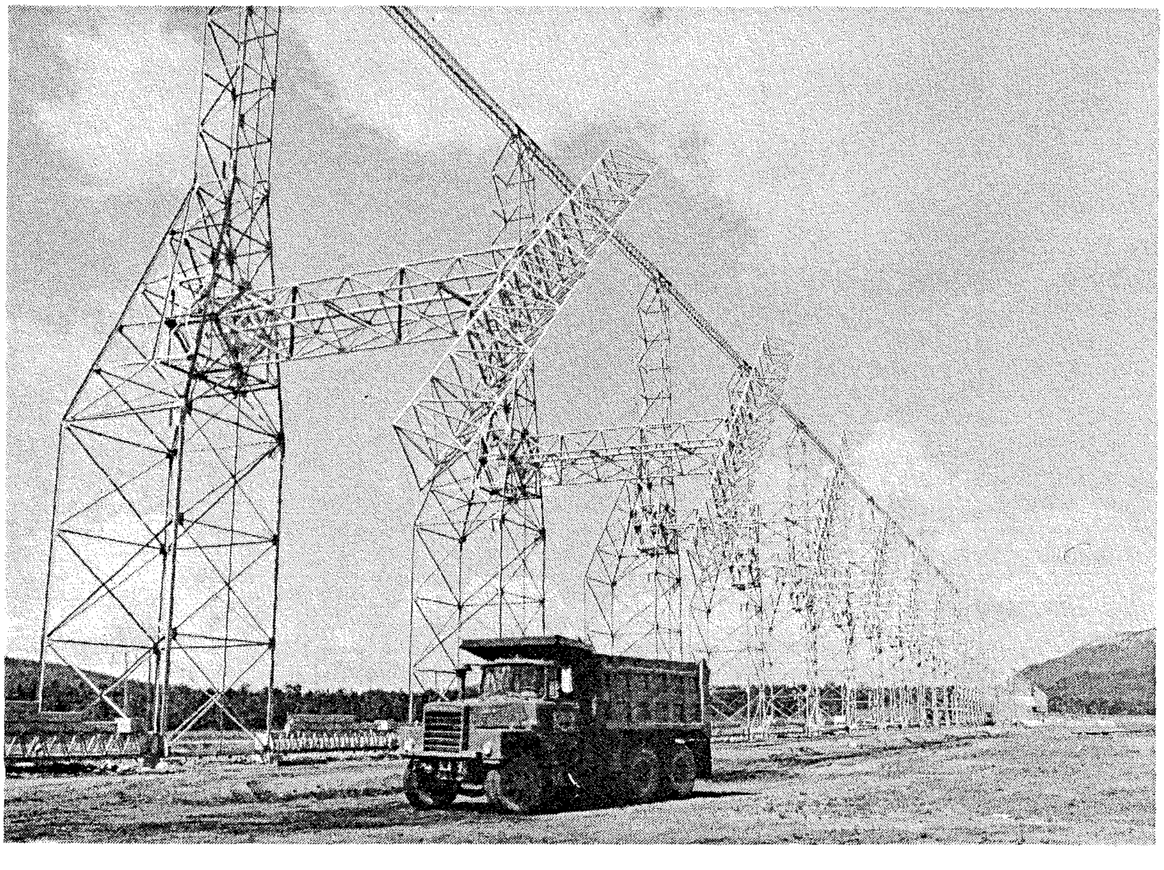


**Figure 14. OTH transmitting antenna array**
This experimental OTH transmitting antenna was located near Moscow/Caratunk, Maine in the USA and had a length of 690 m. The receiving and signal processing site is near Columbia Falls, Maine about 175 km to the south-east. The transmitting antennas are the diagonal elements in front of the 30 m high towers. The antennas were driven by twelve 100 kW transmitters powerful enough to detect and track aircraft at ranges of 3 000 km.[58] The experimental system has been dismantled and is in the process of being upgraded to an operational system.

*Source*: US Air Force.

There are some serious difficulties which restrict rather severely the applicability of such radars to verification. The radar must be located so that the major portion of the beam path is over water, and this naturally restricts the number of areas that can be monitored in this way.[60] There has been speculation, but no firm evidence, that the USA has such a radar deployed in Cyprus, an ideal location for observing flight tests of missiles or aircraft at the Soviet testing centres of Kapustin Yar and Tyuratam.[61] If such a radar exists it may be possible for the USA to monitor the boost phase of Soviet rocket tests and use these data along with information gathered by other sensors to verify SALT limitations on launch weight and throw-weight.

It is interesting to note that given the size and power of such a radar, it would be impossible to keep its existence and capabilities secret from a state like the Soviet Union, which possesses sophisticated satellite and electronic reconnaissance systems. Therefore, if such an installation does exist, there is no military justification for keeping it a secret. However, such secrecy does have political motivations and these are examined in chapter 4.

### Synthetic aperture radar

One feature held in common by all radars which resolve small objects at large distances is the enormous size of their antennas. The OTH radar just described is nearly 700 m long, and the Cobra Dane and Pave Paws radars have diameters of 29 m and 22 m respectively.[62] The areas of the two PARs are 660 $m^2$ and 384 $m^2$, and each must be constructed from thousands or tens of thousands of individual radiating and detecting elements.

The requirement for large antenna size is dictated in part by the need to put large amounts of energy into the beam in order to be able to detect small, distant objects. But it is also required if a beam with a narrow angular spread is to be achieved, that is, if the diffraction spreading phenomenon is to be minimized. As was shown above, even the 29 m diameter of Cobra Dane is capable of only about a 10 km resolution at a distance of 1 000 km. So even if such a large antenna could be carried on a satellite, its ability to resolve small objects on the ground would be quite limited. For example, to achieve a 10 m resolution from an altitude of 200 km would require an angular beam spread of only 50 microradians, 200 times narrower than the Cobra Dane beam. Producing such a narrow beam of radiation with a wavelength of 24 cm would require a circular antenna 200 times the size of Cobra Dane—almost 6 km in diameter. Obviously some other method must be used to obtain high-resolution radar images from satellites.

This method is called ‘synthetic aperture’ radar. It uses a relatively small antenna but takes advantage of the motion of the antenna relative to the ground to create the same effect as that of a very large antenna. Figure 15 illustrates how this works.[63] A satellite or aircraft passing over some region of interest emits radar pulses which are directed downward and to either side

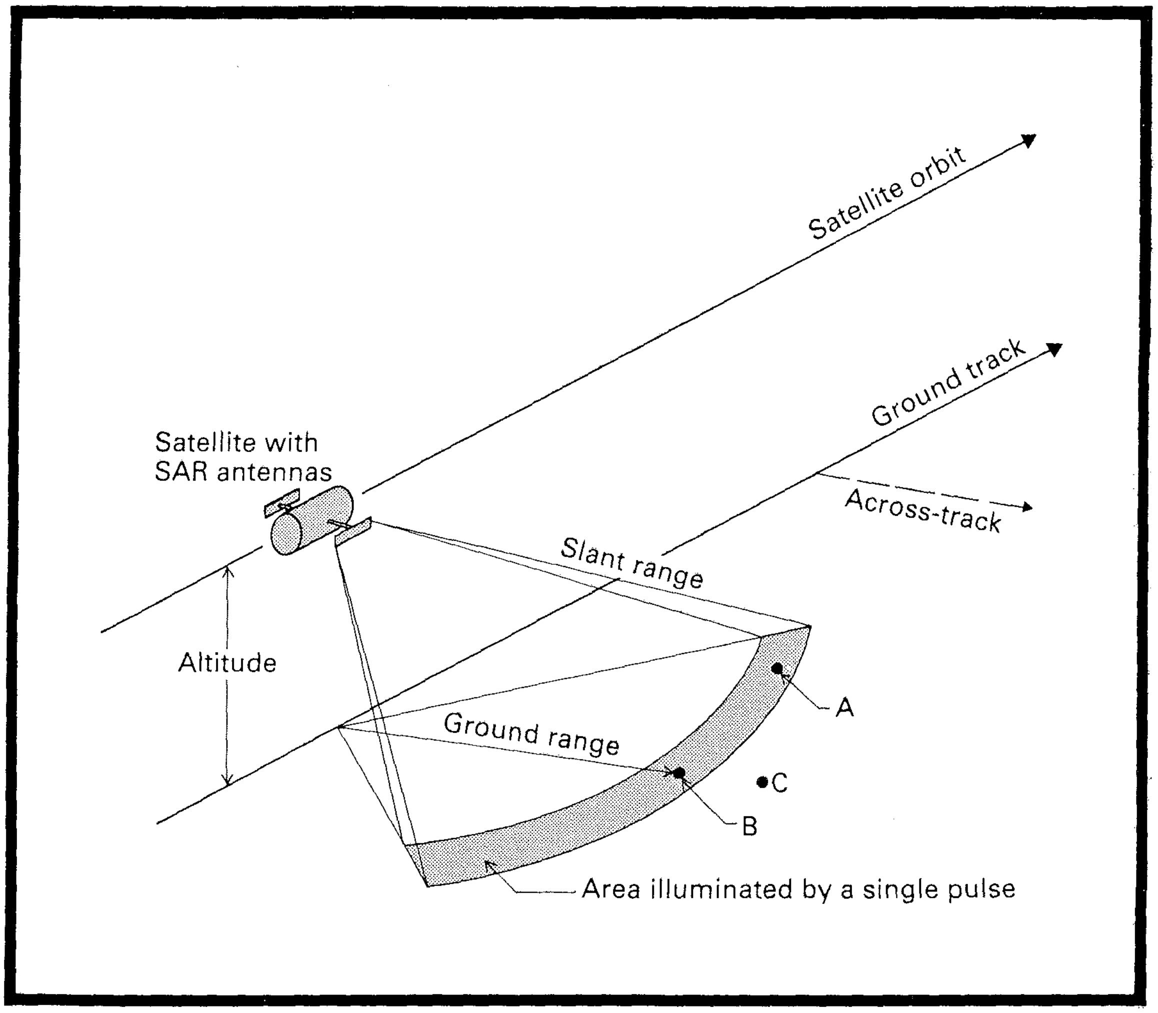


**Figure 15. Synthetic aperture radar**

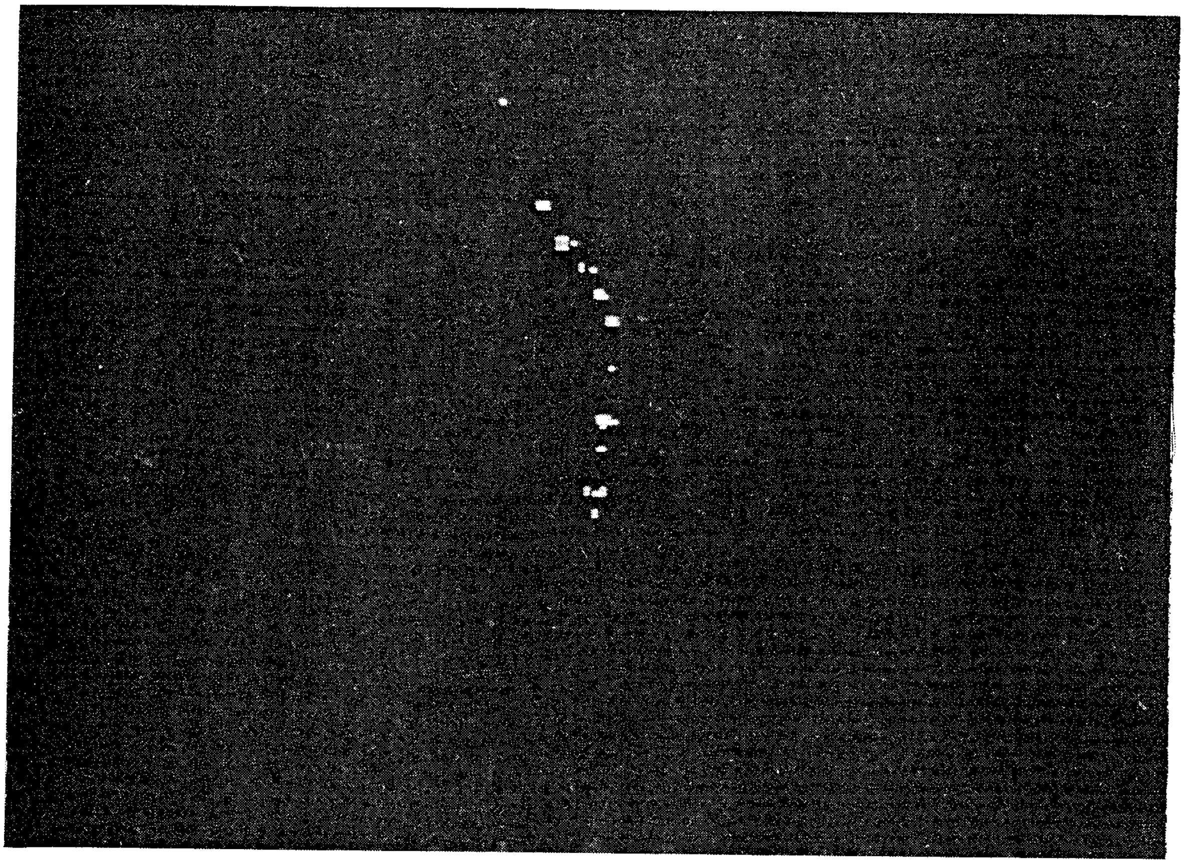
Pulses of radiation emitted by the satellite antenna are reflected from objects on the ground and received by the same antenna. The shaded area on the ground shows the width of a strip defined by the duration of a single pulse. Two objects within that area (A and B) cannot be distinguished by a single pulse from the radar. However, if all the pulses which reflect off both A and B while the satellite is passing overhead are stored in the satellite and processed, these objects can be resolved.

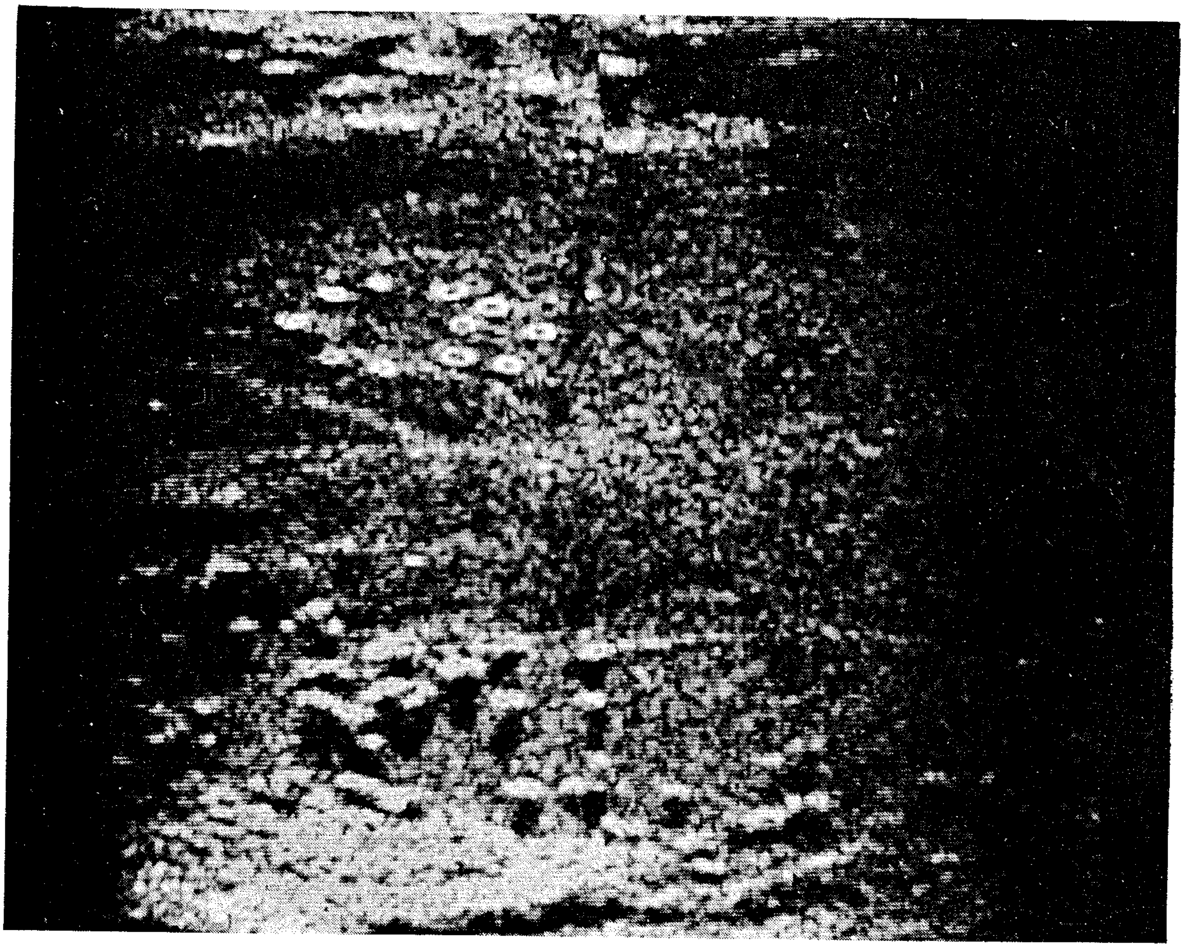
of the ground track. Reflections of these pulses from objects on the ground return to the satellite where they are detected and recorded. The ground resolution problem can be described in terms of three objects on the ground denoted by A, B, and C. Objects A and C are at different distances from the satellite, so their resolution in range is accomplished in the normal way. The pulse is made of very short duration so that the return pulses from two objects very close together can be distinguished. For example, if C is 1.5 m further from the antenna than A, then the pulse to and from C must travel 3 extra metres at the speed of light, a delay of 10 nanoseconds ($10^{-8}$ s). This requires only that the radar pulse be less than 10 nanoseconds long, a goal already seen to be achievable in ground-based radars (see above). Therefore, resolutions in range of the order of 1 m should be achievable from satellite or airborne SARs.

Resolution in angle is another matter, and figure 15 shows the objects A and B to be equal distances from the satellite and both within the angular spread of the beam from the relatively small antenna carried by the satellite. The width of this beam can be estimated if it is assumed that the radar wavelength is 10 cm and the antenna is 5 m long. Such a beam would have an angular spread of 0.02 radians (see equation 2, p. 24) which means that a swath 10 km wide would be illuminated at a slant range of 500 km. Even if A and B were as much as 10 km apart, the return pulses from them would be indistinguishable and they would not be resolved.

The solution to this problem lies in the fact that A and B will stay within the beam for many pulses, and that during this time their spatial relationships to the satellite will change continuously in different ways. As an example suppose that A and B lie at a distance of 400 km from the satellite ground track, and that the satellite is at an altitude of 300 kilometres. The radar pulses to A and B must then travel a round trip of 1 000 km (this example assumes that the Earth is flat; a more accurate treatment would change the numbers somewhat but not the essential features of the example), which takes 3.3 milliseconds, during which time the radar must be in the receiving mode. Once the echo is received from the objects of interest another pulse can be emitted, so if A and B are at the outer boundaries of the region being surveyed the radar can emit pulses at the rate of 300 per second. Since the satellite is moving at the rate of 7.5 km/s and the width of the beam is 10 km (see above) an object such as A will stay within the beam for 1.33 seconds, during which time it will reflect 400 pulses back to the satellite, each from a slightly different location relative to the satellite. If the entire history of 400 pulses from each object could be analysed then enough information would exist to distinguish A from B and in fact distinguish objects even much closer together in the direction parallel to the ground track.

This process of combining the information of many pulses from a moving antenna is exactly equivalent to what can be done using a single pulse from a very long antenna. So the resolution obtainable parallel to the ground track is the same as could be obtained using 24 cm waves from an antenna with a length of 20 km or about $10^{-5}$ radians. At 500 km range this represents a ground resolution of only 5 m, comparable to the size of the antenna. A more rigorous mathematical treatment of this problem shows that the theoretical limit of resolution is one-half the length of the antenna.[64] So in principle this satellite could achieve resolutions of 1.5 m in the cross-track dimension and 2.5 m parallel to the ground track in a swath 800 km wide. Such resolution is about 10 times poorer than that obtainable from optical infra-red photographs, but this sacrifice in resolution is compensated for by the ability of synthetic aperture radars to obtain their pictures through the heaviest cloud cover and at any time of night or day. It is also very important to note that the ground resolution of an SAR, unlike all other imaging techniques discussed so far, does not depend on the distance between the antenna and the





target.[65] This means that SAR images from satellites can have just as high resolutions as those taken from aircraft at less then one-hundredth the altitude. It also means that SAR satellites can be placed in higher, longer-lasting orbits as long as sufficient power is provided to make up for the additional wave-propagation distance.

The problem of supplying the electrical power for an SAR is a serious one. Existing space-based SARs require at least 20 times as much power as optical photographic systems, and if this is to be supplied by arrays of photovoltaic solar cells these arrays must be very large and expensive. Such concerns have led to serious discussion of using nuclear power sources for military SAR systems, and such power sources are under active development.[66]

Reference to table 2 (p. 26) will show that if SAR resolutions of 1–2 m are achievable, they will be very useful in many monitoring tasks, especially if SAR is used in conjunction with other, higher resolution forms of imagery. As to when such resolutions may be available, one forecast predicts 1 m resolutions from space-based radars by the year 2000 and states that "Radar component capabilities and available power sources are such that progress in achievable resolution is mainly paced by available data-handling rates".[67] Meanwhile, SARs mounted on aircraft such as the US RC-135, TR-1, or SR-71 reconnaissance aircraft probably already have at least such resolutions. One source attributes to such airborne SARs a range of 300 nautical miles (560 km) "enabling the TR-1 to 'see' at least into Eastern Poland [from FR Germany] and probably beyond. On surveillance missions the TR-1 can cover 131,800 sq nm [450 000 km$^2$] per hour".[68] Figure 16 shows images of tank formations obtained with one such radar which is produced and advertised by the General Electric Company.

The major difficulty which remains to be solved for satellite-based SAR is the rapid processing of vast amounts of data, and this problem is discussed further in the next section. As this obstacle is overcome much more extensive use of SARs can be expected for a wide variety of Earth survey, military intelligence and arms control verification tasks.

## *V. Image processing*

The information obtained from optical and infra-red photography and radar is in the form of images. These can be photographs, readings from sensors,

**Figure 16. SAR images of tank formations**

Two SAR images, one of a tank/truck column (top) and the other an assembled tank formation, (bottom) were made by an airborne radar system called Multimode Surveillance Radar (MSR) manufactured by the General Electric Corp.

*Source*: Photos courtesy of General Electric Aerospace Electronic Systems, Utica, NY, USA.

or detected radar signals which have been recorded on film, magnetic tape, or in the memory of a computer. These images must now be put into a form in which they can be examined and analysed by skilled interpreters. It is conceivable that some day this process of recognition and interpretation of images might be almost totally automated, and this particular aspect of "artificial intelligence" research is receiving considerable attention.[69] But at present, and for the foreseeable future, the involvement of a skilled and experienced human intelligence is essential for the interpretation of photographic images. Considering that ground resolutions of 10 cm are now possible, and comparing this with the vast areas that are routinely photographed (not all at such high resolutions, of course), it is clear that the number of images being routinely scanned and interpreted by intelligence analysts must be enormous.[70] There are simply not enough analysts to handle the flow of military and commercial information. For example, probably 90 per cent of the data gathered to date by the US Earth Resources Satellite programme has not yet been analysed, and there exists a genuine danger of the intelligence system being swamped by unmanageable amounts of data. A similar flood of data has inundated the US Infrared Astronomy Satellite (IRAS) programme.[71] If such quantities of data are to be effectively utilized, automated analytical methods will have to be devised to reduce the load on human interpreters by filtering and pre-analysing images.[72]

The images received from modern satellites are almost never analysed in their raw form. They are first processed to make the job of interpretation easier. Image processing is a general term which includes two sub-classes of operation: *restoration* and *enhancement*.[73] Image *restoration* is the process of correcting certain image defects caused by transmission through a less than perfectly transparent medium, distortions and limitations of optics, relative motion of camera and target, incorrect exposure times, and so on. Such restoration is generally based on some mathematical model of the processes which have degraded the image. Its object is to produce the highest possible fidelity of the image to the object it represents.

The purpose of image *enhancement* is to alter the image in ways which clarify or accentuate objects of interest and suppress unwanted background or redundant information. While enhancement can also employ mathematical models,[74] the range of possible techniques is far broader, more flexible and more subjective than formal models would permit. Image enhancement has perhaps more appropriately been called a "bag of tricks"[75] whose objective is to produce optimal image "quality", a concept for which no mathematical criterion exists. Such techniques include manipulations of contrast and colour and the sharpening of edges to highlight objects of interest.[76] In such manipulations the skill and imagination of the human interpreter are an essential ingredient, and since different interpreters have different levels of skill and imagination; "an image which causes one analyst to conclude that no enhancement is possible may be treated with great success by another analyst. For images with great significance, such as those which might be used in weapons verification monitoring, it is disturbing to think of the consequences of an analyst failing to produce the optimum visual quality from a given image".[77]

Virtually all image processing is now carried out on digital computers, and the first step is therefore to convert the image into digital form. The only major exception to this generalization is the photographic technique of displaying synthetic aperture radar pictures (see below). Images from sensors which convert light directly into electrical signals can easily be converted into digital form for direct transmission back to Earth. Images recorded on photographic film are digitized by developing the film and scanning the image with a light-sensitive sensor. The digitized electrical signals from this sensor are stored on magnetic tape or in a computer memory for further processing.

The typical digital format for image processing is to have each picture element, called a 'pixel', represented by a binary number of 8–12 bits.[78] An 8-bit number would produce a 'dynamic range' of brightness values from 0 for total black to 255 for full white. There is in fact a wide variation in the dynamic-range capabilities of different sensors. Photographic film or a television screen can cover a dynamic range of only about 100, while modern charge-coupled device sensors can have dynamic ranges of 5 000, that is, they can distinguish 5 000 different levels of brightness.[79] Such discrimination is obviously helpful in situations where subtle differences of brightness are important, but it also adds to the information processing demands, since a dynamic range of 5 000 requires that each pixel be represented by 13 binary digits instead of 8. In practice this would involve the use of chips with 16-bit word lengths, since such chips are relatively cheap and available.

Some idea of the amount of information contained in a single high-resolution photograph can be obtained by imagining a 15 cm × 15 cm photograph with a film resolution of 50 lines per mm. A single pixel on such a photograph would measure only 20 $\mu m^2$ and the entire picture would contain 56 million pixels. A digitized record of such a photograph would therefore contain 56 million 8-bit binary numbers. High-resolution aerial or satellite photographs can contain more than 100 times as much information as this.[80]

Once an image has been digitized there are a wide variety of operations that can be carried out under the general rubric of image restoration or enhancement. A few of these can be mentioned briefly here, and more details can be found in the references.

## Image restoration

Two examples of image restoration are noise suppression and corrections for lens or mirror distortions. 'Noise' is a familiar phenomenon in all signal processing. A television viewer attempting to watch a programme coming from a distant transmitter will see 'snow' on the screen as random noise signals compete with the weak programme signal. The hiss or static on a weak radio station is another example.

Noise is an essentially random phenomenon, so it is amenable to analysis by mathematical techniques which exploit this randomness. Since an information-carrying signal (say a photograph) has a high degree of coherence, it is possible to devise computer routines (called noise-cleaning masks) which accentuate this coherence and suppress random noise signals.[81] Figure 17 illustrates an example of the effects of one such noise-cleaning operation.

