Remote sensing

Definition and Brief History of Remote Sensing

There are a number of definitions of remote sensing and one of the most commonly accepted is as follows: “Remote sensing is the science and art of obtaining information about an object, area, or phenomenon through the analysis of data acquired by a device that is not in contact with the object, area, or phenomenon under investigation” [1]. A working definition of remote sensing refers to the acquisition of information from digital geospatial data acquired from an overhead perspective, using sensors that sample and record electromagnetic radiation in one or more regions of the electromagnetic spectrum that is reflected or emitted from the surface of the earth. The geospatial data might include analog data, other force fields (seismic, acoustic, or gravitational), and nonterrestrial applications (lunar and planetary surfaces and astronomical investigations). Remote sensing can be used to detect the presence, or confirm the absence, of a specific target