All optical systems introduce some distortion into the images they create, although this can be minimized by careful design and construction. The remaining distortion can be analysed mathematically, both from basic optical principles and by empirical measurements on the actual optical system. This analysis can then be translated into a computer program which can be applied to any image produced by the system to remove the distortions. Such routines can also be used to produce a sharp focus in slightly out-of-focus image or to correct for blurring caused by the relative motion of camera and subject (see figure 18). In principle, and almost certainly in practice, any degradation or distortion of an image which can be expressed in an empirical or analytical algorithm can be corrected for in this way. Even diffraction effects can be reduced by such algorithms, although there is no way to obtain information

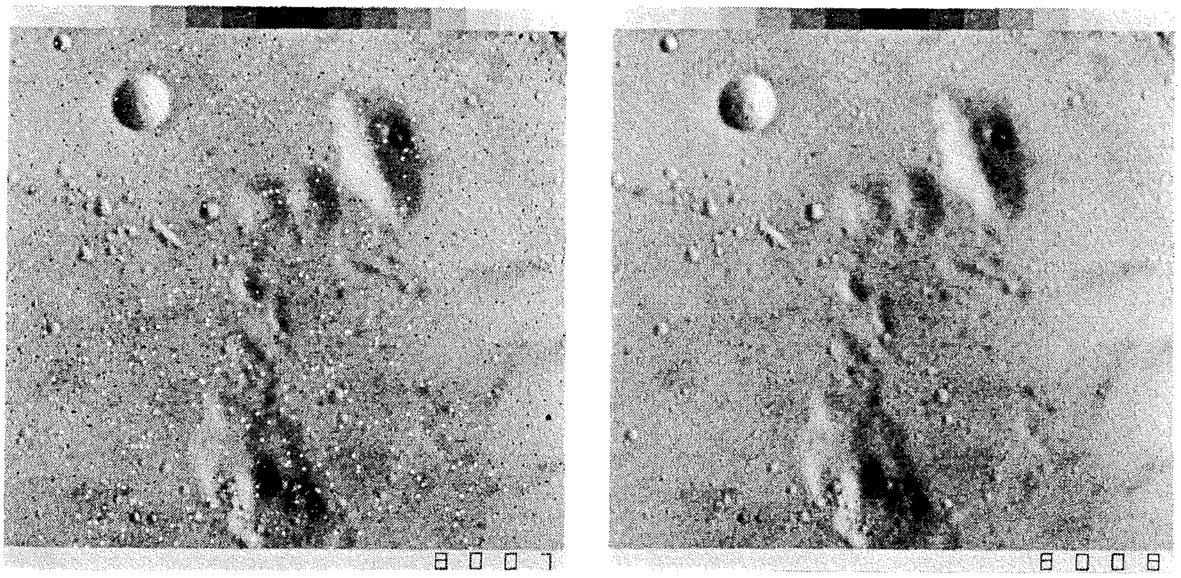


**Figure 17. Noise cleaning**
Image noise caused by sensor or signal errors usually produces random pixels which are very different from their neighbours (see image on left). These noise pixels can be removed by a simple computerized algorithm which computes for each pixel the difference between its brightness value and the average brightness of the eight nearest neighbours. If this difference exceeds some chosen threshold (49 in the images above) the deviant pixel is replaced by the average of the neighbours. The image on the right is the result of the noise cleaning algorithm.

*Source*: Photos courtesy of Vicom Systems, San José, CA, USA.

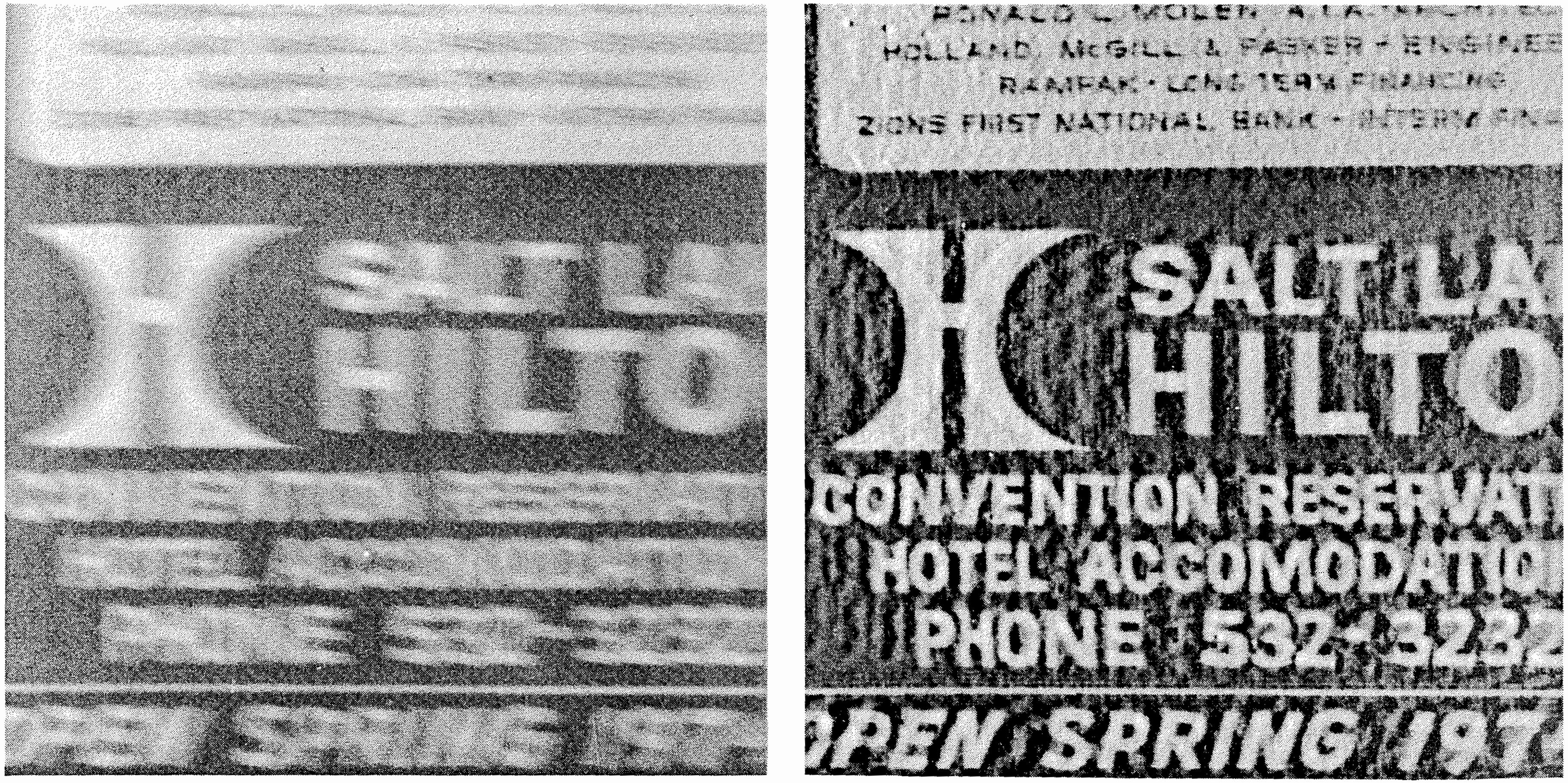


**Figure 18. Motion de-blurring**
The photograph on the left was blurred by intentionally moving the camera during the exposure. A computer program determined the direction and the extent of the camera's motion from the digitized photograph (left) and then produced the restored image (right).

*Source*: Photos courtesy of Michael Cannon, Los Alamos National Laboratory, Los Alamos, NM, USA.

beyond the fundamental limit described above (see pp. 24–25).[82] Still, if the diffraction limit is 5 cm, objects with dimensions of 10–20 cm will be significantly distorted by diffraction effects, and their resolution could be greatly improved with a diffraction-correcting algorithm.

## Image enhancement

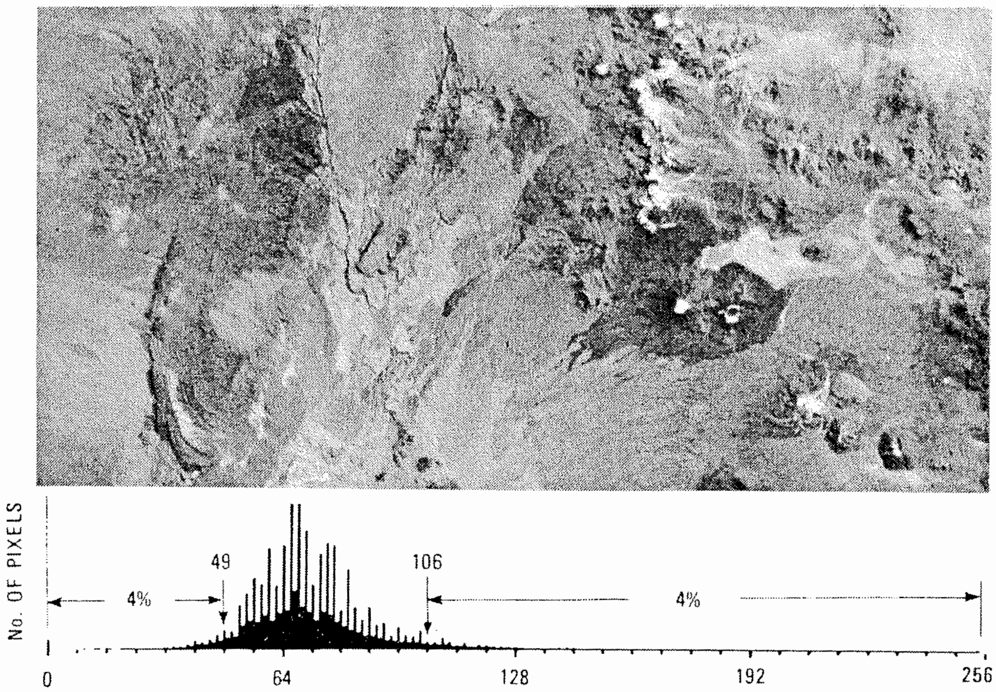
The single most important application of image enhancement is in the manipulation of contrast to increase the visibility of objects in shadow, obscured by haze, or photographed with too much or too little exposure. Contrast enhancement can be done in several ways, only two of which will be mentioned here: histogram equalization and adaptive filtering.

Histogram equalization[83] begins with the construction of a histogram, that is, a frequency distribution of brightness in the picture. This is done by counting the number of pixels having each brightness level and plotting these numbers on a bar graph.[84] An underexposed or low-contrast picture will utilize only a small portion of the available dynamic range of film, and contrast can therefore be enhanced by expanding the histogram to take up the full range. By redistributing some brightness values the histogram can also be levelled. The two processes greatly enhance the contrast in the picture, as illustrated in figure 19.

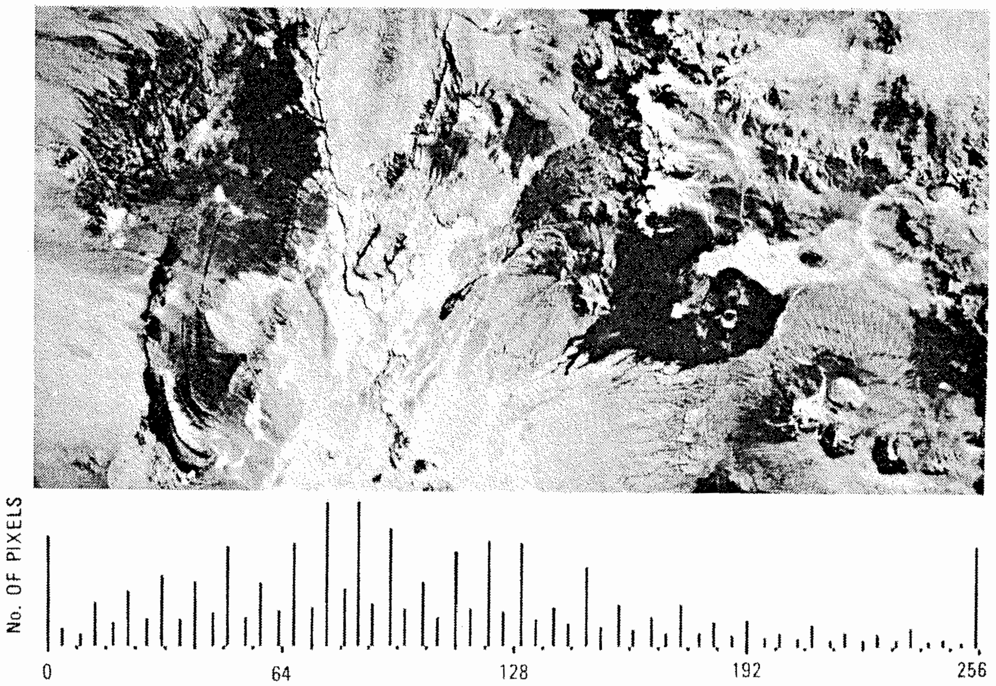
A somewhat more complex technique that achieves similar results is called adaptive filtering and is illustrated in figure 20.[85] The thin cloud cover almost totally obscures the image in two ways: first, it partially obstructs the transmission of light from the ground to the camera; and second, it reflects considerable amounts of light directly back to the camera causing overexposure of the haze relative to that of the ground. The adaptive filtering process first computes the average brightness and the local contrast values for all regions of the picture and then reduces the average brightness and increases the variations in such a way that contrast is greatly enhanced.

One of the most important purposes of image enhancement is the detection of objects, some of which may be so small that their images comprise only a few pixels. Such small, indistinct images are extremely difficult to pick out with the unaided eye, so a number of techniques have been developed to make object detection more reliable and efficient. One such technique is optical image subtraction in which two images of the same scene taken at different times are optically combined in such a way that the earlier image is 'subtracted' from the later one.[86] The result is an image which records only the *changes* in the scene in the interval between the two images, thereby highlighting objects which have been moved into or out of the area.

A second object detection technique uses an intensity prediction algorithm somewhat analogous to the noise cleaning masks described above,[87] but now the purpose is to enhance anomalies instead of eliminating them. In this process the expected intensity of each pixel is predicted from the intensity



A. ORIGINAL IMAGE WITH NO CONTRAST ENHANCEMENT.



B. LINEAR CONTRAST STRETCH WITH LOWER AND UPPER FOUR PERCENT OF PIXELS SATURATED TO BLACK AND WHITE RESPECTIVELY.

**Figure 19. Histogram equalization**
The top picture shows a low-contrast satellite photograph of a region on the Chilean–Bolivian border. Beneath the photograph is the histogram of pixel intensity values, and the narrowness of this histogram is directly related to the lack of contrast in the photograph. The lower picture is the result of the histogram equalization process which makes use of the full dynamic range of the display medium to enhance contrast and emphasize details which are obscure on the unprocessed image.

*Source*: Sabins, F. F., 'Thermal infrared imagery and its application to structural mapping in Southern California', *Geological Society of America Bulletin*, Vol. 80, 1969, pp. 397–404, figure 2.

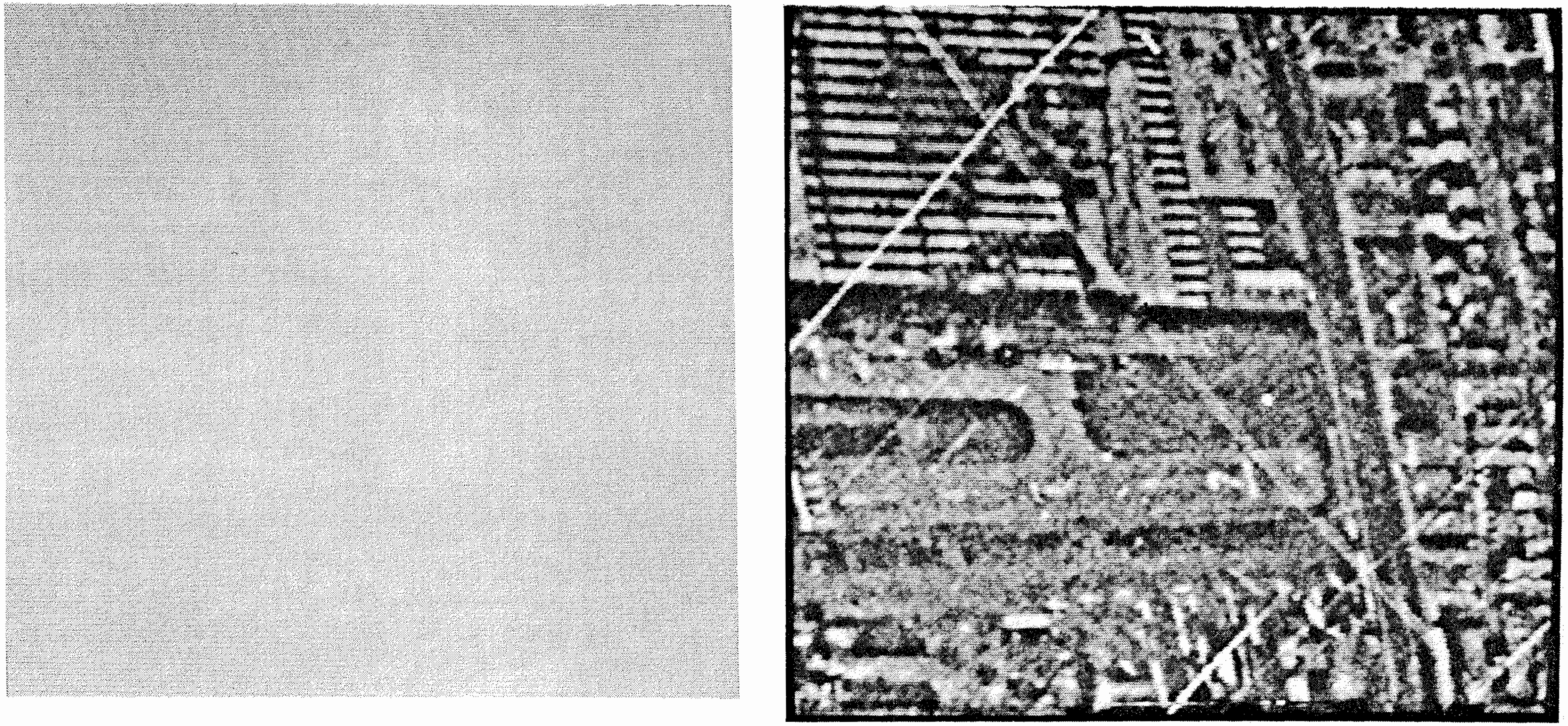


**Figure 20. Adaptive filtering**
The photograph on the left is of an airport runway almost totally obscured by haze. The image on the right is the result of applying the adaptive filtering contrast enhancement process mentioned in the text. The diagonal lines on the enhanced image are called 'digital line-artifacts' and can be removed by yet another restoration procedure.

*Source*: Peli, T. and Verly, J. G., 'Digital line-artifact removal', *Optical Engineering*, Vol. 22, No. 4 July/August 1983, pp. 479–484, figures 12 and 14.

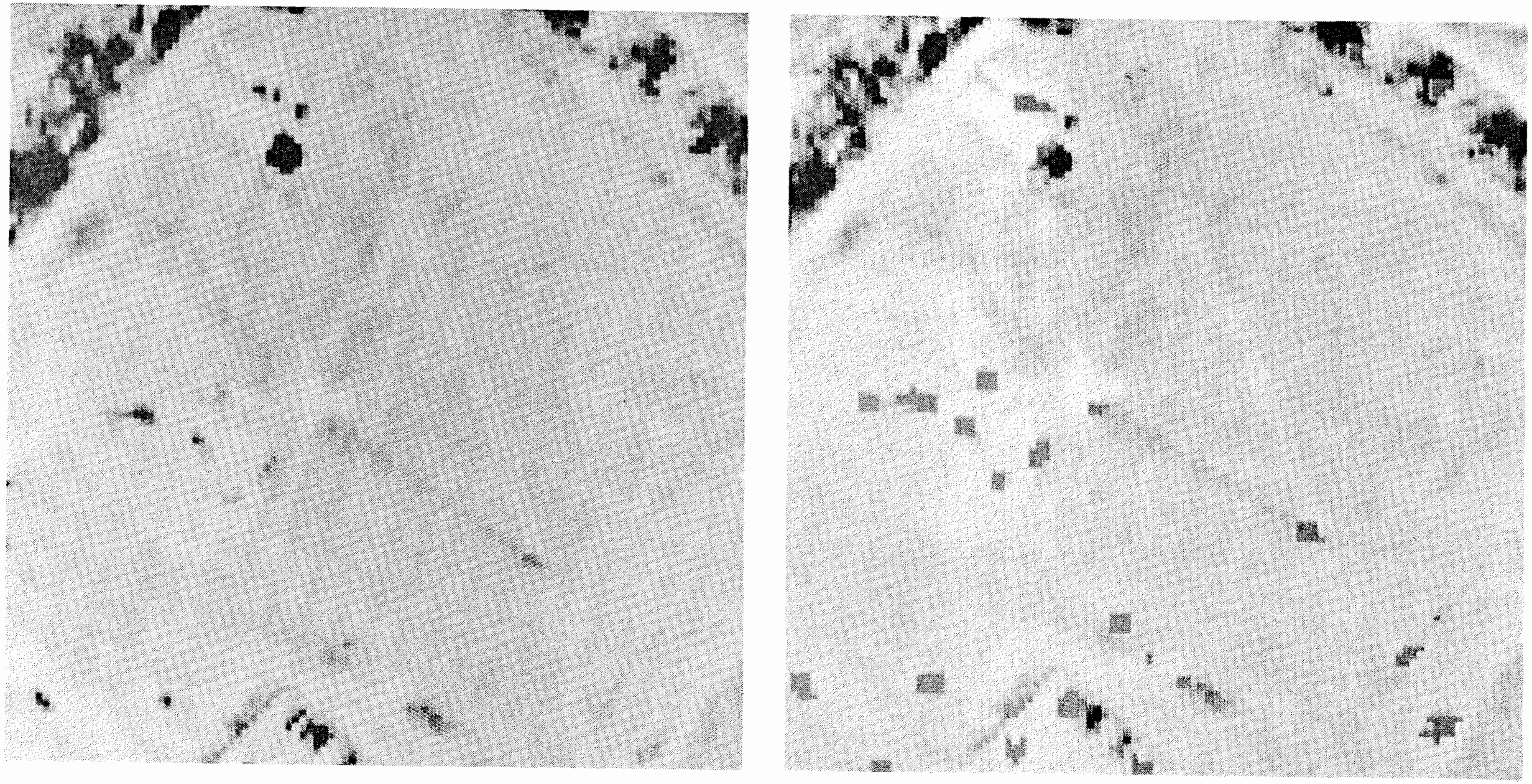


**Figure 21. Object detection**

The image on the right is the result of applying the object detection algorithm to the low-resolution image on the left. The ‘objects’ are the rectangular dark grey spots which are made up of at most a few pixels and are therefore not identifiable. While it is possible that very small ‘objects’ are in fact noise pixels, the algorithm is designed to make the number of false alarms small. Therefore it is highly probable that most of the objects detected are in fact real objects which might be identified by a higher resolution image acquired at some future time.

*Source*: Photos courtesy of Thomas F. Quatieri, Lincoln Laboratories, MIT, Lexington, MA, USA.

values of a large number of pixels in its neighbourhood. A statistical test is then applied to determine if the pixel intensity differs significantly from the predicted value, in which case it is classified as an anomaly, that is, an object (see figure 21). An object detection process such as this would be very useful in examining low-resolution images to determine if there is sufficient interest to warrant the taking of higher resolution photographs to better identify the detected objects.

These are only a sample of the image enhancer's 'bag of tricks', and many more exist.[88] Generally an interpreter working with an important picture will use several such techniques, and it has now become possible to build special computers capable of applying these techniques so rapidly that an interpreter can experiment with various restoration and enhancement techniques sitting at a computer console and observing the changes in the image in real time.[89] Even relatively complex operations such as image subtraction can be accomplished in real time.[90]

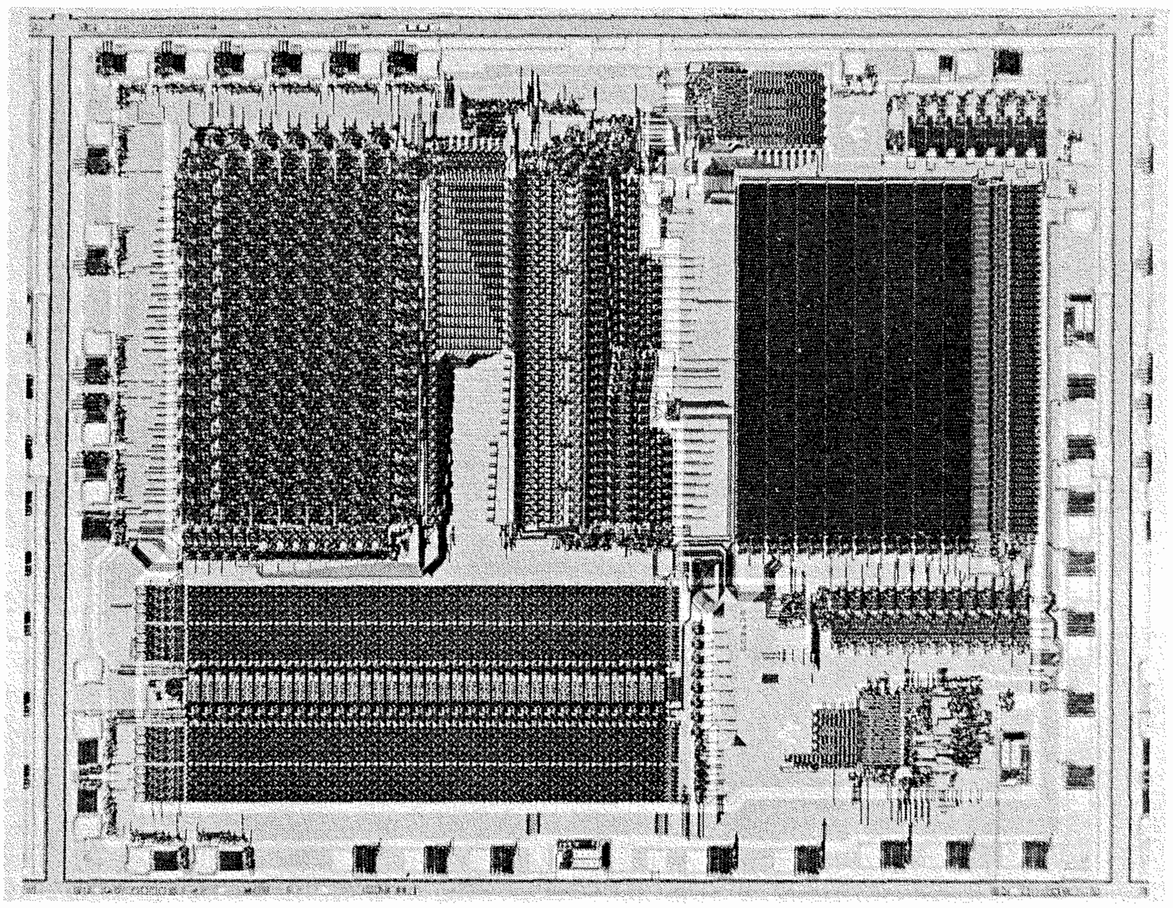


**Figure 22. VLSIC chip**

This very large-scale integrated circuit combines two types of memory and a central processing unit on a chip that measures about 6 mm square. It can execute five million instructions per second, about 100 times faster than conventional microcomputers, making it very useful in image processing applications.

*Source*: Photo courtesy of Texas Instruments Inc.

It is clear from just the small sample of techniques described here that remarkable improvements can be made in satellite photographs as long as sufficient computer capacity and ample numbers of skilled interpreters are available. Given the rapid reductions in size and increases in speed of present-day and projected computers, the problem of sufficient capacity would seem to be soluble. For example, it has been estimated that a contrast enhancement of a single image requires between 100 million and 1 000 million individual computer operations.[91] But one forecast of computing capabilities predicted image processing speeds of $10^{12}$ bits per second by the mid-1980s.[92] Such a computer could perform hundreds of contrast enhancements per second, enabling the kind of real-time interactive interpretation described above. The situation is expected to continue to improve as the development of micro-electronic technology continues. Improvements have in some cases proven to be even more rapid than expected, producing memory chips which can store one million bits of information on a 50 $\text{mm}^2$ chip.[93] Other chips designed for very high-speed processing have achieved rates of 100 million multiplication operations per second.[94] Figure 22 shows a so-called 'very large-scale integrated circuit' (VLSIC) which combines both memory and processing functions on a single chip. Individual feature sizes on such a chip are as small as 5 μm.[95]

## Radar image processing

A radar image is created from radiation which has been reflected off distant objects and returned to a detector. The return signal is in the form of a fluctuating voltage, which must be processed electronically to convert it into a bright spot on a screen calibrated to show the distance and direction to the object. Such radar images are the stock in trade of air traffic controllers, fighter pilots, reconnaissance aircraft, ship navigators and many others. A similar image is obtained from the large phased-array radars, although with these the display must be more sophisticated and the amount of electronic processing much greater.

For synthetic aperture radar the signal processing problem is truly gigantic. As was shown above, the location of a single object with good resolution will require the information contained in many hundreds of complex voltage pulses. Each pulse reflected from the object also contains information on the many other objects encountered by that same pulse, and the creation of a detailed image requires an enormous number of elementary mathematical computations.

A example of the magnitude of the problem is the SAR imagery obtained from Seasat, a US ocean surveillance satellite which was placed in orbit in 1978.[96] Seasat images had 25 m resolution and the satellite could transmit the raw data for a 40 × 40 km image (2.56 million pixels) in 2.5 seconds. The digital processing of these data into a visible image requires the equivalent of

10 billion ($10^{10}$) multiplication and addition operations. As recently as 1979 this process required 25 hours of computer time for each 40 × 40 km image.

Such computational problems explain why digital methods have not been used extensively for SAR image processing until very recently. The traditional method has been to use the returning radar signal to modulate a beam of light which in turn exposes photographic film.[97] The returns from a single radar pulse then appear as a thin vertical strip of varying brightness on the film. The film is moved so that subsequent pulses will produce adjacent strips until an entire piece of film is exposed. The image produced on the film is a 'hologram' of the scene, that is, an image which is related to the scene by a complex mathematical transformation.[98] The transformation can be performed optically by shining a laser beam on the hologram and manipulating the transmitted light with lenses. In this way the hologram can be converted into a high-resolution 'photograph' in a single operation.

The optical process requires no digitizing or computing (it is called an 'analog' process), but it does require the use of film with all of the accompanying inconvenience of chemical development. This has made it awkward to use in satellites, and unsuitable for the production of images in real time. Nevertheless, some very high-quality images have been obtained in this way and in January 1982 the optical process was still six times faster in producing images than the best digital processor available at that time.[99]

However, the rapid increase in speed and compactness of digital computers promises that high-quality, real-time SAR images will be available in the very near future.[100] One system under test is designed for use in fighter aircraft. It will achieve resolutions of 2.5 m and process the images with a computer capable of performing 45 million complex operations per second and storing 3 million bits of information (300 000–400 000 pixels). The computer itself weighs only 32 kg, uses 375 watts of power and occupies a volume of only 0.05 $m^3$ (roughly the size of an office typewriter).[101] It should be emphasized that such computational capabilities are required for real-time imaging, and most monitoring tasks in arms control verification do not require such rapid image analysis. The real-time capability has been mentioned here only to show that existing and projected capabilities are already more than adequate for many verification tasks.

Even when a good radar image is obtained it still requires skilled interpretation. Radar waves reflect differently from many surfaces than do light waves, and a given object can look very different on a radar image if its orientation changes with respect to the radar beam or if it is in motion. One peculiar property of synthetic aperture radar images is that a moving object appears on the image as stationary but in a different location, depending on its velocity relative to the aircraft or satellite taking the picture. Interpreting such images may be very tricky, and considerable effort is being put into classifying various kinds of radar image for more routine interpretation. One technique that promises vast improvements in object identification capability is to merge SAR images with visible and infra-red images of the same scene. Radar images are principally responsive to surface shapes and contours, while visible/infra-red images are more sensitive to surface chemistry. Combining the three types of image provides much more information than can be obtained from any single one, a good example of the synergism between different systems.[102]

Since radar signals are subject to several forms of attenuation and distortion there is also a need for image restoration and enhancement techniques similar to those used in visual photography. Such techniques exist already and are also well on their way to being digitized and automated.

All of these capabilities lead to the unmistakable conclusion that satellite cameras, sensors and radars will permit observation of objects and activities on the surface of the Earth in remarkable detail from altitudes of several hundred kilometres. The major limitation on all of this will remain the number of experienced, talented and reliable human monitors and interpreters. Such people will require training to a high standard of integrity and professionalism, and much of the success of any verification regime will depend on their alertness, skill and integrity. It is safe to assume that there already exists a large number of such people in the intelligence agencies of the USA, the USSR and other countries. The acquisition and retention of such people would be one of the highest priorities for any international satellite monitoring agency (see chapter 4).

## *VI. Seismology*

There is no technical area of verification which can even approach seismology for the volume of detailed analytical studies available in the open literature. Since the early 1950s there has been an active interest in detecting underground nuclear explosions for both intelligence and arms control reasons, and many states have sponsored active research programmes in this area. The United States alone has spent over $600 million on research and instrumentation related to verification of a nuclear test ban agreement[103] and, because of the high degree of international co-operation required for seismological research, most of the knowledge gained from this intensive programme is in the public domain.

A brief review such as this cannot hope to do justice to this interesting and still very active field. Only the basic concepts are introduced here along with an outline of the capabilities and limitations of current technology. For more detailed studies the reader is referred to any of a number of excellent recent reviews.[104]

There are a great many analogies between the basic principles of seismology and those of electromagnetic radiation which have been considered in previous sections. In both cases the fundamental phenomenon is a form of radiation which propagates for long distances in the form of waves. The radiation is emitted from a source, scattered or reflected off of objects in its path, absorbed or dispersed in transmission through a medium and detected by instruments which can record arrival times, frequency spectra, amplitudes and polarizations. In the case of seismology the source is some short-lived release of energy, such as an explosion or an earthquake; the medium of transmission is the interior or surface of the Earth; and the detectors are seismometers, instruments which respond to extremely small displacements of the Earth at their

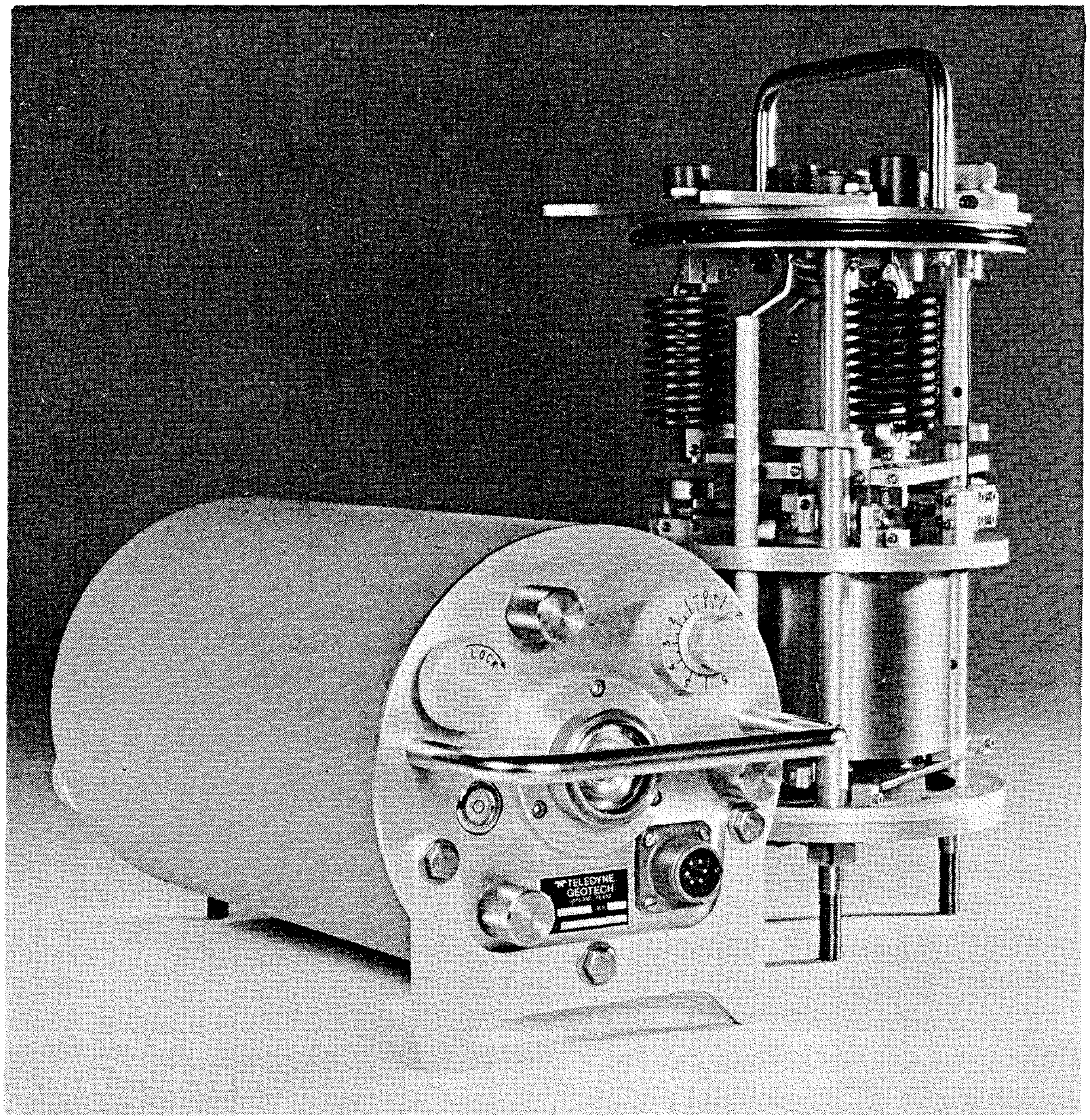


**Figure 23. Portable short-period seismometer**
The uncased instrument to the right shows the suspension of the oscillating mass by springs. This is one of the most widely used instruments for recording the short-period P-waves from earthquakes or underground explosions. Its maximum response is in the neighbourhood of 1 Hz, that is, a period of 1 second.

*Source*: Photo courtesy of Teledyne Geotech, Garland, TX, USA.

point of location. Seismometers can be in the form of individual instruments (analogous to a single infra-red sensor element or radar antenna) or arrays of instruments co-ordinated by electronic processors (analogous to phased-array radars).[105]

The basic design of a seismometer is very simple. A mass is hung from springs which are attached to a frame rigidly fixed to the Earth—preferably on or in solid rock. When the Earth moves the mass is set into movement by the springs and this movement is converted to an electrical signal by a magnet surrounding the mass (see figure 23). The sensitivity of modern seismometers is remarkable. The motion of the Earth in all but the most violent seismic disturbances is imperceptible to human beings, but useful information can be extracted from motions with amplitudes as small as or even smaller than one nanometre ($10^{-9}$ m), comparable to the diameter of a single atom. However, even this is not good enough for the detection of very small events at long distances, so new instruments are being designed capable of faithfully recording displacements 10 000 times smaller than this, that is, between $10^{-14}$ and $10^{-13}$ m.[106] Such instruments must be located where seismic 'noise' (the random disturbances caused by winds, waves, human activity, etc.) is at very low levels. There is a continuing search for such areas, with a major focus in the United States on placing sensitive seismometers in deep 'bore-holes' in the ocean floor.[107]

While there are many similarities between electromagnetic and seismic waves there are also some very importance differences. One of the most fundamental results from the different mechanisms by which the waves are excited and the various media through which they propagate. While electromagnetic waves come in many 'colours' (i.e., frequencies) there is really only one basic wave type involved. Seismic waves on the other hand come in many forms as well as in a wide range of frequencies. There are two types of 'body' wave (i.e., those which pass through the body of the Earth), one which involves compressional motion (P waves) and the other transverse or 'shearing' motion (S waves). Then there are two other waves, which travel only over the surface of the Earth, called Rayleigh waves and Love waves. These are distinguished by the differing motions (vertical and horizontal respectively) of elements of the Earth's surface as the wave passes by (see figures 24 and 25).

Different seismic waves travel on different paths, at different speeds, have different characteristic frequencies and wavelengths, and are absorbed and scattered with different strengths. This means that the signal that reaches a detector at some distance from a source is extremely complex, consisting of several 'phases' which correspond to the arrivals of different types of wave (see figure 26). The nature of these phases depends strongly on the distance between the source and the detector, and seismologists consider the problems of observation and analysis to be quite different at 'regional' (i.e., less than 2 000 km) and 'teleseismic' (i.e., greater than 2 000 km) distances (see figure 24). Most of the research effort in seismological identification since 1960 has been carried out at teleseismic ranges in order to develop 'national technical means' of verification of underground nuclear weapon testing limitations. However, in recent years a number of proposals for extensive seismological networks have revived interest in using regional data, and research at these distances is now quite active.[108] Most seismologists agree that some reliance on regional data would greatly enhance the ability to monitor a comprehensive nuclear test ban, but there are differences of opinion as to how much is needed.[109]

The capabilities needed in a seismological monitoring system depend on the nature of the treaty which must be verified. First, the system must be capable of distinguishing between earthquakes and explosions above some specified level of energy release (i.e., yield). If the treaty is a threshold type, which prohibits explosions only above a certain maximum yield, then the monitoring network must be capable of effective location and discrimination at or above this level as well as able to provide reliable estimates of explosion yields. If the treaty is a comprehensive one prohibiting *all* nuclear explosions, then the system must be able to locate and identify nuclear explosions at such low yields that any explosions smaller than this limit are agreed by all parties to be militarily and politically insignificant.

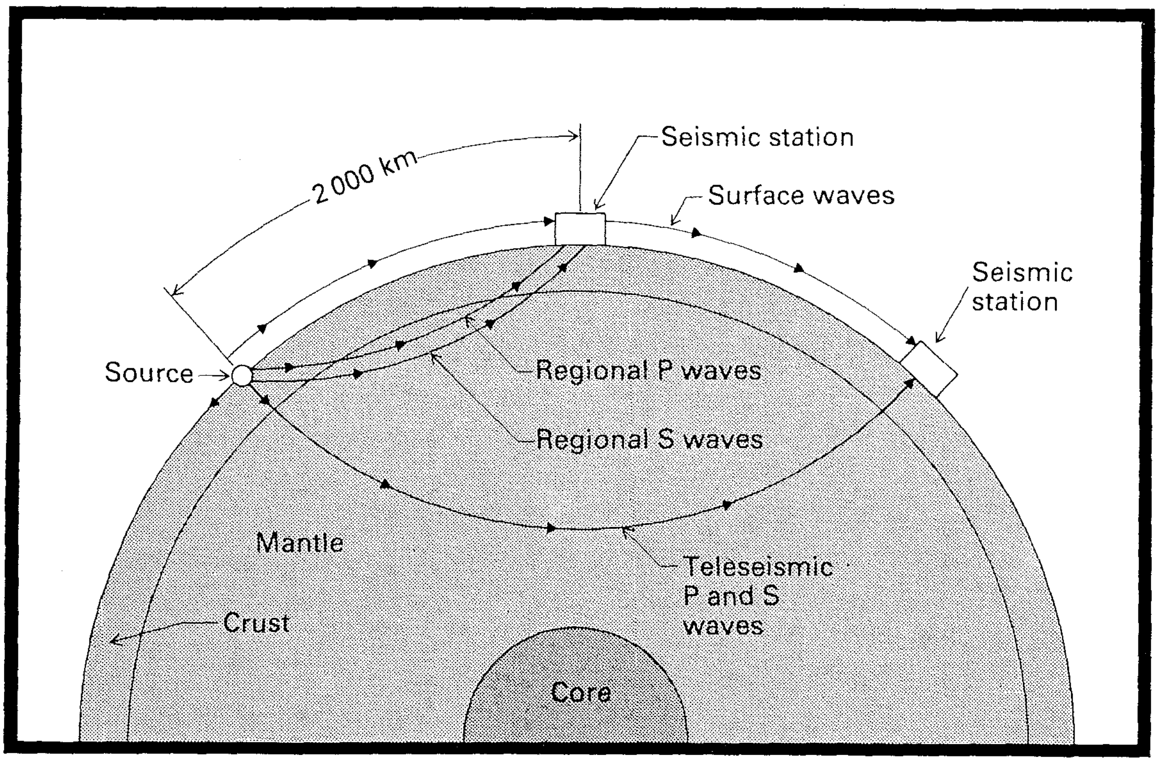


**Figure 24. Seismic wave paths**

The figure indicates the paths of both body and surface waves at regional and teleseismic distances. The bending of the body waves is a result of the variations in density with depth in the Earth's mantle. This effect is analogous to the bending of light rays as they pass through a medium of varying density. Note that the distances and sizes on the drawing are not to proper scale.

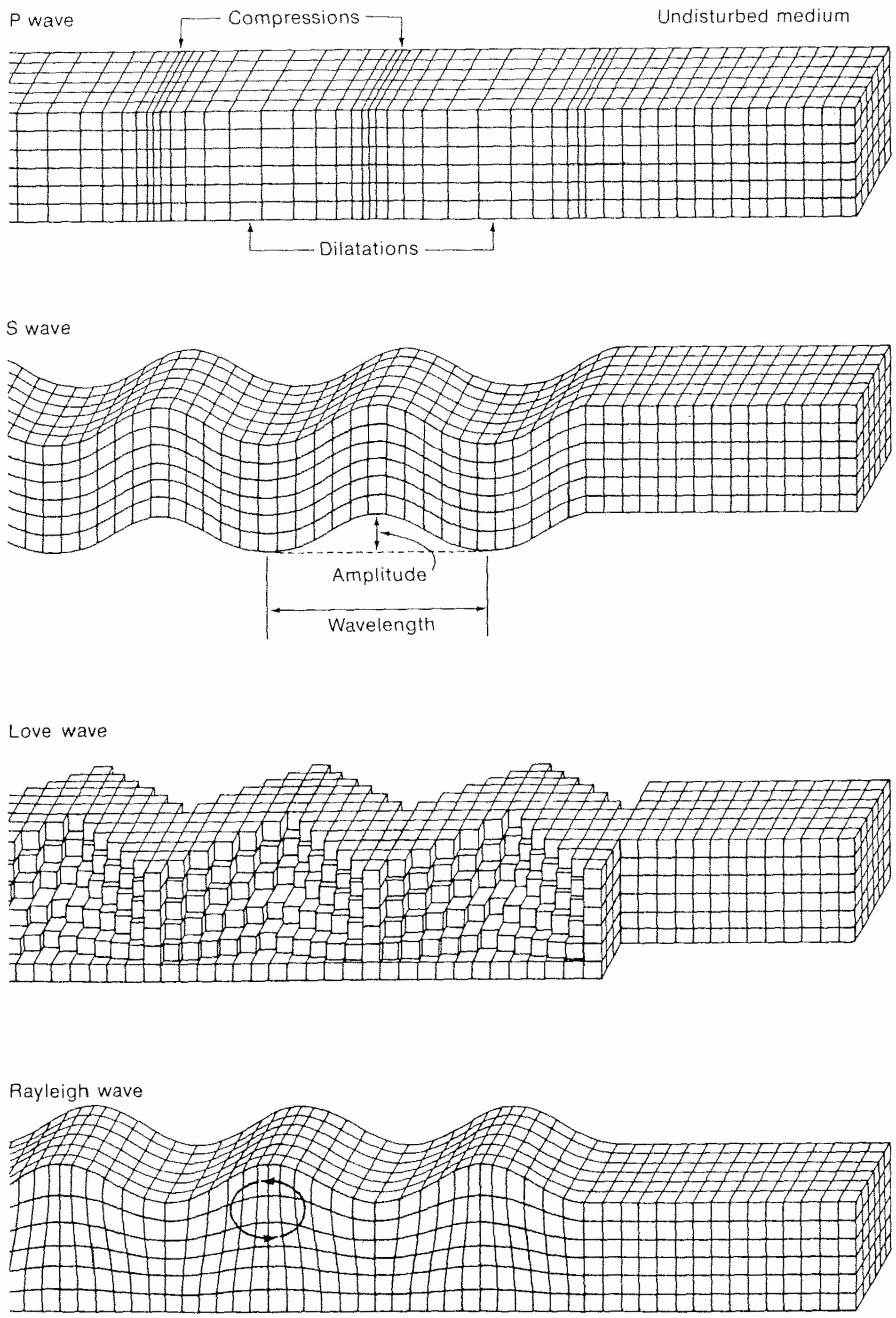


**Figure 25. Seismic wave types**
Seismic waves are differentiated by the medium through which they propagate (P and S waves through the body of the Earth and Love and Rayleigh waves over the surface) and by the relationship between particle motion and wave propagation direction (longitudinal: P waves; transverse: S and Love waves; and vertical/longitudinal: Rayleigh waves). A seismic station capable of detecting all of these waves must have six individual instruments: three short- and three long-period seismometers, with each set oriented in three perpendicular directions.

*Source*: Bolt, B. A., *Nuclear Explosions and Earthquakes. The Parted Veil* (W. H. Freeman, San Francisco, 1976), p. 49, figure 3.5. W. H. Freeman and Company, Copyright 1976.

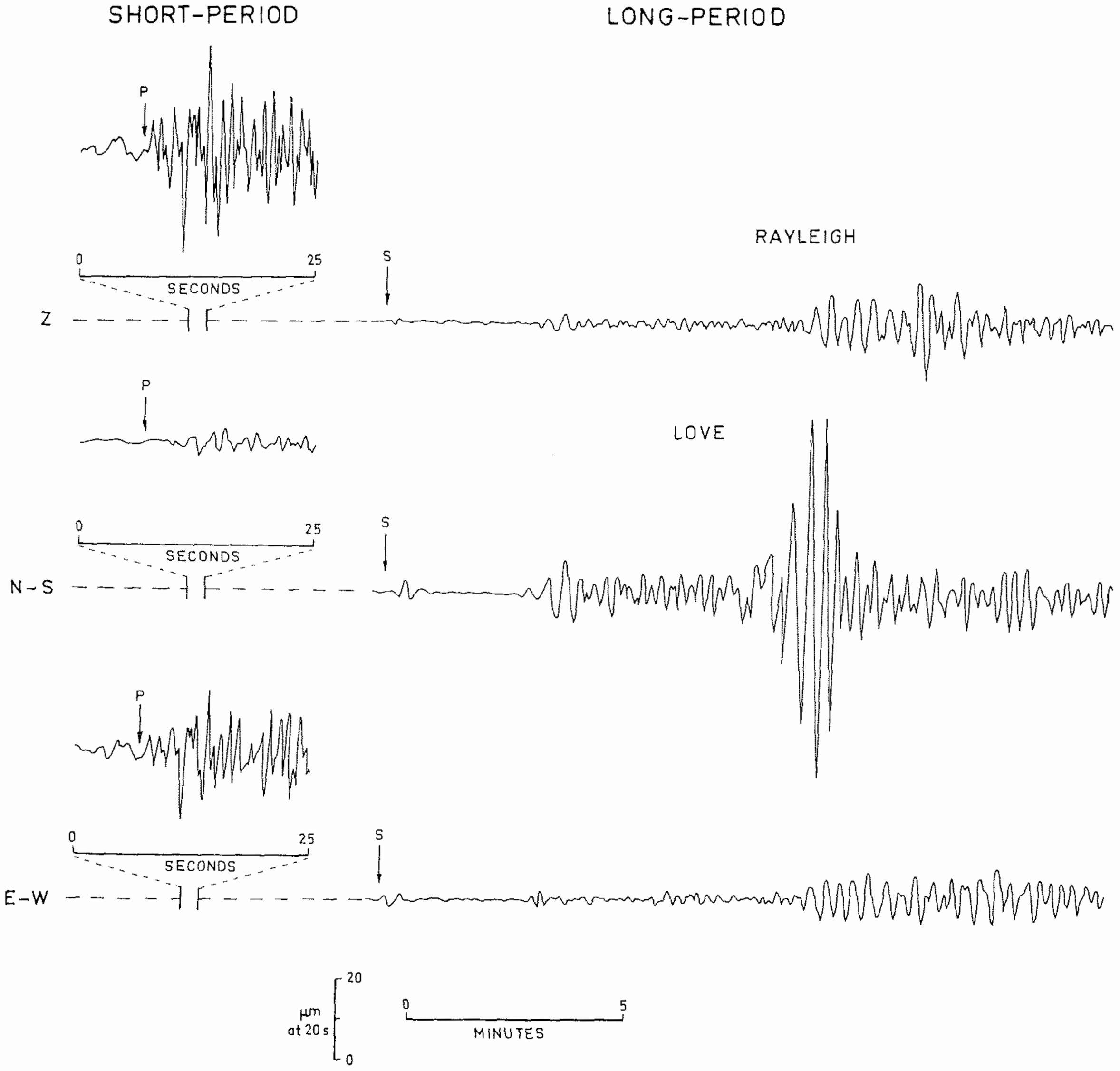


**Figure 26. A typical seismic record**

The three records display Earth motion in three different dimensions. The top record (labelled Z) shows vertical motion and therefore consists almost entirely of the P-wave (short-period) and Rayleigh-wave (long-period) phases. Note the approximately 1-second period of the P-wave phase shown with an expanded time-scale. The horizontal motion records (N–S and E–W) show smaller P-wave and larger S-wave amplitudes as well as the Love-wave phase. Note that the average period of the Love wave is roughly 15–20 seconds and its maximum peak-to-peak amplitude is about 90 $\mu$m.

*Source*: Dahlman, O. and Israelson, H., *Monitoring Underground Nuclear Explosions* (Elsevier, Amsterdam, 1977), p. 60, figure 4.8. Reproduced by permission of Elsevier Science Publishers, Amsterdam.

## Detection and identification

There are two particular phases which are most often used in detecting and identifying nuclear explosions. At teleseismic distances the important phases are the initial P phase which travels through the Earth at a speed of from 8–12 km/s and has frequencies in the neighbourhood of 1 Hz (a period of 1 second), and a Rayleigh-type surface wave which travels at a speed of 3–4 km/s and has frequencies around 0.05 Hz (a period of 20 seconds).[110] The two frequencies mentioned here are the ones for which most seismometers are optimized, because seismic noise levels are significantly lower at these frequencies than at intermediate ones.[111]

Using two phases at different frequencies to discriminate between different kinds of seismic event is analogous to using multi-spectral information in photography to distinguish different objects which all look the same if no frequency separations are made (see p. 30). This is, in essence, the fundamental principle underlying the most successful and most commonly employed earthquake-explosion 'discriminant': the $m_b : M_s$ criterion.

The symbols $m_b$ and $M_s$ refers to the 'magnitudes' of the body-wave and surface-wave phases respectively. Each magnitude is a measure of the local velocity of Earth movements and is determined by first dividing the amplitude of the motion by the period, then taking the logarithm and finally applying corrections for the distance between the detector and the source as well as any biases associated with the equipment used or the location of the seismometer.[112] The magnitude of a particular phase is closely related to the amount of seismic energy in that phase, so to compare the body- and surface-wave magnitudes is equivalent to comparing the relative amounts of energy put into these different forms of ground motion by the source.

Explosions and earthquakes are very different phenomena. An explosion takes place in a very short time in a relatively small region and imparts a strong outward compressional impulse to the Earth in all directions simultaneously. On the other hand an earthquake is a more slowly developing phenomenon which usually involves the release of seismic stresses over a large volume of the Earth and which has a highly directional, that is unsymmetrical, pattern of seismic radiation (see figure 27). While an explosion will produce almost exclusively compressional waves, an earthquake will produce both compressional and shear waves. The latter when they reach the Earth's surface are much more effective in producing surface waves, so the fraction of an earthquake's energy which ends up in surface waves is generally quite a bit larger than for an explosion.

The time during which an event takes place determines the frequency spectrum of the radiation from the event, with short events creating higher frequency radiation than long ones. Since P waves have much higher frequencies than Rayleigh waves, more P waves can be expected from explosions and a much greater generation of low-frequency Rayleigh waves can be expected from earthquakes. The combination of the above effects leads in most cases to clearly distinguishable seismograms for explosions and earthquakes (see figure 28).

The standard procedure for determining whether a given signal came from an earthquake or an explosion is to compute the magnitudes $m_b$ and $M_s$ of the short-period (P) and long-period (Rayleigh) phases respectively and then to display the relationship between the two magnitudes on a graph (see figure 29).

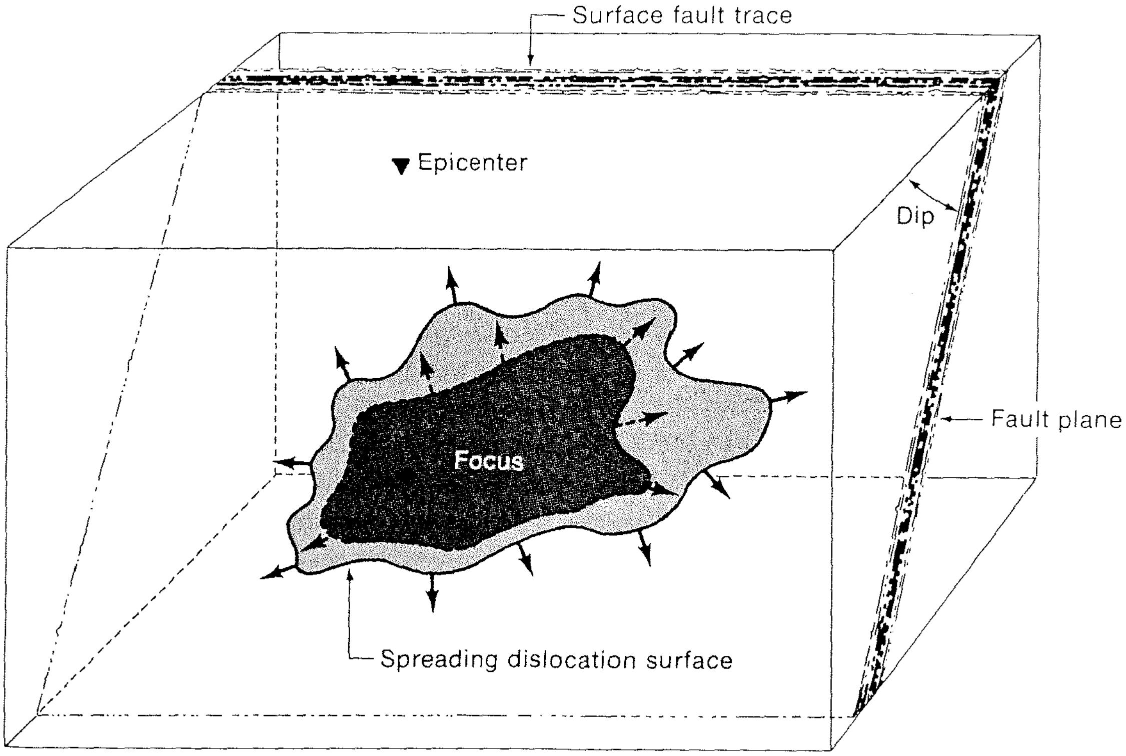


**Figure 27. Earthquake mechanism**
A three-dimensional section of the Earth's crust showing a rupture spreading out from the focus of the earthquake along the fault plane. The release of seismic energy is produced by the relative slippage of the two sides of the fault plane, a slippage which begins with the release of strain at the focus and spreads rapidly outwards. Note the highly non-symmetrical nature of the disturbance.

*Source*: Bolt, B. A., *Nuclear Explosions and Earthquakes. The Parted Veil* (W. H. Freeman, San Francisco, 1976), p. 68, figure 4.3. W. H. Freeman and Company, Copyright 1976.

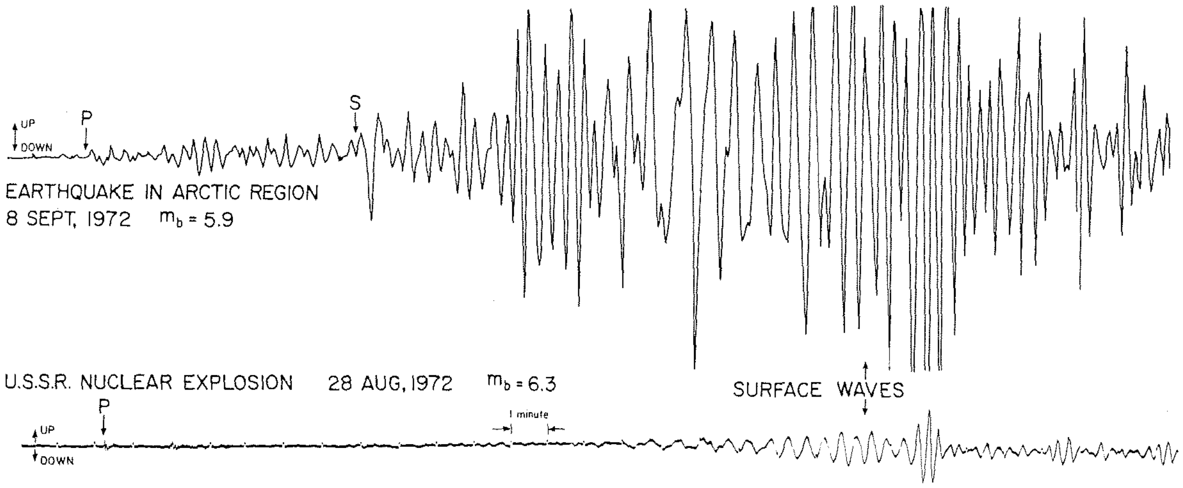


**Figure 28. Earthquake and explosion seismograms**
Note that the P-wave magnitudes of the two events are roughly similar but that the surface-wave magnitude of the earthquake is dramatically larger than that of the explosion. In general, shallow earthquakes couple energy far more strongly into surface waves than do explosions.

*Source*: Courtesy of Lynn R. Sykes, Lamont Doherty Geological Observatory, Columbia University, NY, USA.

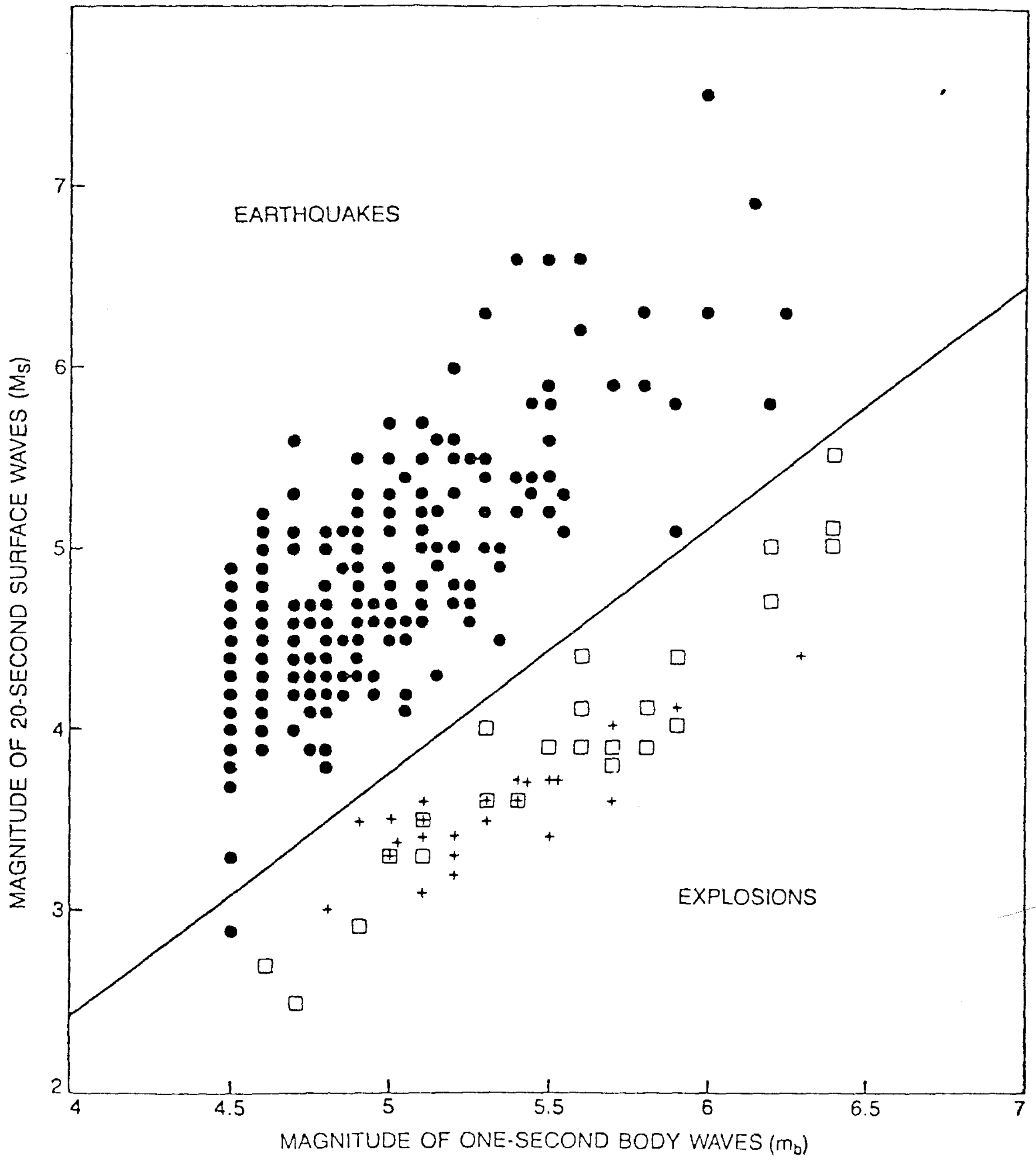


**Figure 29. Application of $m_b:M_s$ discriminant**

The black dots represent a population of 383 earthquakes with focal depths less than 30 km recorded world-wide over a six-month interval. (There are not 383 dots because many events had the same magnitudes). The squares designate explosions in the USA and the crosses explosions in the USSR. The one earthquake which falls on the explosion side of the line was a very weak event which occurred in a region of the south-west Pacific Ocean poorly covered by the existing network of seismological stations.

*Source*: Sykes, L. R. and Evernden, J. F., 'The verification of a comprehensive nuclear test ban', *Scientific American*, Vol. 247, No. 4, October 1982, p. 35. Copyright 1982 by Scientific American, Inc. All rights reserved.

It has been found from many studies that an event can be identified with high confidence by its location on this graph.

The $m_b : M_s$ discriminant is the most effective one yet devised for use at teleseismic distances, and there is a similar discriminant which employs analogous but different phases at regional distances and appears promising.[113] But these discriminants are by no means perfect, and it is highly unlikely that any single discriminant will be found which can distinguish earthquakes from explosions with 100 per cent confidence. The geological medium through which seismic waves travel is far too complex to allow for such hopes.

The answer to this problem is to use several analytical techniques and discriminants to reduce the uncertainty in ambiguous events. For example, some earthquakes have produced $m_b : M_s$ values which made them look like explosions because of unusually clear transmission of P waves from the source to the detector.[114] But when the depth of the sources of these waves was measured from other characteristics of the signal,[115] they were found to be at least 20 km deep, putting them well below the depth at which nuclear explosives can be placed. In fact the deepest known nuclear explosion was conducted at a depth of 2 km,[116] and the limits of modern drilling technology are about 12 km.[117] So any source located with high confidence at a depth below 10 km can be safely identified as an earthquake. This criterion alone can rule out a substantial fraction of 'false alarms'.[118]

Another useful discriminant is the location of the source. With good data from a few seismological stations a source can be located to an accuracy of about 10–20 km.[119] If the location is found to be under the ocean, then the possibility of it being an explosion can be effectively ruled out, since any attempt to conduct a nuclear test under the ocean would be easily detectable by a number of other means. Since over 90 per cent of all earthquakes occur under oceans and/or at depths greater than 30 km only a relatively small number of earthquakes remain to produce false alarms.[120]

The combination of location, depth and the $m_b : M_s$ discriminant is a powerful method for distinguishing earthquakes from explosions. Having applied this method to a very large data sample, one group of analysts summarized their results as follows: "We know of no Eurasian earthquake with 1 second P-wave magnitude of 4 or more of the past 20 years whose waves are classified as those of an explosion . . . Furthermore, to our knowledge not one of several hundred underground nuclear explosions set off in the same period radiated seismic waves that could be mistaken for those of an earthquake."[121]

## Yield estimates

It is quite a bit easier to determine whether or not an explosion has taken place than it is to get a reliable estimate of its yield. There is an enormous variability in the magnitude values recorded at different seismic stations for a single explosion (see figure 30).[122] The variability can be illustrated by computing

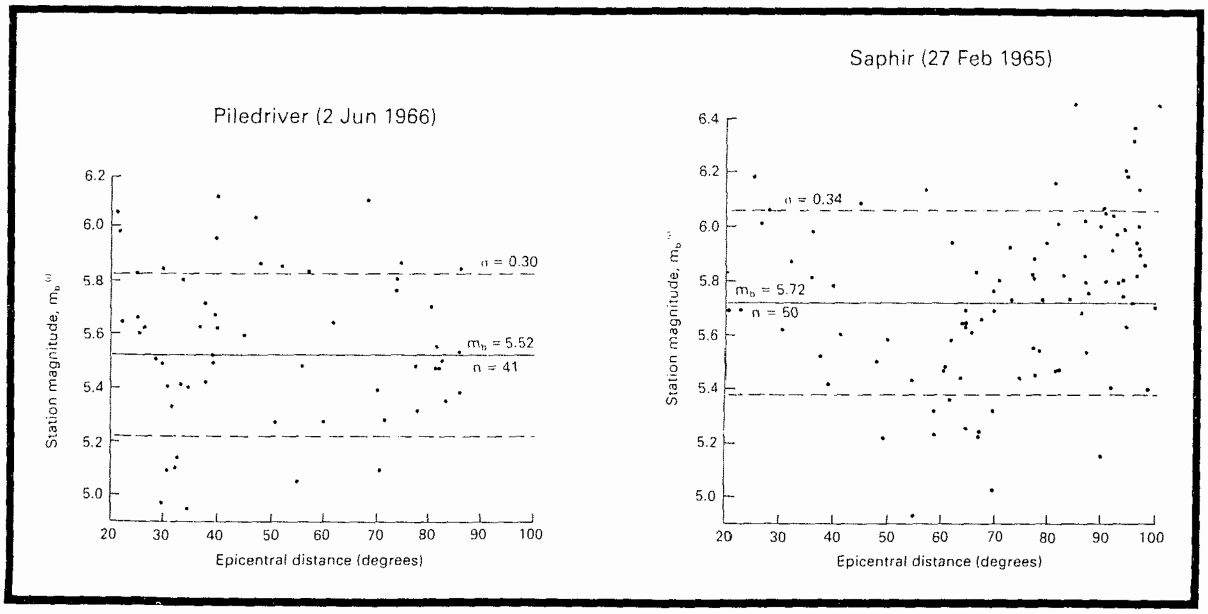


**Figure 30. Variation in body-wave magnitudes**
The $m_b$ values recorded at a large number of seismological stations (41 and 50 respectively) for two test explosions are shown plotted against the distance of the station from the test site. Note that the distance is measured in degrees as is customary in seismology (10 degrees represents a distance of about 1 100 km on the Earth's surface). The average magnitudes for the two explosions are 5.52 and 5.72 and both show standard deviations of at least 0.3, implying an uncertainty in yield estimate of at least a factor of 2 (see text).

*Source*: Based on Bache, T. C., 'Estimating the yield of underground nuclear explosions', *Bulletin of the Seismological Society of America*, Vol. 72, No. 6, Part B, December 1982, p. S113, figure 1.

the difference in yield estimates which would result from using the lowest and highest recorded magnitudes for the Saphir explosion in figure 30. Using a standard average formula relating yield Y, in kilotons, to body-wave magnitude[123]

$$m_b = 3.8 + 0.75 \log_{10} Y$$

and the two extreme magnitudes of 5.1 and 6.4 gives a range of yield estimates from 54 to 2 900 kilotons. The correct value was 120 kt. The same formula can be used to show that an error of only 0.1 in the value of $m_b$ corresponds to an error of 30 per cent in the yield estimate. The sensitivity to small errors illustrates the great danger in using magnitude estimates from only one or a few stations to estimate yields as well as the need for large amounts of data to get yield values which are correct even within error limits of 100 per cent, that is, a magnitude estimate valid within ±0.3 or so.

The scatter in $m_b$ data is only one of many problems facing the yield estimator. Explosions carried out in different geological media will generally be more or less 'decoupled' from the surrounding medium, that is, they will transfer a larger or smaller fraction of their released energy into seismic waves.[124] (The formula used above assumed a well-coupled explosion in hard rock.) For example, explosions in dry alluvium (a soft, porous sedimentary medium) can give magnitudes from 0.5 to 1.0 lower than for the same yield explosion in hard rock. An explosion carried out in a large underground cavity would be even more decoupled, leading to reductions in apparent yield by as much as a factor of 100 (e.g., from 100 kt to 1 kt). It is also known that the yield–magnitude relationship for a given test site is affected by the tectonic history of that site. Recent (on a geological time-scale) tectonic activity causes a site to produce lower magnitudes for a given yield than a site which has been free of such activity for hundreds of millions of years.[125] The US test site in Nevada is an example of the former type, while the Soviet eastern Kazakh test site is one of the latter. The different site properties produce a systematic bias in any attempt to apply Nevada test-site data to estimating the yields of Soviet test explosions. The assumed value of this bias is a crucial factor in evaluating the Reagan Administration charges that the Soviet Union has violated the Threshold Test Ban Treaty by testing weapons with yields over 150 kt (see chapter 4). One recent study of this problem employs surface-wave magnitudes which are subject to less variation in bias to establish that Soviet tests have in fact not exceeded the 150 kt limit.[126]

One more simple application of the average magnitude–yield formula shows that a value of $m_b = 4.0$ corresponds to a yield in hard rock of about 2 kt. On the basis of this value and the quotation on p. 70, a highly reliable existing capability to distinguish between earthquakes and any explosion with a yield greater than 2 kt in hard rock can be assumed. The many estimates of this limit in the literature range from 1 to 5 kt, with most tending towards the lower end of the range.

The possibility that explosions in this yield range or even larger might be concealed by conducting them in large cavities (see above) has for many years been the most commonly mentioned means by which a party to a ban on underground tests could evade detection.[127] It is true that the apparent yield of such a cavity-decoupled explosion is greatly reduced, possibly by a factor of 100 or more, when measured on the usual short-period seismometers optimized to record signals in the neighbourhood of 1 Hz. However, recent studies have shown that the decoupling effect is dramatically reduced at higher frequencies.[128] As has already been noted, explosions are far better generators of high frequencies than are earthquakes, and there is also mounting evidence that the higher frequency seismic waves propagate for much longer distances than had previously been believed. Finally, seismic noise is greatly reduced at frequencies of 30 Hz or higher, allowing for excellent signal-to-noise ratios for even relatively weak high-frequency signals.[129]

This new information has been used to compute the effectiveness of a network of high-frequency seismometers in detecting decoupled explosions. The conclusion of this analysis is that even fully decoupled explosions of fractional kiloton yields are identifiable and therefore "that all discussions of the feasibility and utility for evasion via large cavity decoupling are passé".[130]

### Seismic image processing

Another similarity between seismological and optical or radar observations is worth examining in some detail: the need for image processing. A seismological 'image' consists of the recorded seismometer readings at all stations which received signals from the event. On most present seismographs these readings are still in analog form, that is they are recorded as complex waveforms drawn on paper by a chart recorder. (See for example figures 26 and 28). But rapid technological change, again led by advances in microelecronic and computer technology, is leading to more and more use of digital seismographs. These devices record the seismic signal directly on magnetic tape or into a computer memory in the form of binary numbers, exactly like photographic pixels.

Once the image is recorded, the problem facing the seismologist sounds remarkably similar to that facing the photo-interpreter: "complexities of the earth strongly affect the seismic signal, thus presenting us with a blurred and distorted observational image of the source. To improve this image we have to remove complicating wave propagation and recording effects".[131]

One of the major sources of image degradation is seismic noise. This can be minimized by placing seismometers deep in solid rock formations and by using electronic filters of various types. It can also be reduced by deploying an array of seismometers at a given location and combining the signals from all elements in the array (see figure 31). This technique was originally motivated by the expectation that seismic waves arriving from a source thousands of kilometres away would be coherent over distances of the order of 100 km at the location of the detector. 'Coherence' in this sense means that all of the detectors are excited in the same way, or in a way that is analytically predictable once the distance and direction to the source are known. On the other hand seismic noise is quite incoherent over distances of 1 km or more, so that noise signals from different elements of the array will have no predictable or constant relationship to each other. When the signals from many elements of an array are combined in the appropriate way (in some cases a simple sum might be sufficient) the true signal will be enhanced relative to the noise. Theoretically the enhancement should be proportional to the square root of the number of elements in the array, so a 25-element array should increase the signal-to-noise ratio by a factor of 5.

These considerations led the United States to build three very large arrays in Alaska (ALPA), Montana (LASA) and Norway (NORSAR) under the so-called Vela Uniform Programme.[132] These arrays had diameters of between 100 and 200 km and were made up of hundreds of individual seismometers, quite analogous to the phased-array radars discussed above. When they are accompanied by appropriate data transmission and processing capabilities, seismological arrays can be 'steered' very much in the manner of a phased-array radar in order to be optimally sensitive to seismic waves coming

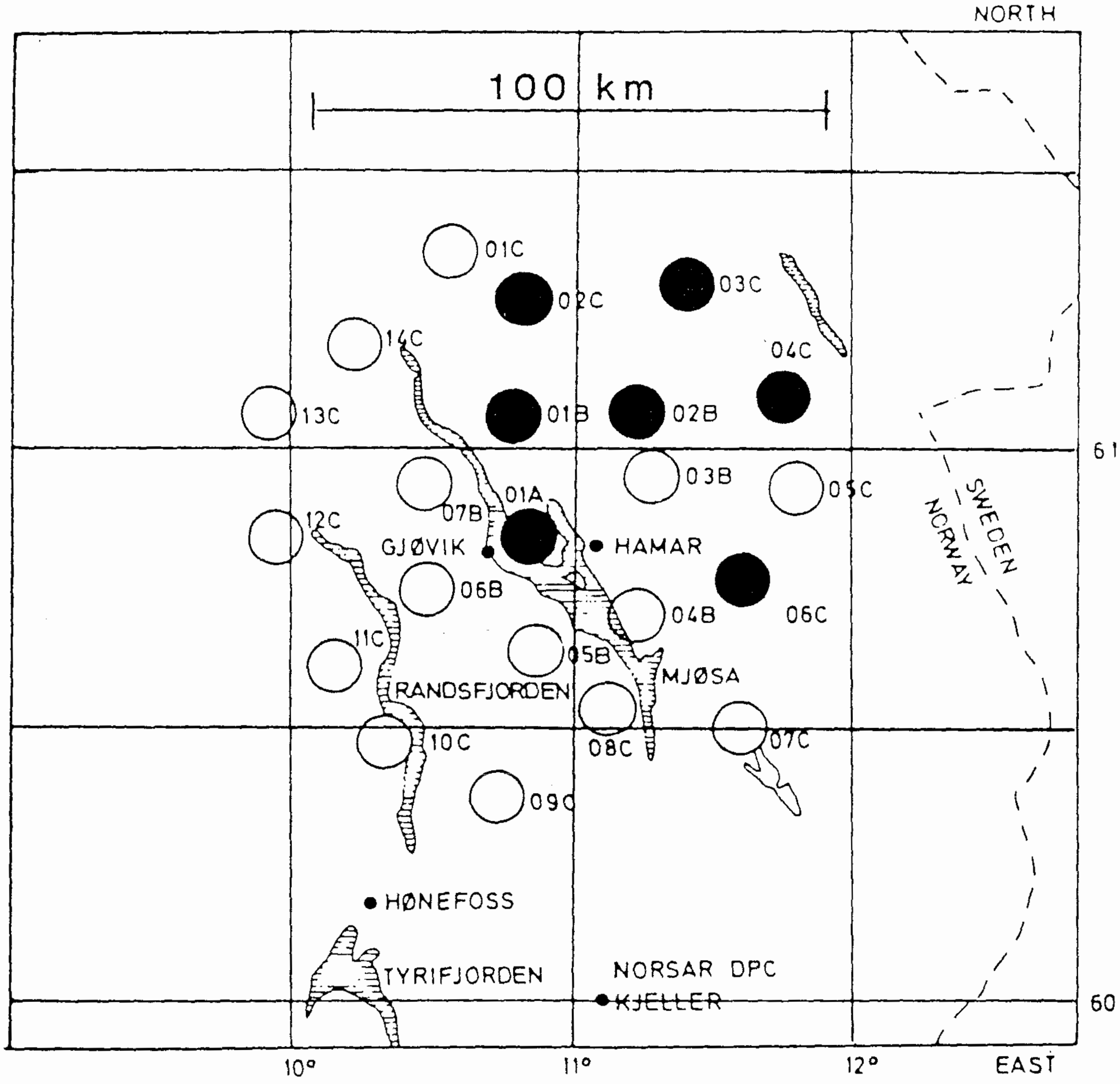


**Figure 31. The NORSAR array**
The original array was made up of 22 sub-stations arranged in a roughly circular pattern about 100 km in diameter. The current array uses only 7 sub-stations and has a diameter of about 60 km. Note the similarity of this array of seismometers to the array of radiating elements in the radars of figures 10 and 11. The principle of beam forming in a seismic array is precisely the same as that for a phased-array radar.

*Source*: Courtesy of NORSAR.

from particular directions (called 'beam forming') or with particular velocities (called 'velocity filtering'). These capabilities allow, at least in principle, for seismic arrays to 'scan' the Earth, much as a phased-array radar scans the skies. Naturally the seismological array does not move; the scanning is done electronically by changing the time-delay relationships among the detectors.[133]

The actual performance of the three very large arrays turned out to be less than was hoped for, largely because the distances over which teleseismic P waves exhibit coherence turned out to be smaller than anticipated.[134] Both the LASA and ALPA arrays have been shut down, while the NORSAR array has been reduced in size from 22 subarrays to 7 (see figure 31). Such smaller arrays are still a considerable improvement on individual seismometers.

Distortions of the seismic image are also caused by absorption and scattering of seismic waves along the path from source to detector and by the specific response features of the seismometer. For example, waves with different frequencies are degraded at different rates as they move through the Earth, and seismometers respond differently to signals at different frequencies. Both of these effects can in principle be modelled mathematically, and these models can be used in an 'image restoration' process quite analogous to those used in photography to remove atmospheric and optical distortions. In seismology the process of applying these corrections to the signal is called 'deconvolving', and many useful features of a seismic signal can be revealed by successful deconvolution. The major difficulty in applying this method is the lack of precise models for seismic-wave propagation through the Earth. Much research remains to be done to improve such models, and it is evident that the Earth will never be as 'transparent' as the atmosphere.

There are many other image restoration techniques which are in various stages of development and application. The key to effective seismic-image processing is the same as for optical and radar images: more, faster and cheaper digital computers. The data processing demands on a world-wide network of seismometers in continuous operation will be at least as severe as for satellite photography. And, since the number of human interpreters will always be far too small to examine all these data, there must be a considerable amount of automatic processing, that is, artificial intelligence, which can make preliminary judgements about the significance of events and leave only the most important or ambiguous for human interpretation. While such capabilities are still far from realization, much progress is being made and much more is expected as the result of current research.[135]

## *VII. Nuclear explosion detectors*

A nuclear explosion in the atmosphere or in outer space is an exceptionally violent event which provides ample evidence of its occurrence. The essence of the explosion is the sudden release of an enormous quantity of energy into a very small volume. For example, a 10 kt explosion will in the first millionth of a second or so release the energy equivalent of 10 000 tons of TNT into a volume no bigger than a grapefruit. This creates extremely high temperatures—at least 10 million degrees Celsius—and as a result of the Wien displacement law (see table 3, p. 29) the average wavelength of the radiation is extremely short, characteristic of X-rays, a form of electromagnetic radiation whose characteristic energies are from 1 000 to 100 000 times as large as those of visible light. These X-rays account for more than half of the total energy released by the explosion, with most of the rest being in the form of fast-moving fragments of the original bomb materials.[136] Nuclear explosions also produce large numbers of an even more energetic form of radiation, called 'gamma-rays'. These can have energies more than one million times as great as visible light. As column 4 of table 3 makes clear, the intensity of the X-rays emitted (that is, the relative 'brightness' of the fireball) is unimaginably large. No comparisons to such intensities exist in human experience, and even one of the most graphic attempts, *Brighter than a Thousand Suns*,[137] is still 10 orders of magnitude too small.

The detection of X- and gamma-rays requires a very different type of detector from the visible and infra-red sensors considered in sections II and III. However, such detectors have existed for many years and have been used to monitor nuclear explosions at least since the early 1960s when both the USA and the Soviet Union were confident that they could verify a ban on nuclear tests in the atmosphere and in space. This confidence led to the signing of the Partial Test Ban Treaty in August 1963.

The most common form of X-ray and gamma-ray detector (see figure 32) uses a material called a 'scintillator' which converts the energy of the incoming photon into a pulse of visible light. When an X- or gamma-photon enters the scintillating material it can cause one or more electrons to be ejected from atoms in the crystal. As these electrons recombine with the positively charged

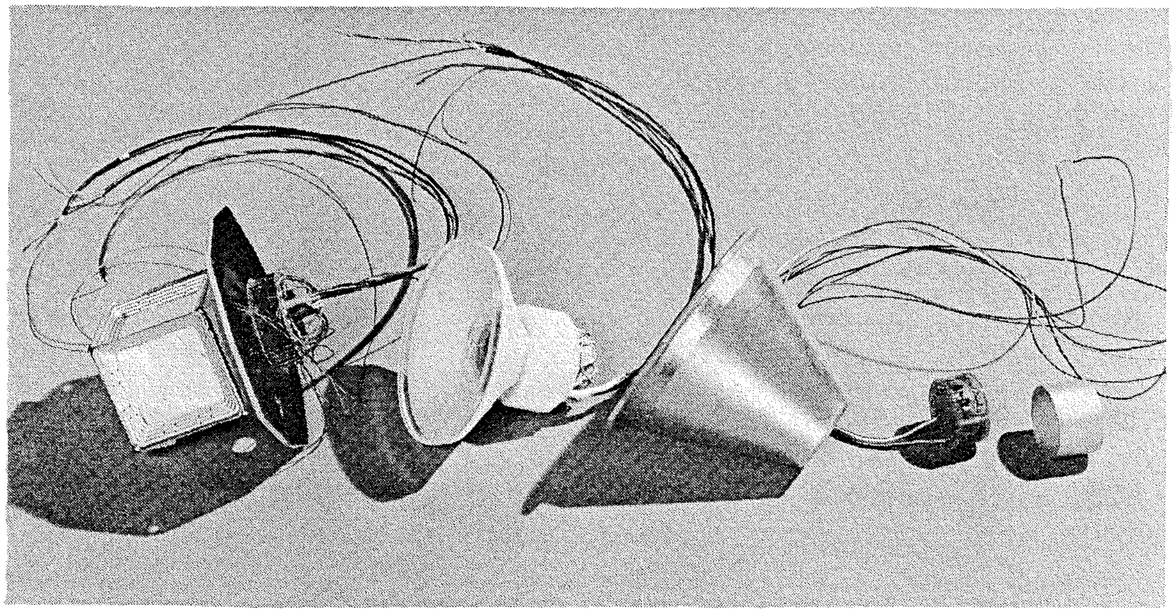


**Figure 32. The M4 X-ray detector**
X-rays enter the cubical box at the left of the detector and interact with the atoms of a caesium iodide (CsI) 'scintillator' producing flashes of light which are then converted to electrical signals by a photomultiplier tube at the base of the cube. The conical shaped scintillator and photomultiplier in the centre form a so-called 'guard' detector whose function is to identify and reject high-energy cosmic ray events which also trigger the main detector and could produce false alarms.

*Source*: Photo courtesy of Los Alamos National Laboratory.

ions from which they were detached, light is emitted. This light is captured by a 'photomultiplier' tube which converts the light energy into an electrical voltage pulse whose magnitude is proportional to the energy delivered by the original X- or gamma-ray photons. The voltage pulses from the photomultiplier can be digitized and stored for later transmission to Earth if the detector is in a satellite, or they can be stored on magnetic tape for computer processing or observed directly on a video screen if the instrument is based on Earth. Scintillation counters can be quite small, light and portable, and they require very little electrical power for their operation.

If the explosion takes place above the atmosphere in the near-perfect vacuum of space these X-rays move outward from the explosion in all directions at the speed of light. Because the total radiated energy is so large, the intensity of X-rays remains large even at great distances from the explosion. This enables a single satellite, such as the US Vela satellite, to detect the X-rays from nuclear explosions at distances comparable to the diameter of the Earth's orbit about the Sun, approximately 300 million kilometres.[138]

If the explosion occurs in the atmosphere, the X-rays are absorbed within a few metres of the centre of the explosion, causing rapid heating and com-

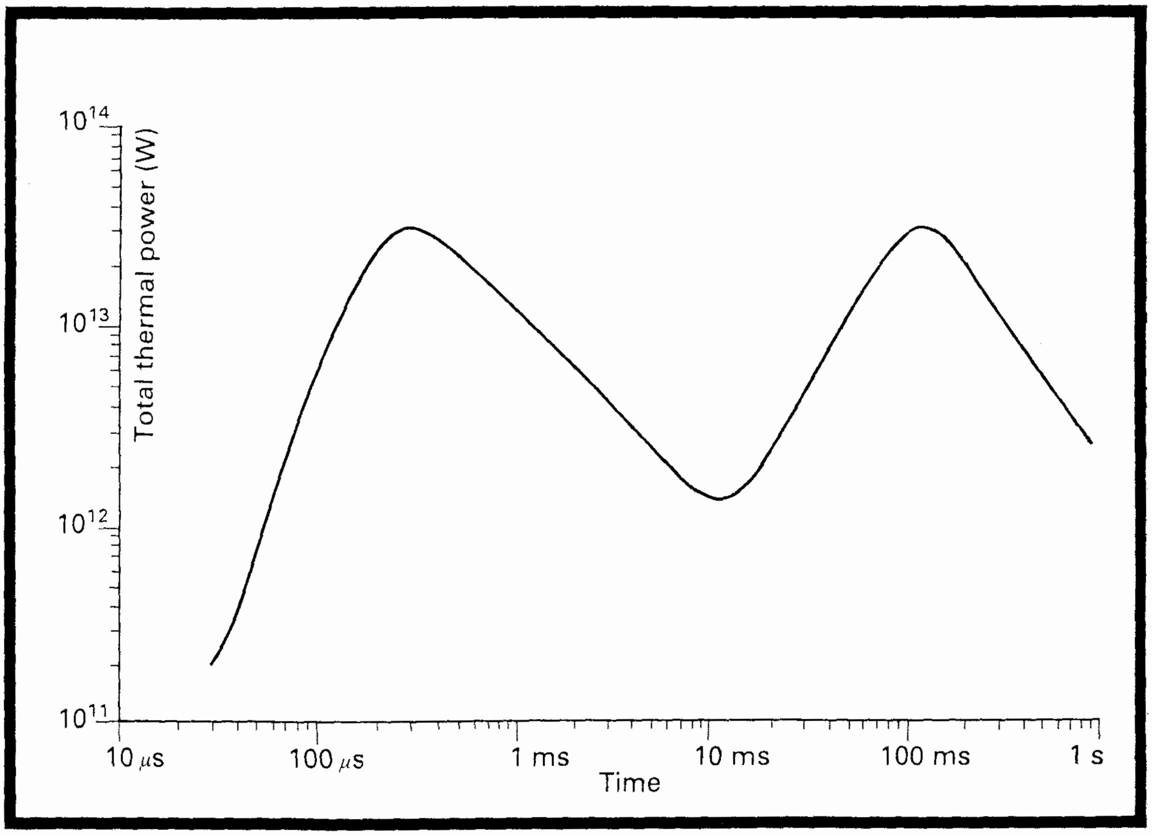


**Figure 33. The double light pulse from a nuclear explosion**
The general shape of this double pulse is the same for all nuclear explosions, and the yield of an explosion can be estimated quite accurately from measurements of the time intervals between the two maxima and the minimum.

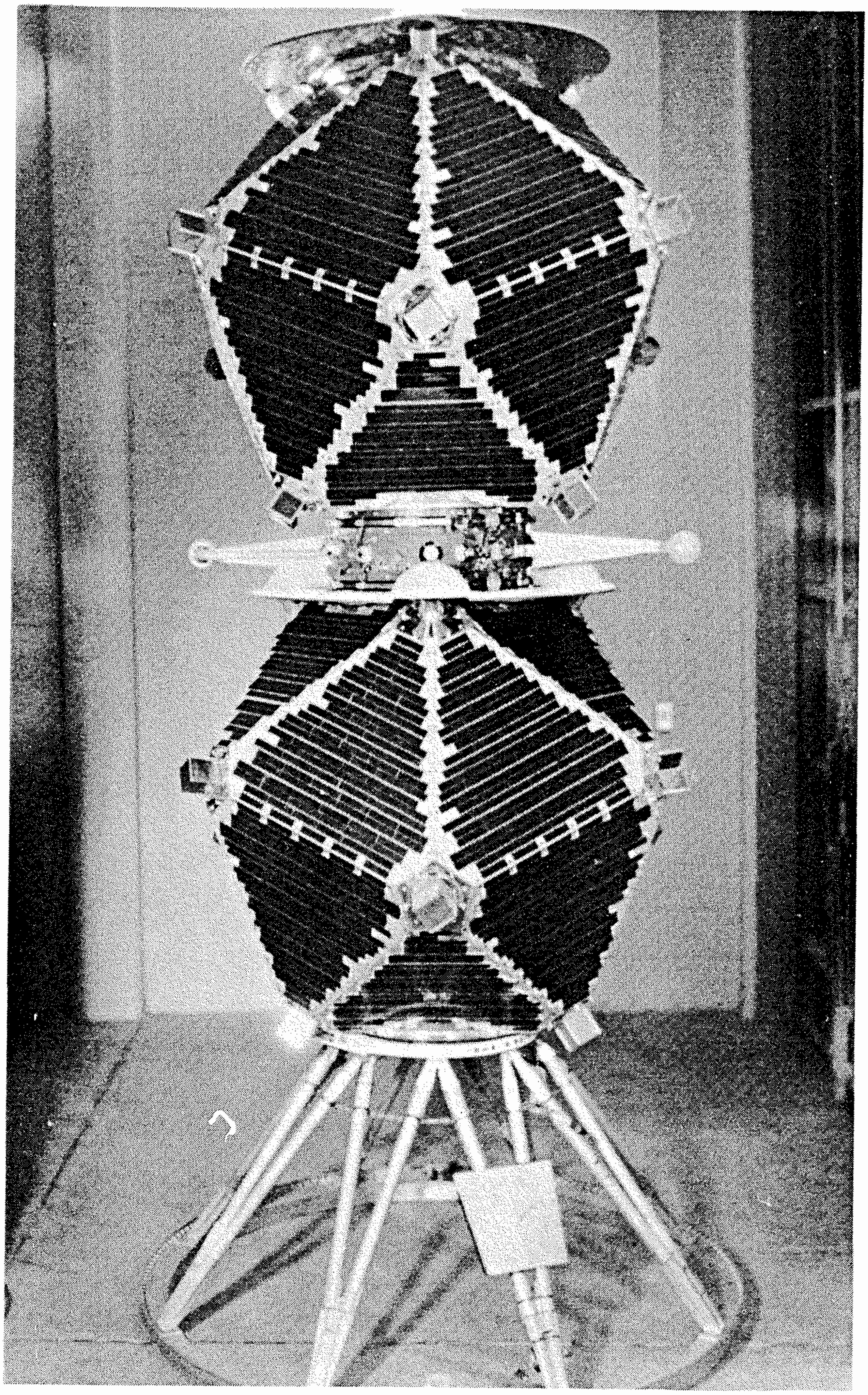


**Figure 34. The Launch 1 Vela spacecraft**
The X-ray detectors are the cubes projecting from the corners of the triangular solar panels which provide the energy source for the detectors and data transmitters.

*Source*: Photo courtesy of Los Alamos National Laboratory.

pression of the surrounding air. The hot gases become incandescent and emit intense visible light. As the shock wave becomes more intense the air around the fireball becomes opaque for a brief period and then, as the shock wave expands, the air becomes transparent again, allowing the release of another burst of light. It is this unique double flash of light that is the most useful signal of a nuclear explosion in the atmosphere (see figure 33). The general shape of the double pulse is the same for all nuclear explosions, regardless of yield or detailed design features, and the times to the two maxima and the minimum are well-known functions of the yield of the weapon.[139]

The double light pulse can be detected from a satellite with a device called a 'bhangmeter',[140] a special kind of photometer which uses sensors similar to the visible and infra-red sensors described earlier. The bhangmeter is focused on the Earth from a circular orbit at an altitude of 115 000 km (roughly one-third of the distance from the Earth to the Moon). From this distance the Earth is a relatively small, but very bright, sphere. Because of its large size and high reflectivity the Earth has a total luminosity which can be several thousand times that of a small nuclear explosion.[141] So the bhangmeter must incorporate electronic circuits which can separate the rapid fluctuations of light intensity characteristic of a nuclear explosion from the nearly constant bright background of light reflected from the Earth.

Both bhangmeters and X- and gamma-ray detectors, along with a number of other detection devices, have been watching the Earth and outer space since the first Vela satellite was launched in 1963[142] (see figure 34). Presumably similar devices are in use by the Soviet Union, and possibly other states as well. The last Vela satellite was launched in 1970, but nuclear detection equipment similar to that carried by Vela has been deployed on other types of satellite since then. The next generation of X- and gamma-ray detectors and bhangmeters will be part of the payload of the Navstar global positioning satellites (GPS), which will therefore also be nuclear detection satellites (NDS). The GPS/NDS satellites will be in operation by 1988 and will include "18 satellites deployed in 6 circular orbits of radius 26,600 kilometres, inclined at 60° to the equator and equally spaced in azimuth [longitude]"[143] (see figure 35). These satellites will serve the dual function of supplying navigational fixes for vehicles, ships and aircraft and watching for nuclear explosions in the atmosphere or in outer space.[144] The likelihood of any such explosion escaping detection by this system is extremely small.

## *VIII. Electronic reconnaissance*

The monitoring of radio and radar signals is at once the easiest to explain of all the technologies so far described and the one about which the fewest specific details are known. No technical intelligence-gathering methods are as sensitive and closely guarded as the signals (SIGINT) and communications (COMINT) intelligence techniques and devices used by many countries to

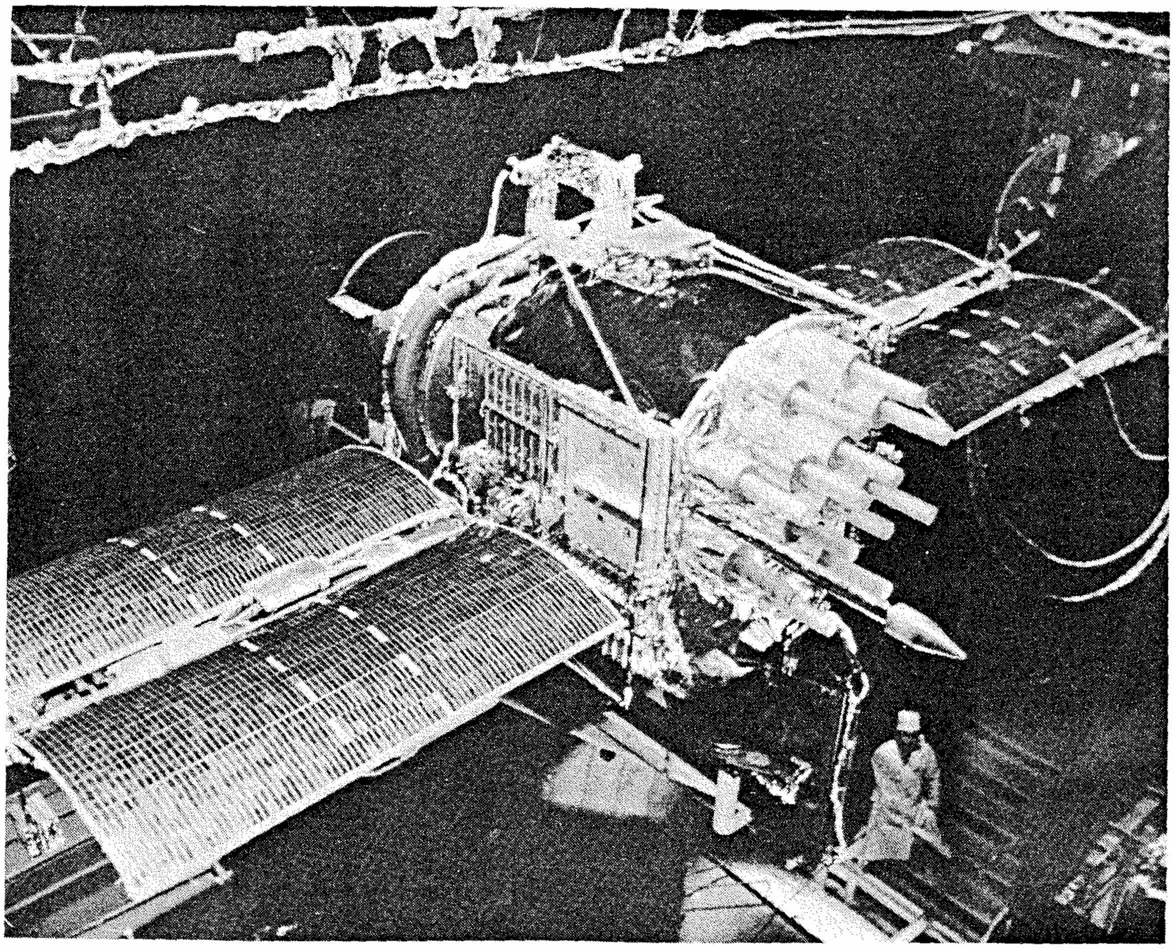


**Figure 35 a. A GPS/NDS satellite**
The satellite shown under construction is one of 21 which will ultimately be deployed in space by the USA. The wings of the satellite carry the solar power source, and the cylindrical structures pointing to the right are the various communication antennas used to provide navigational fixes for ships, aircraft, missiles and land vehicles. Not visible on the photograph are the bhangmeter and X- and gamma-ray detectors which will enable the satellite to detect nuclear explosions in the Earth's atmosphere or in outer space.

*Source*: Photo courtesy of US Air Force.

intercept the communications and radar signals of friends and enemies alike. Because of this secrecy the available literature on the subject is skimpy on details and riddled with speculations and contradictory assertions. In view of this situation it is not possible to give a satisfactory picture of the capabilities and limitations of electronic reconnaissance techniques, and one must be satisfied with a few rather superficial comments.

The most widespread use of electronic intelligence is in the interception of communications (COMINT). This ranges all the way from the tapping of telephones to the monitoring by satellites of microwave transmissions from Earth-based transmitters. Somewhere in this broad range lies the indistinct but significant border between legitimate national technical means of verification and illegitimate espionage. The precise location of this border is an issue which

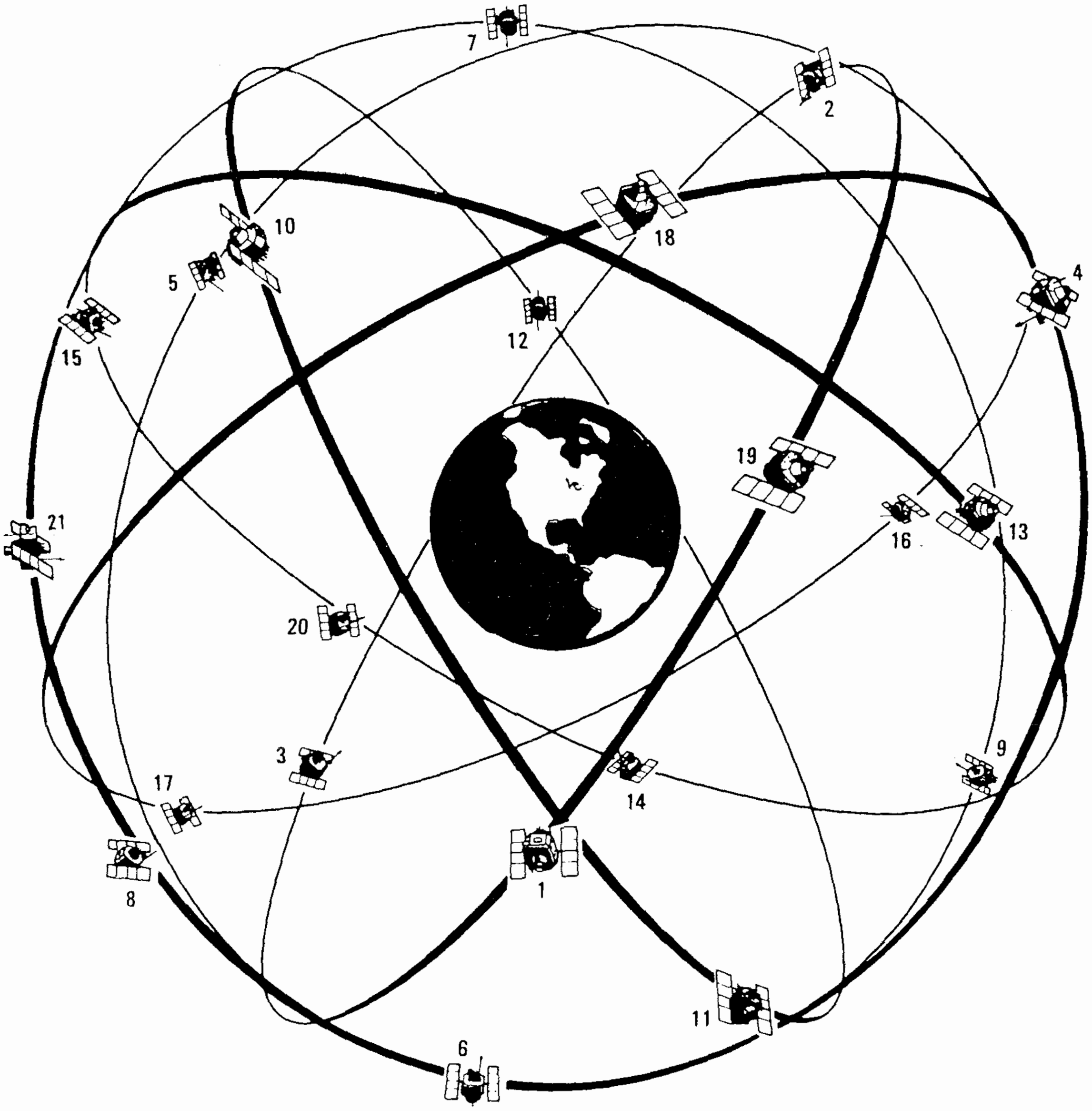


**Figure 35 b. Deployment of GPS/NDS satellites**
The GPS/NDS system, also called the Navstar system, will consist of 21 satellites deployed in 6 orbits as shown. The primary purpose of the system is to provide precise positioning information for a wide variety of applications, including such military uses as weapon delivery, rendezvous, precision mapping and point-to-point navigation. It is clear from the figure that the system will also provide a highly redundant capability for nuclear explosion detection.

*Source*: Courtesy of US Air Force.

no country seems anxious to discuss, and the phrase "national technical means of verification" has never been clearly defined, mainly to avoid any serious examination of these techniques.

It is remarkable that such a highly secret and sensitive activity is carried out on such an enormous scale. For example, the US National Security Agency (NSA) attempts to collect and preserve on magnetic tape ("more or less forever") *all* Soviet radio transmissions, including "the full daily broadcast of every conventional radio station in all the Soviet republics, every transmission to every Soviet embassy abroad, every broadcast to a ship at sea, every transmission by military units on maneuvers in Eastern Europe, the radio traffic of every control tower at Soviet airports".[145] To this must be added the substantial efforts applied to monitoring the communications of many other states as well.[146] Soviet SIGINT/COMINT activities are obviously also extensive, although almost nothing has been written about them in the open literature.

The collection, decoding (decrypting), monitoring and storing of this vast volume of radio traffic requires an enormous organization including tens-of-thousands of people and facilities distributed all around the Earth, on land, on sea, in the air and in space. The main NSA facility at Fort Meade, Maryland has a floor area of 180 000 square metres of which some 25 per cent (45 000 square metres), is devoted to computers used for code breaking, traffic and content analysis and record keeping.[147]

The vast majority of this information has little or nothing to do with the verification of arms control agreements, and is related to political, military and economic intelligence gathering. But it is important to understand that this capability to monitor virtually all of the communications of another state must act as a powerful inhibiting factor on any attempts by that state to carry out clandestine activities, especially activities which require the co-operation of several separate facilities and substantial numbers of people. Almost any significant violation of an arms control agreement would fit this definition and would therefore face serious risks of discovery unless highly elaborate and expensive precautions were taken, precautions which would not only reduce the efficiency of the clandestine activity, but which might in themselves arouse suspicion and increased attention. An interesting historical example was the realization in 1942 by Soviet scientists that the USA was working on an atomic bomb. The very secrecy of the project, which resulted in the disappearance of many prominent physicists and the sudden absence of articles on nuclear fission in physics journals, alerted Soviet researchers to the Manhattan Project.[148]

The aspects of SIGINT/COMINT which are most relevant to present arms control problems and which seem to have become accepted as legitimate national technical means are the monitoring of radar signals and the radio transmissions (telemetry) from missile test-flights. Radar signals must be monitored to verify that large phased-array radars are not being tested in an 'ABM mode' as forbidden in the SALT I Treaty, and the monitoring of telemetry is important, possibly essential, in verifying compliance with the highly detailed and complex prohibitions against 'new types' of ICBMs and limits on multiple warhead deployments included in the SALT II Treaty.

## Radar signals

Radars emit pulses of electromagnetic energy with distinctive frequency and amplitude characteristics, which are reflected off objects and returned to receivers designed to interpret them. But these pulses can also be intercepted by antennas deployed on aircraft or satellites, and much can be learned in this way about the location, purpose and capabilities of the radar.

The shooting down of a Korean Airlines passenger aircraft by Soviet air defence forces in September 1983 called attention to the use of aircraft by intelligence agencies to monitor radar transmissions. Most commonly the aircraft stay just outside the airspace of the state being observed, but execute maneouvres designed to alert air defences. For example, "About seventy times each year big Soviet Tu-95 'Bear' reconnaissance aircraft veer inside the [US] Air Force's Aerospace Defence Indentification Zone (ADIZ), a buffer area surrounding US airspace, which ranges from 60 to 200 miles [96–320 km] wide . . . The US intercepts more than 300 Soviet aircraft each year in the ADIZ".[149] More rare, but still surprisingly frequent, are the actual penetrations of national airspace by hostile reconnaissance aircraft.

Such activities clearly cross the boundary between legitimate and illegitimate national technical means, and a substantial number of aircraft have been shot down as a result of such activities.[150] Soviet officials publicly charged that the Korean airliner was being used for just such a mission, but no persuasive evidence has been produced to support this charge.

Satellites used for SIGINT are often called 'ferret' satellites.[151] They are usually placed in orbits slightly higher than those used for photographic satellites and are sometimes used in pairs with one satellite at relatively high altitude and the other at a low altitude (around 200 km).[152] Very little is publicly known about the configurations and capabilities of these satellites, and the literature abounds with contradictory and confusing statements. For example a recent generation of US satellites called 'Rhyolite' has been described by one source as designed primarily "to scan the Soviet Union with infra-red sensors to detect missile booster exhaust plumes"[153] and by another as "pure SIGINT".[154] Most likely the Rhyolite satellites carry out both missions, but in the absence of hard information on the real missions and capabilities of Rhyolite and other ferret satellites there is no way to resolve such contradictory statements. And the level of secrecy surrounding these satellites is increasing rather than decreasing. In June 1983 the US government stopped releasing even the launch-times and orbital parameters of its own military satellites.[155]

The use of ferret satellites and other SIGINT monitors to verify the ABM Treaty (SALT I) involves determining whether or not a given phased-array radar exceeds certain power and size limitations or is tested in an "ABM mode".[156] The power of a radar can be measured by determining the strength of the signal at a known distance from the radar if the radiation pattern emitted by it is known. This is a straightforward measurement which can be carried out by several types of monitor.

The question of whether or not a given radar is being tested in an ABM mode is far more complex. Attempts to clarify this notion in the SALT I negotiations were not successful, as evidenced by a "unilateral statement" attached to the Treaty by the United States in which the US definition of "tested in an ABM mode" is spelled out. No indication of Soviet agreement with this unilateral statement is given, and there is no alternative Soviet definition. Therefore the monitoring of compliance with Article II of the ABM Treaty, which defines an ABM radar as one which has been tested in an ABM mode, remains an ambiguous process.

The USA in 1973–74 gathered evidence which it believed revealed Soviet testing of a phased-array radar in an ABM mode and made a complaint in the Standing Consultative Commission.[157] However, it was very difficult to build such evidence into a case for a violation, as indicated by the need to observe and analyse 40 incidents of Soviet testing of radars before a pattern could be established to show that that radar was being tested in an ABM mode.

So while ferret satellites and other SIGINT 'platforms' are capable of very thorough and precise monitoring of radar emissions, there do remain limits to their application to monitoring arms control agreements, especially when what is important to know is not simply the characteristics of the radar itself but its interaction with other components in a complex weapon system.

### Telemetry monitoring

In the context of arms control, telemetry generally refers to the radio data transmissions from missiles which are being flight-tested. In such a test it is important to monitor a great many components and sub-systems in the missile to see if they are functioning according to design and to locate malfunctions and design flaws. A missile under test contains all sorts of devices for measuring temperatures, voltages, currents, accelerations, vibrations, stresses, and so forth. Each one of these devices can be connected to a 'transducer' which converts its output into an electrical signal which is fed to a radio transmitter. The signal is then broadcast from one or more atennas mounted on the missile or its payload to receivers on land, on ships at sea, on aircraft or on satellites. An example of such a system for the monitoring of Trident I (C4) missile tests is shown in figure 36.[158]

The telemetry is radiated in many directions, so it can be received by anyone with an appropriate receiver in a line-of-sight from the missile. It has been shown by a group at the Kettering School in Great Britain that some very interesting and useful information can be obtained from satellite telemetry even with relatively inexpensive equipment and unsophisticated analytical techniques.[159]

A modern ballistic missile is a complex object, and in a test-flight of such a missile there are so many systems to be monitored that the data from several of them must be combined and transmitted together on a single channel. (A channel is characterized by a central frequency, called the 'carrier', and a

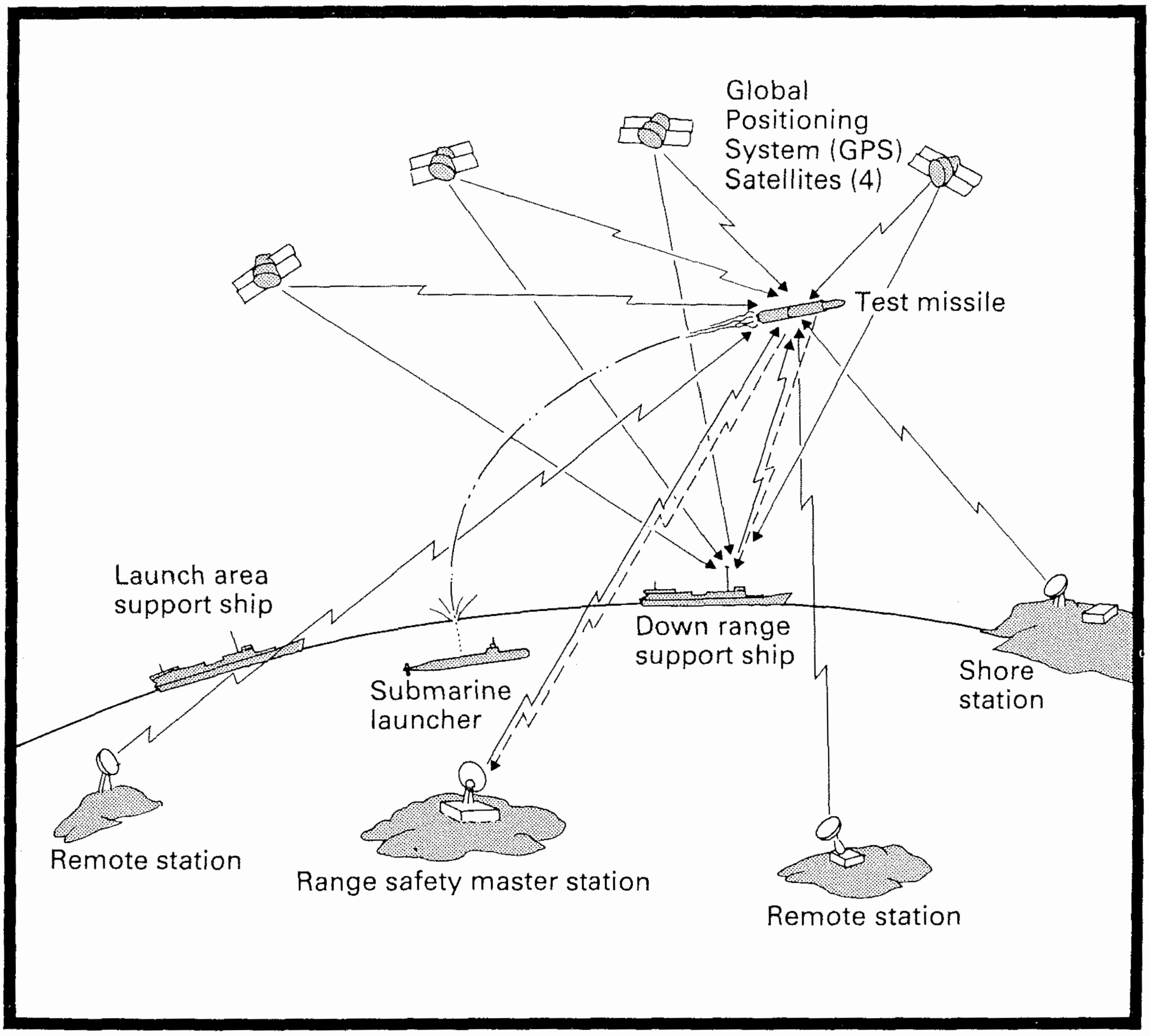


**Figure 36. Trident missile test telemetry**
The figure shows radio communications both to and from the missile being tested. Telemetry data from the missile is shown being received by the range safety master station and the down-range support ship. The transmissions to the missile from the GPS satellites and ground stations are tracking signals used to monitor the missile's precise position and velocity. If there is a malfunction and the missile begins to wander off course, a signal can be sent from a ground station to destroy it.

'bandwidth', that is, a certain spread of frequencies on either side of this central frequency. The bandwidth is related to the rate of transmission of information on the channel: the greater the information rate the greater the necessary bandwidth.) The mixing of several signals on a single channel is called 'multiplexing', and it makes intercepted telemetry signals very difficult to interpret.

There are a number of different modes for the transmission of telemetry. One, called pulse duration modulation (PDM), uses a signal of a fixed amplitude which can be turned on for varying fractions of a known time interval. The on-time can then be made proportional to some quantity of interest on the missile. The Soyuz satellite telemetry interpreted at the Kettering School was in PDM form, and the pulse durations appeared to be related to heart and respiration rates of the cosmonauts on board the satellite.[160]

An increasingly common mode of data transmission is called pulse code modulation (PCM). In PCM the data are first digitized, that is, put into the form of binary numbers (see section V, above), which are then transmitted on a channel made up of two separate carrier frequencies very close together. The ones are transmitted as pulses on one of these frequencies and the zeros as pulses on the other.

The Trident I tests use PCM on 192 channels, each of which is sampled 400 times per second by the multiplexer. The data stream consists of 76 800 numbers per second, each of which is in the form of an 8-bit digital 'word' (i.e., a number between 0 and 255), resulting in a data transmission rate of 614 400 bits per second.[161] Telemetry from Soviet missile tests has also moved increasingly to digital formats and PCM in recent years. One analyst has even speculated that reports of Soviet 'encoding' of telemetry were a result of this change from PDM to digital format.[162] It is true that converting data to digital form can be described as 'encoding' but this must be seen as different from 'encryption', which is a form of encoding whose sole purpose is to make the data incomprehensible to observers who are not supposed to receive them. It seems clear from the intensity of the controversy over the alleged encryption of missile test telemetry by the Soviet Union (see chapter 4, pp. 186–91) that the issue concerns more than the simple conversion from PDM to digital data formats.

While the analysis of missile telemetry is clearly a complex and difficult process, it is also clear that such analysis is conducted routinely and has been for many years.[163] The ability to intercept and analyse telemetry is an integral part of each side's national technical means of verification and is explicitly included in the SALT II Treaty.[164]

Telemetry is broadcast throughout the flight-test of an ICBM, starting with pre-launch preparations and ending with the impact of the warheads at their targets. Pre-launch and early boost-phase telemetry is important for an accurate determination of the launch-weight of the missile, a feature restricted by SALT II. But telemetry from low altitudes must be monitored from systems which are not too far away from the launching site, because the curvature of the Earth prevents reception of telemetry at great distances. This is no problem for Soviet intelligence, since both US launching sites are on coastlines, and Soviet 'trawlers' equipped with sophisticated receivers can, and do, approach these areas, as well as the target areas, quite closely.[165]

For the United States the problem is more difficult, since Soviet test sites are deep in the interior of Soviet territory. The USA uses ground stations in Norway, Turkey, China and other states bordering on the Soviet Union as well as aircraft patrolling border areas to monitor Soviet telemetry.[166] The loss of two ground stations in Iran in the revolution of 1979 dealt a severe blow to the hopes of ratification of SALT II in the US Senate, as a number of Senators argued that without the Iranian ground stations the Treaty could not be verified adequately.[167] Proponents of SALT II verification argued that the loss of these sites did not prevent adequate verification, but they were unable to make these arguments convincing, either because the loss of Iranian sites really *did* seriously degrade US monitoring capabilities, or because the alternative systems which could replace them were too secret to reveal in public. This problem of attempting to gain public confidence in verification while maintaining maximum secrecy about capabilities is analysed in chapter 3.

Once the missile rises above about 100 km its telemetry can be monitored by more distant stations and by aircraft and satellites. Telemetry from high altitudes carries information on the detailed manoeuvres of the MIRV 'bus', the vehicle that releases the individual re-entry vehicles. This telemetry is useful for monitoring the SALT limitations on numbers of warheads (fractionation), but it is also very useful in monitoring accuracy, a property not covered by arms control agreements.[168]

Most accounts of the US Rhyolite satellite assume that it has the capability to monitor Soviet missile telemetry, and one even suggests that it could monitor boost-phase telemetry.[169] Given the great height of the orbit (the Rhyolite is in geostationary orbit 37 300 km above the Earth) it can be shown that a very large antenna is necessary to achieve a satisfactory signal to noise ratio.[170] It is reported that the main Rhyolite antenna is a concave dish 21.3 m in diameter, and that the satellite carries other antennas as well as "a number of other appendages".[171]

Telemetry can more easily be monitored from the lower orbits used by ferret satellites. Again the open literature is confusing and sometimes contradictory on the capabilities of ferret satellites to monitor missile tests. A 1979 source predicted a new satellite with a 20 m antenna under development and scheduled for deployment in 1982,[172] while a 1982 source referred to a "new ferret satellite equipped with a long antenna tailored for telemetry interception" which was "reportedly under development".[173]

Given the secrecy surrounding the launching of military satellites it may be that such a ferret has already been deployed, or, if the deployment of the satellite was dependent on the use of the US space shuttle, the deployment may have been delayed until January 1985. On 24 January a highly secret satellite with the orbital characteristics of a new type of ferret satellite[174] was placed into orbit as part of a space shuttle mission.

Whatever the current capabilities of US systems (and even less is known about Soviet SIGINT systems) it is clear that the interception and interpretation of missile-test telemetry is a high priority mission for US and Soviet national intelligence services, a mission deemed worthy of large expenditures of money and talent. Apparently the benefits to be gained from unrestricted access to such telemetry are substantial. This would explain the intense reaction generated in the US intelligence community to the alleged encryption by the Soviet Union of some portions of its telemetry. Given the importance of this issue in the arms control debate it is worth giving a short introduction to the techniques of encryption and decryption.

### Data encryption

The first point that must be made is that even when there is no intent to encrypt digital telemetry data, its interpretation can be very difficult. A given channel will carry multiplexed information from many instruments, some of which may be continuously monitored while others are only sampled at longer intervals. In addition, the relationship between the binary number transmitted for a given quantity and the actual value of that quantity may be very obscure. For example, a temperature at some location in the missile may vary under normal conditions over a range of 10 degrees, but if it is necessary to detect small variations in this temperature, then the 10-degree range can be divided into 256 intervals and the temperature value transmitted as an 8-bit binary number. To transform this number back into a temperature it is necessary to know both the nominal range and the temperature value assigned to the lower end. If, for example, the temperature being measured is known to be in the range $50^\circ$ to $60^\circ$, and the binary number 00110010 (50) is received, then the temperature can be read as 50/256 of 10 degrees over the base of 50 degrees, that is, 51.95 degrees. Unless an unauthorized listener knows the nominal base value and range the number 50 is of no value.

The only way in which one side can interpret another's telemetry is if certain standard channels and parameters are used repeatedly and it is possible to observe many tests and look for patterns in the data. Once patterns are found it is often possible to infer what quantities are being measured as did the Kettering group in the case of Soyuz telemetry.

There are considerable advantages in adopting standard procedures for telemetry broadcasts. Hardware can be standardized and computer analysis of received data simplified. It is reasonable to infer that both the USA and the Soviet Union have used such standard procedures for many years and that each is capable of interpreting significant amounts, if not all, of the other's telemetry.

But with advances in digital computer technology the ease with which telemetry can be encrypted has been greatly increased. To encrypt digital telemetry it is necessary only to put the digitized data through a process in which a secret binary 'key' number is added to the correct number.[175] The key can be an extremely long string of 1s and 0s or a shorter string repeated over and over again. When the key has been added the data become totally obscured and can be deciphered only by subtracting the identical key.

There are many routine uses of codes (for example in the transmission of financial records or diplomatic messages) in which the same key is used repeatedly.[176] Such codes can often be broken if the code-breaker has access to a large computer which can generate thousands of keys per second until the correct one is found. Such 'brute force' techniques are used routinely by the US National Security Agency which possesses an enormous computer capability. For certain other types of commercial code, clever mathematical techniques (algorithms) have been devised to break supposedly unbreakable codes with surprising speed.[177]

However, for relatively infrequent events such as missile tests there is no need to use the same key repeatedly, and the key can be changed for each test (presumably even *during* a test). Such 'one-time keys' are to all intents and purposes unbreakable, especially when the data being sent are not in the form of text but are already only strings of numbers. Therefore any state that wishes to withhold test data from unauthorized listeners can certainly do so with little risk that the data can be decrypted. In older encryption methods there was always the possibility that a spy might communicate the keys to the other side. But in modern computerized encryption a one-time key can be generated entirely within the computer, and there is nothing for a spy to communicate.

This ability to encrypt telemetry casts considerable doubt on the ability to monitor some of the more detailed restrictions embodied in existing arms control agreements. One solution to this problem would be to agree that telemetry encryption itself be banned, but both sides have shown resistance to such an agreement (see chapter 4). In the absence of a total ban on encryption the only solution is for the interested parties to negotiate detailed rules governing such encryption, but such negotiations would encounter deep resistance from the intelligence community who would want to prevent any discussion with the other side of existing SIGINT/COMINT capabilities.

## *IX. Safeguards*

The final set of verification technologies to be discussed depends far less on sophisticated hardware than those so far described and much more on an elaborate set of record-keeping and administrative techniques. These are the so-called 'safeguards' administered by the International Atomic Energy Agency (IAEA) in Vienna, an operation which began on a very small scale in 1957 and which by February 1984 had negotiated safeguards agreements with 84 states.[178] These agreements and the monitoring activities carried out under them constitute an unprecedented international co-operation to attempt to prevent the proliferation of nuclear weapons.

### Objectives

The objectives of the IAEA safeguards are: "the timely detection of diversion of significant quantities of *nuclear material* from peaceful nuclear activities to the manufacture of nuclear weapons or of other nuclear explosive devices or for purposes unknown, and deterrence of such diversion by the risk of early detection".[179] This statement has been carefully drafted not only to specify the Agency's responsibility but also to make clear the very significant limits on its responsibility. An understanding of these limits is crucial to an appreciation of the role that the safeguards system plays in current arms control verification and of how it might be extended or adapted to other arms control situations.

According to the statement of purpose, safeguards apply only to 'peaceful nuclear activities', that is, to nuclear facilities and materials devoted to non-military functions such as electric power generation or research. This means that the military nuclear facilities of the so-called 'nuclear weapon states' (USA, USSR, UK, France and China) are not subjected to safeguards, even for the three states (USA, USSR, UK) that have signed the Non-Proliferation Treaty. Nor are the civilian nuclear facilities of these states required to be under safeguards. It has only been in the past several years that the USA, UK and USSR have agreed to place some of their non-military nuclear activities under IAEA safeguards.

The clear separation of military and civilian applications of nuclear energy implied by the statement of objectives has been questioned by many people ever since the earliest days of the nuclear age. In fact, the original study on which the US Baruch Plan for international control of atomic energy was based (the Acheson–Lilienthal Report) denied the practicality of making this separation, emphasizing that "*safe* [i.e. non-explosive] *operations are possible only because dangerous ones are being carried out concurrently*".[180]

A second limitation of safeguards is that they are intended to deter diversions of sensitive materials, not prevent them. Actual prevention of diversion requires the authority of a sovereign state and falls under the concept of 'physical protection' of such materials, not safeguards.[181] Because the IAEA is an international organization it does not have the authority to use force or other coercive measures to modify the behaviour of nuclear facility operators or states. It can only serve as a deterrent by threatening exposure of an attempt to divert nuclear materials from non-military to military purposes. Just how effective this deterrence function is cannot be assessed accurately, since it depends not only on the probability of detection but on the potential benefits a state might see in cheating successfully as compared with the costs of being exposed prematurely as a violator.

The next set of limitations on safeguards derives from the definitions of the phrases 'timely detection' and 'significant quantities of nuclear materials'. A significant quantity (SQ) of a nuclear material is defined as the approximate amount needed to produce a nuclear weapon after account is taken for whatever processing must be done to put the nuclear material into usable form as an explosive.[182] Values for 'significant quantities' of nuclear explosive materials can be inferred from the data in table 4. For example, a total amount of 92.9 tonnes of plutonium in irradiated fuel represents 11 600 SQs, implying that 1 SQ = 8.0 kg for plutonium in this form.

'Timely detection' has turned out to be much more difficult to define. This

**Table 4. Approximate quantities of material subject to IAEA safeguards except that covered by voluntary-offer agreements with nuclear weapon states at the end of 1983**

| Type of material | Quantity of material (t) in NNWS | Quantity of material (t) in NWS[a] | Quantity in SQ |
|---|---|---|---|
| *Nuclear material* | | | |
| Plutonium[b] contained in irradiated fuel | 85.8 | 7.1 | 11 600 |
| Separated plutonium | 5.3 | 1.5 | 850 |
| HEU (equal to or greater than 20% uranium-235) | 11.0 | 0 | 260 |
| LEU (less than 20% uranium-235) | 17 600 | 990 | 5 820 |
| Source material[c] (natural or depleted uranium and thorium) | 28 000 | 0 | 2 270 |
| *Total significant quantity* | | | 20 800 |
| *Non-nuclear material*[d] | | | |
| Heavy water | 1 307 | 0 | –[e] |

[a] Material in facilities in nuclear weapon states subject to safeguards under safeguards transfer agreements.
[b] The quantity includes an estimated 39.7 t (4 970 SQ) of plutonium in irradiated fuel, which is not reported to the Agency under the reporting procedures agreed to (the non-reported plutonium is contained in irradiated fuel assemblies to which item accountancy and containment and surveillance measures are applied).
[c] This table does not include material within the terms of sub-paragraphs 34(*a*) and (*b*) of INFCIRC/153 (Corrected)—in essence, yellowcake.
[d] Non-nuclear material subject to Agency safeguards under INFCIRC/66/Rev.2-type agreements.
[e] "Quantity in SQ" does not apply to non-nuclear material.

*Source: IAEA Annual Report for 1983* (IAEA, Vienna, 1984), p. 68.

is not so much a technical problem as a political and adminstrative one, since the timeliness criterion is used to set the frequency of on-site inspections, and facility operators have a strong interest in keeping these to a minimum.[183] The compromise solution arrived at by the IAEA has been to define the necessary detection time to have the same 'order of magnitude' as the 'conversion time',[184] which in turn is defined as the time required to convert a given material into the 'metallic components of a nuclear explosive device'.[185] The official conversion times for various materials are listed in table 5.

Most of the conversion times are reasonable and have led to arrangements for relatively frequent inspection visits. In fact, the IAEA has decided that diversion possibilities are so great at facilities that process plutonium that the continuous presence of inspectors is necessary.[186] However, one conversion time, the one year allowed for conversion of low-enriched uranium to highly enriched nuclear explosive, is quite unrealistic given the capabilities of modern ultra-centrifuge enrichment facilities. Using a small clandestine centrifuge plant a state could produce enough very pure uranium-235 for a bomb in less

**Table 5. Estimated material conversion times**

| Material classification | Beginning material form | End process form | Estimated conversion time |
|---|---|---|---|
| 1 | Pu, HEU[a], or $^{233}$U metal | Finished plutonium or uranium metal components | Order of days (7–10) |
| 2 | $PuO_2$, $Pu(NO_3)_4$ or other pure compounds; HEU or $^{233}$U oxide or other pure compounds; MOX or other non-irradiated pure mixtures of Pu or U [($^{233}$U + $^{235}$U) > 20%]; Pu, HEU and/or $^{233}$U in scrap or other miscellaneous impure compounds | Finished plutonium or uranium metal components | Order of weeks[b] (1–3) |
| 3 | Pu, HEU or $^{233}$U in irradiated fuels[c] | Finished plutonium or uranium metal components | Order of months (1–3) |
| 4 | U containing <20% $^{235}$U and $^{233}$U; thorium | | Order of one year |

[a] Uranium enriched to 20 per cent or more in the isotope $^{235}$U.
[b] While no single factor is completely responsible for the indicated range of 1–3 weeks for conversion of these plutonium and uranium compounds, the pure compounds will tend to be at the lower end of the range and the mixtures and scrap at the higher end.
[c] Irradiation level is chosen on a case-by-case basis.

*Source: IAEA Bulletin*, Vol. 22, No. 3/4, August 1980, p. 6.

than three weeks using an amount of low-enriched or natural uranium whose diversion from a large bulk-handling facility would be very difficult to detect.[187] Unfortunately, the process of changing such a number once it is set is extremely difficult in an agency as large and politically diverse as the IAEA. This kind of inflexibility is an important disadvantage of international approaches to verification (see chapter 4).

## Materials accounting

The major technique employed by the IAEA in carrying out its verification responsibilities is 'nuclear materials accountancy'.[188] It begins with a detailed agreement between the IAEA and the state in which the facility to be safeguarded is located. The IAEA is first given design information on the facility, which is used to designate a number of 'material balance areas' (MBAs) and 'key measurement points' (KMPs). An MBA is an area where nuclear materials are stored, for example the spent fuel pool of a nuclear power reactor or the product storage area of an enrichment plant. A KMP is generally a point of transition at which nuclear materials move from one MBA to another or into or out of the facility, for example a pipe carrying waste out of a reprocessing facility or a loading dock at a fuel fabrication plant.

Also part of the IAEA-state agreement is the creation by the state of its own system of accounting for nuclear materials at the facility.[189] The state agrees to maintain accurate data on inventories in MBAs and flow or transport through KMPs. The IAEA keeps its own set of records based on the initial inventories established at the opening of a safeguarded facility and the subsequent reports of material flows and inventories submitted by the operators. The records kept by the operators are periodically verified by independent on-site measurements made by IAEA inspectors. In a typical site visit the inspectors will audit the records of the facility and make their own measurements of inventories in MBAs and flow rates through KMPs as well as sample measurements to verify the declared compositions of materials in the facility. Some of these composition measurements can be made on-site by so-called 'non-destructive assay' (NDA), while others must be made by taking samples which are sent for chemical, spectroscopic or radiometric analysis to the IAEA's own laboratory in Siebersdorf, Austria. This laboratory is capable of processing about 2 000 samples per year.[190]

Because such remote analysis is costly in both time and money it is desirable to make as many on-site non-destructive measurements as possible. Most of these are intended to measure the percentage composition of uranium and plutonium isotopes contained in fuel rods, casks, tanks and so on. The most commonly used devices for these measurements rely on gamma-ray counters similar to the X-ray detectors described in section VII.[191]

Every radioactive isotope has a unique 'signature' which is carried by the radiation it emits. Sensitive detectors can read this signature at considerable distances even if the material being monitored is shielded by barriers. For example, it was with a simple, portable gamma-ray detector that Swedish researchers were able to detect the presence of uranium—and therefore possibly a nuclear weapon—aboard a Soviet submarine which ran aground near Karlskrona in October 1981. The monitoring was carried out from a small boat next to the submarine, and enough radiation passed through the hull to allow positive identification of the presence of 10 kg of uranium-238 and the reasonable inference that this was part of a nuclear weapon carried in the submarine's torpedo room.[192]

The total on-site inspection effort of the IAEA in 1983 consisted of about 1 840 inspections carried out at 520 nuclear installations in 53 states. Non-destructive assays were conducted as part of 26 per cent of these inspections and more than 1 150 analyses of plutonium and uranium samples were performed at the Siebersdorf Laboratory.[193] The impressive scope of application of IAEA safeguards can be seen in tables 4 and 6 which show the amounts of

**Table 6. Installations in non-nuclear weapon states under safeguards or containing safeguarded material at the end of 1983**

| Installation category | Number of installations<br>INFCIRC/153[a] | INFCIRC/66/Rev.2 | Total[b] |
|---|---|---|---|
| A. Power reactors | 121 | 26 | 147 (143) |
| B. Research reactors and critical assemblies | 151 | 26 | 177 (177) |
| C. Conversion plants | 5 | 2 | 7 (6) |
| D. Fuel fabrication plants | 32 | 8 | 40 (39) |
| E. Reprocessing plants | 4 | 2 | 6 (6) |
| F. Enrichment plants | 4 | 0 | 4 (4) |
| G. Separate storage facilities | 26 | 2 | 28 (23) |
| H. Other facilities | 45 | 1 | 46 (42) |
| I. Other locations | 398 | 27 | 425 (404) |
| J. Non-nuclear installations | 0 | 1 | 1 (0) |
| Totals | **786** | **95** | **881 (844)** |

[a] Covering safeguards agreements pursuant to NPT and/or Tlatelolco Treaty.
[b] Numbers for 1982 are indicated in parentheses for comparison.

*Source: IAEA Annual Report for 1983* (IAEA, Vienna, 1984), p. 69.

nuclear materials and the numbers and types of facilities under safeguards at the end of 1983.[194]

The end result of all of this measuring and accounting is a set of values for 'material unaccounted for' (MUF) at each MBA. The MUF value is the discrepancy between the 'book inventory' derived from accounting records and the 'physical inventory' as measured at the end of each 'material balance period'.[195] Every value of MUF must be accompanied by an estimate of the expected range of error so that standard statistical tests can be applied to determine whether or not the MUF is significant. A significant value of MUF is called an 'anomaly', and unless it is satisfactorily resolved by further investigation such an anomaly can lead to the conclusion that a diversion of nuclear materials has occurred and initiate the IAEA sanctions procedures.[196] The IAEA's Annual Report notes that 420 such anomalies were found during 1983 and that all but one had been satisfactorily explained at the time of publication of the report.[197]

## Containment and surveillance

The total number of facilities under IAEA safeguards at the end of 1983 was 881 (see table 6), and as non-military facilities in the USA and USSR are added the number will rise appreciably. At the same time the IAEA is constrained in its ability to add new inspectors by budgetary restraints and the difficulty of recruiting and retaining qualified personnel.[198] It is not surprising, therefore, that the IAEA has placed increasing emphasis on containment and surveillance technology to limit the demand for human inspectors.

Containment is the process of restricting the movement of nuclear materials by the use of various kinds of physical barriers, such as walls, transport flasks, containers and so on.[199] The primary technology for containment purposes is a simple seal which is designed to reveal any attempt to break or tamper with it. As of December 1981, IAEA inspectors were applying over 3 000 such seals per year, and the computerized history of more than 10 000 seals had already been accumulated at IAEA headquarters.[200] One disadvantage of these seals is that they must be sent back to the laboratory to check for tampering or replacement. This has led to an effort to develop fibre optic (see figure 37) or electronic seals which could be checked either on-site or by remote control.[201] However, development of these new devices has been slow, and the Agency still relies almost entirely on the traditional seals. In 1983, 6 600 were used, more than double the number in 1981, illustrating both the expansion of the Agency's responsibilities and its increasing emphasis on containment measures.[202]

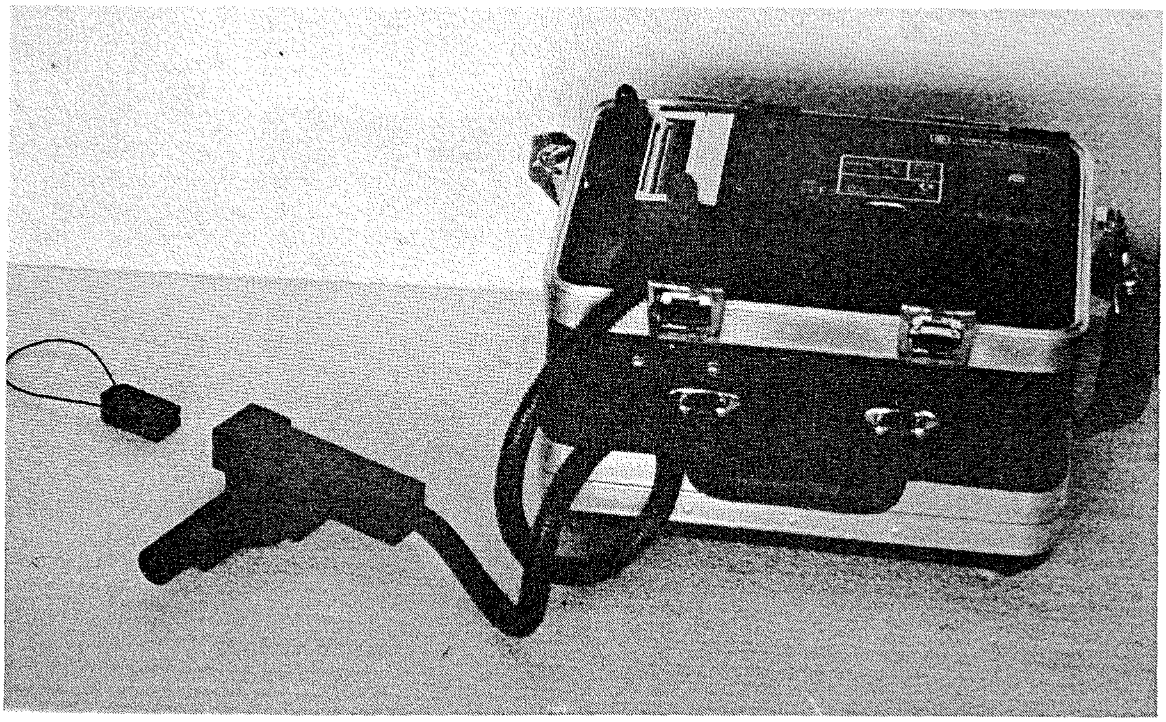


**Figure 37. COBRA prototype fibre-optic seal and verifier**

In a fibre-optic seal the seal wire is replaced by a multi-strand plastic fibre-optic loop, the ends of which are enclosed in a seal in such a way that a unique random pattern of fibres is formed. This can be verified by shining a light into the ends of the loop and observing the magnified pattern of the fibre ends, either visually or photographically; development is also being directed towards television recording of images.

*Source*: IAEA, *IAEA Safeguards, Safeguards Techniques and Equipment,* IAEA/SG/INF/5 (IAEA, Vienna, 1984), p. 29, figure 18.

The word ‘surveillance’ has much the same meaning when applied to safeguards as it has for more general intelligence activities. It is the collection of information through the use of monitoring devices (or on-site inspectors) in order to detect undeclared movements of nuclear material or tampering with safeguards devices.[203] For many years the most commonly used surveillance device was a dual super-8 motion picture camera which could take single-frame photographs every 20 minutes for 100 days before film reloading (7 200 frames).[204] A new generation of surveillance systems uses a closed circuit television monitor and magnetic tape recording (see figure 38), which not only eliminates the need for film developing but also permits remote monitoring.[205] Such television monitors can also be equipped with their own infra-red light source to allow surveillance at all light levels.

The concept of remote monitoring of containment and surveillance devices is a very attractive one and has been embodied in a project called RECOVER (remote continuous verification). In this system all electronic seals and television monitors would be connected via international telephone lines or relay satellites to IAEA headquarters, where their performance and status could be checked periodically by simply dialling a phone number.[206] The present RECOVER concept does not involve the actual transmission of data, but there

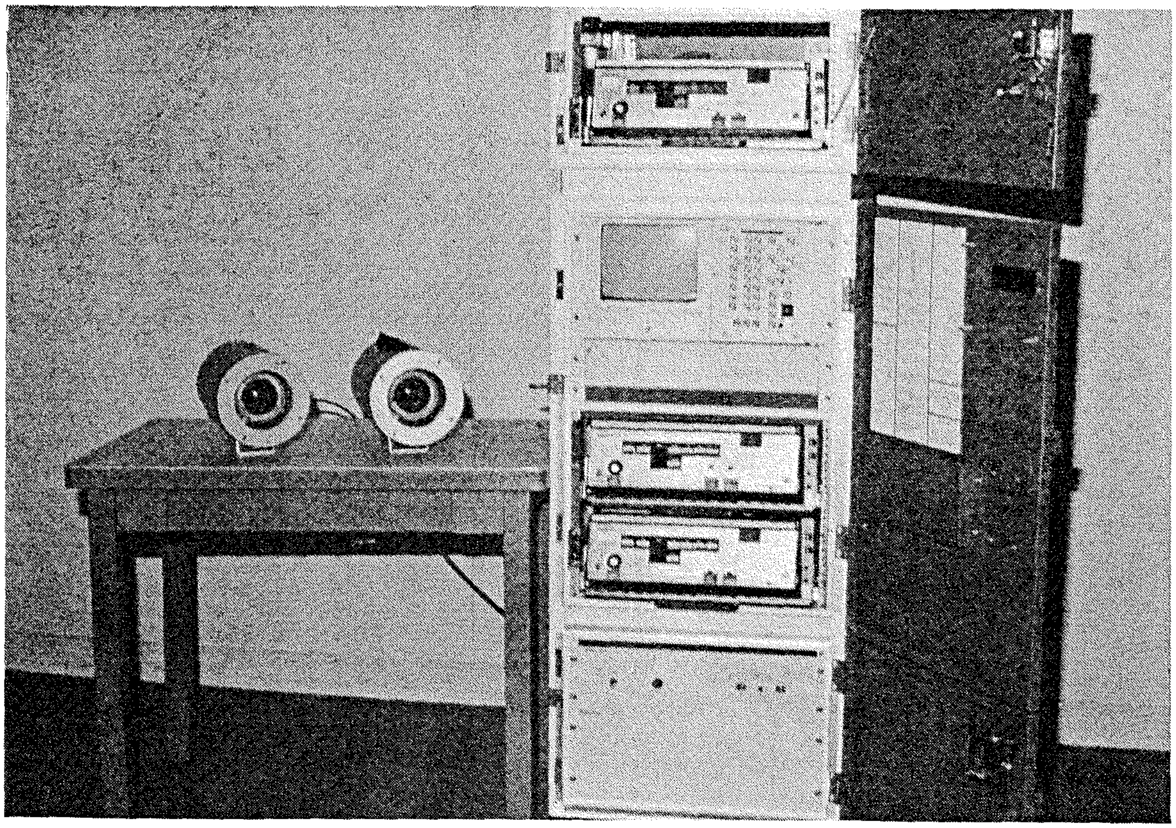


**Figure 38. Closed circuit television monitor**

*Source*: Photo courtesy of IAEA.

is no doubt that the system could be designed to do this, much in the same manner as the international seismic data exchange discussed in chapter 4. The RECOVER system has already been tested using encrypted digital data transmitted over secure communication lines.[207]

The RECOVER programme began in the USA during the Carter Administration with the original research contract going to the TRW Corporation.[208] Unfortunately, the programme has received very little support from the Reagan Administration, and the only serious research currently being done on it is in Japan, where it is being studied for possible adaptation to the Japanese national safeguards programme.[209]

Developments in modern containment and surveillance technology, along with remote monitoring concepts like RECOVER, have led to suggestions that the IAEA safeguards system might be extended to include the verification of a total ban on production of nuclear explosives or bans on chemical, biological or radiological weapons.[210] From the purely technical point of view there do seem to be some interesting possibilities for the adaptation of IAEA surveillance and containment concepts to, for example, the monitoring of certain chemical or nuclear production facilities which are shut down and moth-balled under an agreement.

One idea which has already been studied in some detail is the remote monitoring of a chemical weapon destruction facility.[211] Given the high toxicity of the materials being processed, such a facility would have to be highly automated, and the monitoring systems would have to be automated as well. However, the instruments for monitoring chemical substances must be very different from those used to monitor radioactive nuclear substances. Where the latter can often be assayed with non-destructive techniques this will usually not be possible with chemicals. So instead of the fuel bundle counter and gamma-ray spectrometer which could verify the input stream to a nuclear reprocessing plant, a chemical weapon destruction plant would need flow meters and gas chromatographs, instruments which are generally less precise and less convenient to use. However, this may not be a serious problem, since precision in measuring quantities is less important for chemicals than for nuclear materials. Adding a RECOVER system to allow remote monitoring of several facilities from a central location would certainly be feasible as well.[212]

This example shows that opportunities do exist to apply IAEA safeguards experience in new fields of arms control, but at the same time the prospects for such applications should not be exaggerated. Quite aside from the political and administrative problems which are analysed in chapter 4, there are also technical obstacles which will limit the use of on-site inspection, containment and surveillance, as well as remote monitoring in verifying bans on chemical or biological weapons. In contrast to the world nuclear industry, which involves fewer than one thousand facilities, the chemical industry comprises many thousands of facilities of all types and sizes. To attempt to inspect and monitor all of these would create data management problems at least on the scale of those faced by international satellite photography or an international seismic network. It is appropriate to ask whether the danger of chemical weapons is serious enough or the monitoring of declared facilities comprehensive enough to warrant the great expense and complexity such a system would entail.

## *X. The importance of synergism*

The variety and sophistication of the monitoring systems just surveyed must be viewed as contributing to a powerful, integrated intelligence-gathering capability for any state which possesses them. There are many detailed discussions of individual systems such as photographic satellites or seismic networks in the literature on verification, but much less common are studies which point out the many interactions among these various systems in making the whole considerably greater than the sum of its parts.

It is relatively rare that a single piece of evidence gathered by a single monitoring system can be the basis for a charge of violation. Much more often the individual bits of evidence are ambiguous when taken separately and only acquire significance when assembled together in a pattern with other ambiguous bits of evidence. The art of intelligence is the ability to assemble such patterns, and this same art is necessary in analysing the vast quantities of data produced by so many different monitoring systems. A few simple examples will be considered here.

The verification of a comprehensive nuclear test ban would certainly involve a world-wide network of seismic detectors, but even such a network will inevitably have some threshold explosive yield below which the identification of a seismic event as an explosion or an earthquake becomes highly ambiguous. This ambiguity will lead to a certain rate of 'suspicious' events, and the usual remedy suggested for this problem is on-site inspection. Since this remedy may be considered by some to be either technically unfeasible or politically undesirable, or both, it is important to reduce the number of suspicious events by other means in order to keep the demands for on-site inspection to a minimum.

Such other means exist in the form of photographic satellites which can often detect the preparations for nuclear tests. Such a detection was made by both US and Soviet reconnaissance satellites when what appeared to be preparations for a nuclear test were discovered in South Africa in 1977.[213] Preparations for such a test involve drilling a deep hole, placing instruments around the test site and delivering and arming the device. Such activity inevitably takes at least several days, possibly much more, and is difficult to conceal from the prying eyes of photo-reconnaissance satellites.[214] In addition, an underground nuclear explosion often leaves a 'subsidence crater' as the Earth's surface above the explosion collapses into the newly created cavity (see figure 39).[215] Such craters are easily observed from satellites and may be

**Figure 39. Subsidence craters at the Nevada Test Site**
An underground nuclear explosion creates a large cavity into which the rock above the explosion can fall. This creates a 'chimney' which can reach all the way to the surface, causing a visible crater. In order to prevent such craters with high confidence, explosions must be set off at sufficiently great depths and in the proper geological environment.

*Source*: Courtesy of Lawrence Livermore National Laboratory, CA.

difficult to prevent with confidence in an untested geological area. Next there is the added risk that radioactive materials will be released into the atmosphere by the explosion and be detected downwind of the test site. Such releases are not uncommon even for underground tests, and the monitoring of ambient radioactivity is now carried on as a routine aspect of weather and air-quality monitoring in many states around the world. Finally, a space-based thermal infra-red sensor could possibly notice the increase in temperature of the test area after the explosion. Therefore, photographic satellites and radiation detectors can add considerably to the effectiveness and acceptability of a seismic network by reducing the number of ambiguous events and consequently the number of demands for on-site inspection.

A second example of synergistic interactions among a number of systems is the monitoring of ICBM tests. Such tests are observed by infra-red sensors from geostationary satellites, by ground- and sea-based radars and by the interception of communications and telemetry. All these systems interact to produce a much more detailed and complete picture of the test than any of the systems could provide by itself. In addition, it is possible to measure or observe the same feature with more than one of the systems and compare the results. Such cross-checks can greatly increase confidence that the measured values are accurate and can make far more difficult any attempt to disguise or hide the data. Another useful synergism, the superposition of radar, visual and infra-red images to aid in object detection and camouflage penetration, has already been mentioned in section V.

A third example is the monitoring of troop or weapon restrictions in certain zones, such as in Central Europe. Here again, photographic and infra-red satellites can play an important role along with ground-, air- and space-based radars and signals and communications intelligence. While clever camouflage might hide certain things from satellites, the challenge of hiding militarily significant activities from the combined vigilance of all of these systems is far more difficult and risky.

Many other examples of the advantages of synergism could be listed, and there is no doubt that taken together the technologies described here constitute a highly reliable mechanism for monitoring arms control agreements. Yet even this vast array of techniques is not perfect. For example, Argentina was able to construct a uranium enrichment facility in total secrecy over a period of several years.[216] The plant is based on the gaseous diffusion process and is said to be only the first module of a small commercial facility, which can explain why it was not recognized by satellites. A completed gaseous diffusion plant, even of relatively low capacity, would be extremely difficult to conceal. But if the gas-centrifuge process had been chosen instead of gaseous diffusion, it is quite possible that the existence of a militarily significant facility could have been kept a secret even longer.

Incidents such as this do not call into question the great value of verification in support of arms control agreements. They do, however, serve as useful reminders that no system will ever be perfect and that to demand perfection is, as it always has been, to make the best the enemy of the good.

## *XI. The technological dimension of verification*

This chapter has described a remarkable range of monitoring technologies whose sophistication and comprehensiveness have been steadily increasing for many years and can be expected to continue to increase for many more. The impressive capabilities of individual systems combined with the synergistic effects of their interactions with each other give an encouraging picture of the existing and potential prospects for effectively monitoring many kinds of arms control or disarmament agreements.

However, this encouragement must be tempered by the realization that technological progress in the weapons to be monitored is proceeding at least as rapidly as is that of the monitoring systems. One cannot escape the intimate connection between arms control monitoring and military intelligence gathering, and as long as efforts continue to frustrate the latter process the former process will inevitably be made more difficult. There is in fact a qualitative arms race going on between "hiders and finders",[217] and it is not at all clear who the ultimate winners of this race will be.

A review of the monitoring devices and techniques just described will show that the easiest objects or events to monitor are those of large size, fixed location, substantial energy release, high temperature and distinctive appearance or signature. Fortunately this includes a considerable range of weapons and military preparations such as fixed-site ICBMs, nuclear missile submarines, nuclear weapon tests (both above and below the surface of the Earth), missile launches (whether for testing the missiles themselves or for experimenting with anti-satellite or ballistic missile defence systems), large phased-array radars, and most nuclear facilities. All of these objects are extremely difficult to hide from regular monitoring by the remote-sensing devices described here, and efforts to cheat on an agreement involving such weapons or activities would involve a very high risk of detection.

This positive view of verification must be balanced by some very important negative factors. First, there is the problem of political context in which evidence acquired by technical means is evaluated. History has shown that even the kinds of evidence just described as highly reliable can lead to intense political controversy as a result of differing attitudes towards arms control and military doctrine as well as differing assessment of the capabilities, motivations and intentions of rival states. This problem is important enough to deserve an entire chapter of its own, and it is the subject of chapter 3.

A second problem, also largely political, is the great expense and technical sophistication of most of these monitoring technologies. This means that they can be developed and deployed by only the richest and most technically advanced states, while less developed states, with security concerns which are at least as great of those of the major powers, must live in far greater uncertainty and/or become dependent on one or the other of the great powers for information vital to their national security. Such uncertainty and dependence are a source of increasing international concern and also require a separate discussion (see chapter 4).

The third problem is purely technical and involves the growing gap between the capabilities of monitoring systems and the qualitative features of newer generations of weapon. Probably the most important example of this trend is in land-based nuclear missiles where the same features of large size and stationary location which lead to easy monitoring also lead to high vulnerability to pre-emptive attack. As nuclear missiles have become more accurate this problem of vulnerability, and its accompanying sense of crisis instability, has become more acute. It has been suggested that the rational solution to this problem is to eliminate land-based ICBMs (either bilaterally or unilaterally) and rely on less vulnerable submarine-based systems for deterrence of nuclear attack.[218] However, the actual course taken by both the USA and USSR has been to develop smaller, more mobile and more flexible nuclear missiles such as the US cruise missiles, the Soviet SS-20 and SS-X-25, and the proposed US 'Midgetman' missile, which will be much smaller and more mobile than the present generation of Minuteman and MX ICBMs.

In this connection it has been revealed that actual deployment of US nuclear-armed cruise missiles on submarines has begun.[219] The presence of such missiles on submarines is simply impossible to verify by any means short of physical inspection, and even this method would not provide a high degree of reassurance. There is an inherent difficulty in attempting to use on-site inspections to verify the presence or absence of highly portable objects.

Another important trend is towards so-called 'dual-capable' weapons, that is, weapons which can be armed with either nuclear or conventional warheads. An example of such a system is the cruise missile, whose relatively low cost and potentially high accuracy and flexibility make it suitable for delivery of conventional explosives as well as nuclear warheads. An agreement which attempted to ban only nuclear warheads on cruise missiles while permitting conventional ones would present more serious technical verification problems. It has even been suggested that the Minuteman missiles displaced by the new MX missiles could be redeployed to bases in the United Kingdom and armed with conventional warheads instead of their current nuclear payload. While the US Defence Department spokesman who announced this proposal was confident that "we can solve verification issues,"[220] a certain amount of scepticism on this question is probably warranted.

At the same time these increased difficulties should not be exaggerated. Arming cruise or Minuteman missiles with nuclear warheads requires more elaborate and distinctive storage and support facilities which are vulnerable to detection.[221] In addition, this, like all efforts to cheat, faces the risk of exposure by leaks, spies or defectors. Clandestinely deploying nuclear warheads would inevitably involve many people, and the difficulty of keeping secrets increases rapidly with the number of people involved.

Efforts to develop anti-satellite (ASAT) weapons illustrate another technical problem in verification. The United States is currently developing a so-called 'direct-ascent' ASAT system which uses a small missile launched from an ordinary fighter aircraft. The testing of such a weapon is observable by Soviet national technical means, so a ban on the testing of such devices would be relatively easy to verify.[222] But once the weapon is developed and deployed its small size and non-distinctive deployment mode would make verification of a limit or ban on its deployment impossible. Such problems of timing are extremely important in verification, a lesson that has already been learned in the experience with multiple independently targetable re-entry vehicles (MIRVs). An agreement to ban the testing of multiple warhead missiles when they were under development in the 1960s would have been easily verifiable. Now that they are developed and deployed the problem of verifying limits on their numbers and qualitative capabilities is far more difficult.

To these examples of the tendency for qualitative weapon developments to outrun monitoring capabilities must be added the important class of weapon for which national technical means of verification have always had, and are certain to continue to have, extremely limited application. In this class fall chemical and biological weapons as well as the production or diversion of small quantities of nuclear explosives such as plutonium or highly enriched uranium. None of these activities is characterized by the kinds of distinctive and visible signatures associated with large missiles, submarines, aircraft or radars. Any attempt to monitor the production or stockpiling of such weapons or materials must inevitably involve more intrusive and politically sensitive measures than those associated with satellites or seismographs. While some technical measures can aid in the process and are certainly worthy of further study, the search for a purely technical solution to the problem of control of dangerous chemical, biological or nuclear materials is doomed to failure.

There is one area in which the capabilities of monitoring technology have developed faster than techniques for evasion: the monitoring of nuclear explosions. Seismological instruments, data analysis and information processing allow very reliable detection of nuclear explosions, and all the currently discussed schemes for clandestine testing seem highly implausible, especially when the synergistic effects of other monitoring processes are taken into account. A comprehensive nuclear test ban does seem to be verifiable down to very low yield tests (fractions of a kiloton) with a high degree of reliability. Therefore, no serious technical barriers remain to the verification of a comprehensive test ban.[223]

In summary, from a purely technical perspective it can be said with confidence that limits or bans on a substantial number of highly significant weapon systems could be verified with a high degree of reliability. If technical concerns about verifiability were the only obstacle to such agreements, there would be no reason not to have negotiated and signed them already and, in fact, a number of such agreements have been signed. The limited and threshold nuclear test bans, SALT I and II, and other treaties have been negotiated largely because of the existence of these national technical means.

Unfortunately, technological trends seem to be moving in a direction away from such agreements. One analysis of future developments in the 1980s concluded that "the . . . direction of weapons technology is . . . away from, not toward, greater certainty in surveillance".[224] These trends have caused the beginnings of a re-evaluation of the concept of arms control in the United States, not only among those who have traditionally been critical of it, but by those who were previously identified with efforts to achieve agreements. Former officials of both the Nixon and Carter Administrations have recently questioned the usefulness of the SALT approach largely on their assessment of the technological trends referred to above.[225] It is now being suggested that 'informal' restraints be agreed to under which each side would rely on purely unilateral verification measures and decide unilaterally how to deal with ambiguous or incriminating evidence.

Such a reversion to unilateralism would represent a serious setback to what had been painfully slow but still significant progress towards greater co-operation among states in arms control, both bilaterally between the USA and USSR and internationally through the Conference on Disarmament. Such a drastic step does not seem to be warranted on purely technical grounds, since the limitations of technical surveillance measures are only one, and probably not the most important, of the factors determining the likelihood of non-compliance with arms control agreements.[226] While technological trends certainly provide grounds for serious concern about a number of weapon systems, the notion that "we have come to the end of the road with traditional arms control agreements"[227] cannot be sustained on technical grounds. Such pessimism has its roots much more in political than technological developments, and these are the subject of the next chapter.

## Notes and references

1. York, H.A., 'US air and space reconnaissance programmes', in SIPRI, *World Armaments and Disarmament, SIPRI Yearbook 1977* (Almqvist & Wiksell, Stockholm, 1977), Appendix 5A, p. 180.
2. Scoville, H., Jr, 'A leap forward in verification', in Willrich, M. and Rhinelander, J.B. (eds), *SALT—The Moscow Agreements and Beyond* (Free Press, New York, 1974), p. 162.
3. Evvard, J.C., *Limits on Observational Capabilities of Aerospacecraft*, NASA-TN-D-2933, National Aeronautics and Space Administration, Washington, DC, 1965, p. 6.
4. See orbit parameter tables in SIPRI, *Outer Space—Battlefield of the Future?* (Taylor & Francis, London, 1978), pp. 49-81; and Jasani, B. (ed.), *Outer*

*Space—A New Dimension of the Arms Race* (Taylor & Francis, London, 1982), pp. 331-40 [a SIPRI book].

5. Perry, G.E., 'Identification of military components within the Soviet space programme', in Jasani (note 4), p. 142.
6. Katz, A.H., *Verification and SALT: The State of the Art and the Art of the State* (Heritage Foundation, Washington, DC, 1979), p. 16.
7. SIPRI 1978 (note 4), p. 17.
8. Jasani (note 4), p. 332; see SIPRI, *World Armaments and Disarmament, SIPRI Yearbook 1983* (Taylor & Francis, London, 1983), p. 448.
9. Richelson, J.T., *The Keyhole Satellite Program,* School of Government and Public Administration, The American University, Washington, DC, March 1984, pp. 27-33.
10. Most discussions of satellite photography show the light passing through a focusing lens, but large, high-resolution satellite cameras almost certainly use mirrors, since these can be made much lighter than lenses.
11. Bahcall, J.N. and Spitzer, L., Jr, 'The Space Telescope', *Scientific American,* Vol. 247, No. 1, July 1982, p. 39.
12. Mordoff, K.F., 'Company ends electrical tests on telescope', *Aviation Week & Space Technology,* Vol. 121, No. 13, 24 September 1984, p. 56.
13. Ondrjeka, R., 'Photoreconnaissance', in Tsipis, K., Hafemeister, D. and Janeway, P. (eds), *Arms Control Verification: The Technologies that Make it Possible* (Pergamon, Elmsford, NY, 1985).
14. Gelb, L., 'Keeping an eye on Russia', *New York Times Magazine,* 29 November 1981, p. 148.
15. Bahcall & Spitzer (note 11), p. 42.
16. Whitney, W.M., Mackin, R.J., Jr and Spear, A.J., *A Forecast of Space Technology; 1980–2000, Part Three: Management of Information,* NASA SP-387, National Aeronautics and Space Administration, Washington, DC, January 1976, p. 3-13.
17. Fante, R.L., 'Electromagnetic beam propagation in turbulent media', *Proceedings of the IEEE,* Vol. 63, No. 12, December 1975, p. 1669.
18. Richelson, J.T., *United States Strategic Reconnaissance: Photographic/Imaging Satellites,* ACIS Working Paper No. 38, Center for International and Strategic Affairs, UCLA, Los Angeles, May 1983, p. 13.
19. Bahcall & Spitzer (note 11), p. 42.
20. Hafemeister, D., 'Advances in verification technology', *Bulletin of the Atomic Scientists,* Vol. 41, No. 1, January 1985, p. 36.
21. Colby, W., in *Military Implications of the Treaty on the Limitation of Strategic Offensive Arms and Protocol Thereto (Salt II Treaty),* Hearings before the Committee on Armed Services (SASC Hearings), US Senate, 96th Congress, First Session (US Government Printing Office, Washington, DC, 1979), Part 3, 9, 10, 11 and 16 October, p. 1015.
22. Previously published in Jasani (note 4), p. 47.
23. Scoville (note 2).
24. Scoville (note 2), p. 172.
25. 'Satellite pictures show Soviet CVN towering above Nikolaiev shipyard', *Jane's Defence Weekly,* Vol. 2, No. 5, 11 August 1984, pp. 171-73.
26. Hudson, R.D. and Hudson, J.W., 'The military applications of remote sensing by infrared', *Proceedings of the IEEE,* Vol. 63, No. 1, January 1975, p. 123.
27. Sabins, F.F., Jr, *Remote Sensing: Principles and Interpretation* (W.H. Freeman, San Francisco, 1978), p. 121.
28. Kellogg, W.W., Buettner, K.J.K. and May, E.C., *Meteorological Satellite Obser-*

*vation of Thermal Emission,* RM-4392-NASA, Rand Corp., Santa Monica, CA, December 1964, p. 47.

29. Hudson & Hudson (note 26), p. 121.
30. Cannon, T.M. and Hunt, B.R., 'Image processing by computer', *Scientific American,* Vol. 245, No. 4, October 1981, p. 215; see also Sabins (note 27), pp. 49-50.
31. Greenwood, T., 'Reconnaissance and arms control', *Scientific American,* Vol. 228, No. 2, February 1973, p. 19.
32. Hudson & Hudson (note 26), p. 120.
33. Sabins (note 27), p. 120.
34. Constant, J.N., *Fundamentals of Strategic Weapons: Offense and Defence Systems,* Part 2 (Martinus Nijhoff, The Hague, 1981), p. 350.
35. 'USAF admits weather satellite mission', *Aviation Week & Space Technology,* Vol. 98, No. 11, 12 March 1973, p. 18.
36. 'Defense infrared payload loaded in Shuttle Orbiter', *Aviation Week & Space Technology,* Vol. 116, No. 24, 14 June 1982, p. 21.
37. Von Hippel, F. and Levi, B., *Controlling the Source: Verification of a Cutoff in the Production of Plutonium and High-Enriched Uranium for Nuclear Weapons,* Report PU/CEES-167, Princeton University Center for Energy and Environmental Studies, Princeton, NJ, July 1984, p. 25.
38. Hudson & Hudson (note 26), p. 124.
39. Kellogg, Buettner & May (note 28), p. 33.
40. Hudson & Hudson (note 26), p. 112.
41. Hudson & Hudson (note 26), p. 110.
42. Schultz, J. and Russell, D., 'New staring sensors', *Defence Electronics,* Vol. 16, No. 7, July 1984, p. 45.
43. Schultz & Russell (note 42), pp. 46-47.
44. Amelio, G.F., 'Charge-coupled devices', *Scientific American,* Vol. 230, No. 2, February 1974, pp. 22-31
45. Klass, P.J., 'Interest rises in infrared search, track', *Aviation Week & Space Technology,* Vol. 117, No. 9, 30 August 1982, p. 68.
46. Hudson & Hudson (note 26), p. 112.
47. Smith, B.A., 'Tests confirm Teal Ruby design', *Aviation Week & Space Technology,* Vol. 115, No. 1, 6 July 1981, pp. 23-24.
48. Smith, B.A., 'Problems double costs of Teal Ruby', *Aviation Week & Space Technology* (note 47), pp. 23-24.
49. Elachi, C., 'Spaceborne imaging radar: geologic and oceanographic applications', *Science,* Vol. 209, No. 4461, 5 September 1980, p. 1073.
50. Brookner, E., 'A review of array radars', *Microwave Journal,* Vol. 24, No. 10, October 1981, pp. 25-28; see also Brookner, E., *Radar Technology* (Artech House, Dedham, MA, 1977), chapter 1.
51. Brookner, *Microwave Journal* (note 50), p. 25.
52. Rihaczek, A.W., *Principles of High-Resolution Radar* (McGraw Hill, New York, 1969), pp. 51-56.
53. Figure from Brookner, E., 'Radar of the '80s and beyond', Raytheon Company Equipment Division, Wayland, MA, May 1984, p. 4.
54. *McGraw Hill Encyclopedia of Science and Technology,* Vol. 11 (McGraw Hill, New York, 1982), p. 230.
55. *Jane's Weapon Systems, 1982–83* (Macdonald & Co, London, 1982), p. 499.
56. Aspin, L. and Kaplan, F.M., 'Verification in perspective', in Potter, W.C. (ed.), *Verification and SALT: The Challenge of Strategic Deception* (Westview Press, Boulder, CO, 1980), pp. 180-81.

57. SIPRI, *World Armaments and Disarmaments, SIPRI Yearbook 1980* (Taylor & Francis, London, 1980), p. 298.
58. Stein, K.J., 'Backscatter radar unit enters production phase', *Aviation Week & Space Technology,* Vol. 117, No. 7, 16 August 1982, p. 68.
59. Stein (note 58), p. 69.
60. SIPRI (note 57), p. 295.
61. SIPRI (note 57), p. 295.
62. Brookner, *Microwave Journal* (note 50), p. 28.
63. Clear explanations of the principles of synthetic aperture radar are also given in Brown, W.M. and Porcello, L.J., 'An introduction to synthetic-aperture radar', *IEEE Spectrum,* Vol. 6, No. 9, September 1969, pp. 52-62; Jensen, H. *et al.*, 'Side-looking airborne radar', *Scientific American,* Vol. 237, No. 4, October 1977, pp. 84-95; and Brookner in *Radar Technology* (note 50), chapters 16-18.
64. Sherwin, C.W., Ruina, J.P. and Rawcliffe, R.D., 'Some early developments in synthetic aperture radar systems', *IRE Transactions on Military Electronics,* Vol. MIL-6, April 1962, p. 113; Brookner, *Radar Technology* (note 50), chapter 17.
65. Elachi (note 49), p. 1074.
66. Ulsamer, E., 'Military imperatives in space', *Air Force,* Vol. 68, No. 1, January 1985, pp. 92-93.
67. Whitney, Mackin & Spear (note 16), p. 3-12.
68. Chapman, J., 'New eyes for NATO', *Military Technology,* October 1983, p. 42.
69. Kinnucan, P., 'Machines that see', *High Technology,* Vol. 3, No. 4, April 1983, pp. 30-36; Rhea, J., 'DARPA pushes strategic computing technology', *Defense Electronics,* Vol. 16, No. 6, June 1984, pp. 112-30.
70. Constant (note 34), p. 356.
71. Waldrop, M.M., 'The infrared astronomy satellite (II)', *Science*, Vol. 221, 1 July 1983, pp. 43-44.
72. Constant (note 34), p. 343.
73. Hunt, B.R., 'Image enhancement by digital computer', in Hafemeister, D.W. and Schroeer, D. (eds), *Physics, Technology and the Nuclear Arms Race,* AIP Conference Proceedings No. 104 (American Institute of Physics, New York, 1983), p. 135.
74. Oppenheim, A.V., Schafer, R.W. and Stockham, T.G., 'Nonlinear filtering of multiplied and convolved signals', *Proceedings of the IEEE,* Vol. 56, No. 8, August 1968, pp. 1264-92.
75. Hunt (note 73).
76. Pratt, W.K., *Digital Image Processing* (Wiley, New York, 1978), chapter 12.
77. Hunt (note 73), pp. 147-48.
78. A binary number is related to the 'base' 2 in the same way that an ordinary decimal number is related to the base 10. Binary numbers are therefore strings of zeros and ones. For example, the decimal number 2 is written in binary form as 10, and the decimal 4 is written as 100.
    Each digit of a binary number is called a 'bit', so an 8-bit binary number would encompass all the decimal numbers from 0 (written 00000000) to 255 (written 11111111).
79. Kristian, I. and Blouke, M., 'Charge-coupled devices in astronomy', *Scientific American,* Vol. 247, No. 4, October 1982, p. 54.
80. Boehm, B.W., *Some Information Processing Implications of Air Force Space Missions: 1970–1980,* Rand Memorandum RM-6213-PR, January 1970, Rand Corp., Santa Monica, CA, p. 13.
81. Pratt (note 76), pp. 319-21.
82. Hunt (note 73), p. 145.

83. Pratt (note 76), pp. 311-18; Sabins (note 27), pp. 248-53.
84. Sabins (note 27), p. 250.
85. Peli, T. and Lim, J.S., 'Adaptive filtering for image enhancement', *Optical Engineering,* Vol. 21, No. 1, January/February 1982, pp. 108-12.
86. Ebersole, J.F., 'Optical image subtraction', *Optical Engineering,* Vol. 14, No. 5, September/October 1975, pp. 436-47.
87. Quatieri, T.F., *Object Detection by Two-Dimensional Linear Prediction,* Technical Report 632, Massachusetts Institute of Technology Lincoln Laboratory, Lexington, MA, 28 January 1983.
88. See Pratt (note 76) for a more complete survey.
89. Hunt (note 73), p. 143; Woods, R.E. and Gonzalez, R.C., 'Real-time digital image enhancement', *Proceedings of the IEEE,* Vol. 69, No. 5, May 1981, pp. 643-54.
90. Ebersole, J.F. and Wyant, J.C., 'Real-time optical subtraction of photographic imagery for difference detection', *Applied Optics,* Vol. 15, No. 4, April 1976, pp. 871-76.
91. Boehm (note 80), p. 12.
92. Whitney, Mackin & Spear (note 16), p. 3-83.
93. Robinson, A.L., 'Experimental memory chips reach 1 megabit', *Science,* Vol. 224, No. 4649, 11 May 1984, pp. 590-92.
94. 'New circuits expected to exceed projections', *Aviation Week & Space Technology,* Vol. 121, No. 5, 30 July 1984, p. 47.
95. Preston, G.W., 'The very large scale integrated circuit', *American Scientist,* Vol. 71, No. 5, September/October 1983, pp. 466-73.
96. Brookner, E. and Hunt, B.R., 'Synthetic aperture radar processing', *Trends and Perspectives in Signal Processing,* Vol. 2, No. 1, January 1982, p. 3; see also Elachi (note 49), pp. 1073-82.
97. Cutrona, L.J. *et al.*, 'On the application of coherent optical processing techniques to synthetic aperture radar', *Proceedings of the IEEE,* Vol. 54, No. 8, August 1966, pp. 1026-32.
98. Jensen, H. *et al.*, 'Side-looking airborne radar', *Scientific American,* Vol. 237, No. 4, October 1977, p. 94.
99. Brookner & Hunt (note 96), p. 3.
100. Bicknell, T., 'Real-time digital processing of SAR data', Spaceborne Imaging Radar Symposium, 17-20 January 1983, Jet Propulsion Laboratory, Pasadena, CA, 1 July 1983, pp. 96-97.
101. Brookner & Hunt (note 96), p. 4.
102. Carver, K.R., 'Spaceborne SAR sensor architecture', Spaceborne Imaging Radar Symposium (note 100), p. 19.
103. Sykes, L.R., Evernden, J.F. and Cifuentes, I., 'Seismic methods for verifying nuclear test bans', in Hafemeister & Schroeer (note 73), p. 86.
104. Dahlman, O. and Israelson, H., *Monitoring Underground Nuclear Explosions* (Elsevier, Amsterdam, 1977); Bolt, B.A., *Nuclear Explosions and Earthquakes: The Parted Veil* (W.H. Freeman, San Francisco, 1976); Douglas, A., 'Seismic source identification: a review of past and present research efforts', in Husebye, E.S. and Mykkeltveit, S. (eds), *Identification of Seismic Sources—Earthquake or Underground Explosion* (D. Reidel, Dordrecht, 1981), pp. 1-48; Sykes, L.R. and Evernden, J.F., 'The verification of a comprehensive nuclear test ban', *Scientific American,* Vol. 247, No. 4, October 1982, pp. 47-55; *Bulletin of the Seismological Society of America,* Vol. 72, No. 6, December 1982.
105. Ringdal, F. and Husebye, E.S., 'Application of arrays in the detection, location and identification of seismic events', in *Bulletin of the Seismological Society of America* (note 104), pp. S201-S224.

106. Herrin, E., 'The resolution of seismic instruments used in treaty verification research', in *Bulletin of the Seismological Society of America* (note 104), pp. S61-S67.
107. Wallerstedt, R.L., *Marine Seismic System (MSS) Deployment. Phase IV. Investigation of Techniques and Deployment Scenarios for Installation of Tri-Axial Seismometer in a Borehole in the Deep Ocean,* Report No. MSSA02-SYS-R003, Global Marine Development, Inc., Newport Beach, CA, 31 May 1983.
108. Taylor, S.R., 'The regional seismic test network', *Energy and Technology Review,* Lawrence Livermore National Laboratory, Livermore, CA, May 1983, pp. 20-29; Pomeroy, P.W., Best, W.J. and McEvilly, T.V., 'Test ban treaty verification with regional data—a review', in *Bulletin of the Seismological Society of America* (note 104), pp. S89-S129; Blandford, R.R., 'Seismic discrimination problems at regional distances', in Husebye & Mykkeltveit (note 104), pp. 695-740.
109. See for example Sykes & Evernden (note 104), p. 53 and Hannon, W.J., Jr, 'Seismic verification of a comprehensive test ban', *Energy and Technology Review* (note 108), pp. 50-65.
110. Dahlman & Israelson (note 104), pp. 56 and 67.
111. Dahlman & Israelson (note 104), p. 136.
112. Dahlman & Israelson (note 104), pp. 66 and 69.
113. Pomeroy, Best & McEvilly (note 108), p. S111 and Blandford (note 108).
114. Douglas, A. *et al.*, 'Earthquakes that look like explosions',*Geophys. J. R. Astr. Soc.*, Vol. 36, 1974, pp. 227-33.
115. See, for example, Dahlman & Israelson (note 104), chapter 9.
116. Sykes, Evernden & Cifuentes (note 103), p. 96.
117. Kerr, R.A., 'The deepest hole in the world', *Science,* Vol. 224, 29 June 1984, p. 1420.
118. Sykes, Evernden & Cifuentes (note 103), p. 106.
119. Sykes, Evernden & Cifuentes (note 103), p. 94.
120. Sykes & Evernden (note 104), p. 49.
121. Sykes, Evernden & Cifuentes (note 103), p. 108.
122. Bache, T.C., 'Estimating the yield of underground nuclear explosions', *Bulletin of the Seismological Society of America* (note 104), p. S133.
123. Bolt (note 104), p. 39.
124. Dahlman & Israelson (note 104), pp. 306-16.
125. Sykes, L.R. and Cifuentes, I.L., *Yields of Soviet Underground Nuclear Explosions from Seismic Surface Waves: Compliance with Threshold Test Ban Treaty,* Lamont-Doherty Geological Observatory and Department of Geological Sciences, Columbia University, NY, 7 December 1983, p. 3.
126. Sykes & Cifuentes (note 125), p. 15.
127. Dahlman & Israelson (note 104), chapter 13.
128. Glenn, L.A., 'Verification limits for test-ban treaty', *Nature*, Vol. 310, 2 August 1984, pp. 361-62.
129. Evernden, J.F. and Archambeau, C.B., 'Seismological aspects of monitoring a comprehensive test ban treaty', in Tsipis, Hafemeister & Janeway (note 13).
130. Evernden & Archambeau (note 129).
131. Husebye & Mykkeltveit (note 104), p. xii.
132. Ringdal & Husebye (note 105), p. S203.
133. Dahlman & Israelson (note 104), pp. 149-56.
134. Dahlman, O., private communication, 21 August 1984.
135. Ringdal, F., 'Automatic processing methods in the analysis of data from a global seismic network', in Husebye & Mikkeltveit (note 4), p. 807.

136. Singer, S., 'The Vela satellite program for detection of high-altitude nuclear detonations', *Proceedings of the IEEE,* Vol. 53, No. 12, December 1965, p. 1935.
137. Jungk, R., *Brighter than a Thousand Suns* (Harcourt Brace Jovanovich, New York, 1958).
138. Singer (note 136).
139. Argo, H.V., 'Monitoring atmospheric nuclear explosions', in Tsipis, Hafemeister & Janeway (note 13).
140. Argo (note 139). The origins of the term 'bhangmeter' are obscure, but according to one source it was coined in the early years of the US nuclear testing programme as a pun on the words 'bang' for an explosion and 'bhang', a narcotic drug made from hemp (private communication from H.V. Argo).
141. Argo (note 139).
142. Singer (note 136), p. 1942.
143. Argo (note 139).
144. Bell, J., 'Navigation for everyman, and his bomb', *New Scientist,* Vol. 104, No. 1425, 11 October 1984, pp. 36-40.
145. Powers, T., 'The ears of America', *New York Review of Books,* 3 February 1983, p. 12.
146. Bamford, J., *The Puzzle Palace: America's National Security Agency and its Special Relationship with Britain's GCHQ* (Sidgwick and Jackson, London, 1983), Chapter 6.
147. Powers (note 145).
148. Holloway, D., *The Soviet Union and the Arms Race* (Yale University Press, New Haven, CT, 1983), p. 18.
149. Pearson, D., 'KAL 007: what the US knew and when we knew it', *The Nation,* Vol. 239, No. 4, 18-25 August 1984, pp. 113-14.
150. Bamford (note 146), chapter 5.
151. Greenwood, T., *Reconnaissance, Surveillance and Arms Control,* Adelphi Paper No. 88 (IISS, London, 1972), p. 23.
152. SIPRI (note 4), pp. 39-42.
153. Moncrief, F.J., 'SALT verification: how we monitor the Soviet arsenal', *Microwaves,* September 1979, p. 42.
154. Bamford (note 146), p. 197.
155. *Two-line Orbital Elements,* NASA–Goddard Space Flight Center, Greenbelt, MD, daily until June 1983.
156. Treaty between the USA and the USSR on the limitation of Anti-Ballistic Missile Systems, reprinted in Goldblat, J., *Agreements for Arms Control: A Critical Survey* (Taylor & Francis, London 1982) [a SIPRI book]. See the Agreed Interpretation F (p. 199) and Article II (p. 197).
157. Compliance with SALT I Agreements, US Department of State, Bureau of Public Affairs, Special Report No. 55, Washington, DC, July 1979, pp. 2-3.
158. Moncrief (note 153), p. 46.
159. Perry, G.E. and Flagg, R.S., 'Telemetry from Russian Spacecraft', *Journal of the British Interplanetary Society,* Vol. 23, 1970, pp. 451-64; Perry, G.E., 'Identification of military components within the Soviet space programme', in Jasani (note 4), pp. 135-54.
160. Perry & Flagg (note 159), p. 458.
161. Moncrief (note 153), p. 46.
162. Hussain, F., *The Impact of Weapons Test Restrictions,* Adelphi Paper No. 165 (IISS, London, 1981), p. 45.
163. SASC Hearings (note 21), Part 1, p. 138.
164. Treaty between the USA and the USSR on the Limitation of Strategic Offensive

Arms (SALT II Treaty), Second Common Understanding to Paragraph 3 of Article XV, in Goldblat (note 156), p. 280.

165. Hussain (note 162), p. 39.
166. SIPRI, (note 57), pp. 293-300.
167. *The SALT II Treaty,* Hearings before the Committee on Foreign Relations (SFRC Hearings), US Senate, 96th Congress, First Session (US Government Printing Office, Washington, DC, 1979), Part 2, p. 275.
168. Talbot, S., *End Game: The Inside Story of SALT II* (Harper & Row, New York, 1979), p. 266.
169. Blair, B.G. and Brewer, G.D., 'Verifying SALT agreements', in Potter, W.C. (ed.), *Verification and SALT—The Challenge of Strategic Deception* (Westview Press, Boulder, CO, 1980), p. 33.
170. Moncrief (note 153), p. 47.
171. Ball, D., *The Rhyolite Programme,* Reference Paper No. 86, Research School of Pacific Studies, Canberra, November 1981, p. 4.
172. Moncrief (note 153), p. 47.
173. Blair, B.G., 'Reconnaissance satellites', in Jasani (note 4), p. 133.
174. Biddle, W., 'Initial orbit of new satellite different from usual', *New York Times,* 26 April 1985, p. A19; see also Richelson, J.T., *The US Intelligence Community* (Ballinger, Cambridge, MA, 1985), p. 122.
175. Marriott, J., 'Tactical crypt: the art of secure communications', *Military Technology,* Vol. 5, No. 26, August 1981, pp. 12-14.
176. Sullivan, J., 'Cryptography: securing computer transmissions', *High Technology,* Vol. 3, No. 11, November 1983, p. 31-36.
177. Kolata, G., 'Factoring gets easier', *Science,* Vol. 222, 2 December 1983, pp. 999-1001; 'Another promising code falls', *Science,* 16 December 1983, Vol. 222, p. 1224.
178. NPT Newsletter, *IAEA Bulletin,* Vol. 26, No. 1, March 1984, p. 36.
179. *The Structure and Content of Agreements between the Agency and States Required in Connection with the Treaty on the Non-Proliferation of Nuclear Weapons,* INFCIRC/153, IAEA, Vienna, May 1971, para. 28.
180. Barnard, C.I. *et al.* 'A report on the international control of atomic energy, 16 March 1946', *Peaceful Nuclear Exports and Weapons Proliferation: A Compendium,* Committee on Government Operations, US Senate (US Government Printing Office, Washington, DC, April 1975), p. 167 [emphasis in original].
181. Grümm, H., 'IAEA safeguards—where do we stand today?', *IAEA Bulletin,* Vol. 21, No. 4, August 1979, p. 35.
182. *IAEA Safeguards Glossary* (IAEA, Vienna, 1980), para. 89.
183. Jennekens, J., 'International safeguards—the quantification issue', *IAEA Bulletin,* Vol. 22, No. 3/4, August 1980, pp. 42-43.
184. IAEA (note 182), para. 90.
185. IAEA (note 182), para. 87.
186. 'The present status of IAEA safeguards on nuclear fuel cycle facilities', *IAEA Bulletin* (note 183), p. 24.
187. Krass, A.S., Boskma, P. , Elzen, B. and Smit, W.A., *Uranium Enrichment and Nuclear Weapon Proliferation* (Taylor & Francis, London, 1983) [a SIPRI book], p. 45.
188. IAEA (note 182), paras 109-49.
189. IAEA (note 179), paras 31-32.
190. 'The safeguards analytical laboratory: its functions and analytical facilities', *IAEA Bulletin,* Vol. 19, No. 5, October 1977, pp. 38-47.

191. IAEA, *IAEA Safeguards: Techniques and Equipment*, IAEA/SG/INF/5 (IAEA, Vienna, 1984), pp. 2-11.
192. Prial, F.J., 'Sweden releases Soviet sub; finds signs of nuclear arms', *New York Times,* November 1981, pp. 1, 12.
193. *The Annual Report for 1983* (IAEA, Vienna, 1984), p. 59.
194. IAEA (note 193), pp. 68-69.
195. IAEA (note 182), paras 119, 135-37.
196. IAEA (note 182), para. 12.
197. IAEA (note 193).
198. Imber, M.F., 'Arms control verification: the special case of IAEA-NPT "special inspections"', *Arms Control,* Vol. 3, No. 3, December 1982, pp. 66-67.
199. IAEA (note 182), para. 216.
200. Klik, F., 'Field experience of safeguards inspectors', *IAEA Bulletin,* Vol. 23, No. 4, December 1981, p. 20.
201. 'Surveillance and containment measures to support IAEA safeguards', *IAEA Bulletin* (note 190), p. 22.
202. IAEA (note 193), p. 59.
203. IAEA (note 182), para. 217.
204. *IAEA Bulletin* (note 201), pp. 21, 24.
205. *IAEA Bulletin* (note 201), p. 22.
206. Pieragostini, K., 'RECOVERing verification', *Arms Control Today,* Vol. 12, No. 11, December, 1982, pp. 4-5.
207. United States of America, United Kingdom and Australia, 'Technical evaluation of "RECOVER" techniques for CW verification', Committee on Disarmament documents CD/271, CD/CW/WP.32, Committee on Disarmament, Geneva, 7 April 1982, p. 1.
208. Pieragostini (note 206), p. 5.
209. Sheaks, O.J., US Arms Control and Disarmament Agency, private communication, 15 October 1984.
210. Fischer, D.A.V., 'Safeguards—a model for general arms control?', *IAEA Bulletin,* Vol. 24, No. 2, June 1982, pp. 45-49.
211. United States of America, 'Illustrative on-site inspection procedures for verification of chemical weapons stockpile destruction', Committee on Disarmament document CD/387, Committee on Disarmament, Geneva, 6 July 1983.
212. United States of America, 'The declaration and interim monitoring of chemical weapons stockpiles', Conference on Disarmament document CD/516, Conference on Disarmament, 13 July 1984.
213. SIPRI, *World Armaments and Disarmament, SIPRI Yearbook 1978* (Taylor & Francis, London, 1978), pp. 73-79.
214. Scoville (note 2), pp. 178-79.
215. 'Debate on a comprehensive nuclear weapons test ban: pro and con', *Physics Today,* Vol. 36, No. 8, August 1983, pp. 24-25.
216. Kessler, R. and Knapik, M., 'Argentine enrichment pronouncement characterized as startling', *Nucleonics Week,* Vol. 24, No. 47, 24 November 1983, pp. 1-2.
217. Katz (note 6), p. 32.
218. Kahan, J.H., *Security in the Nuclear Age* (Brookings Institution, Washington, DC, 1975), p. 333.
219. Biddle, W., 'Senate defeats a proposal to ban sea-based nuclear cruise missiles', *New York Times,* 20 June 1984, p. 1.
220. 'Minuteman studied for airfield attack', *Aviation Week & Space Technology,* Vol. 121, No. 2, 9 July 1984, p. 21.
221. Rowny, E. in SASC Hearings (note 21), Part 2, p. 687.

222. FAS Public Interest Report, Vol. 36, No. 9, November 1983, pp. 13-16.
223. Sykes & Evernden (note 104), p. 55.
224. Kincade, W.H., 'Over the technological horizon', *Daedalus,* Vol. 110, No. 1, Winter 1981, p. 116.
225. Toth, R.C., 'Some experts doubt value of arms pacts', *Los Angeles Times,* 23 May 1984, p. 1.
226. Chayes, A., 'An inquiry into the workings of arms control agreements', *Harvard Law Review,* Vol. 85, No. 5, March 1972, p. 968.
227. Toth (note 225).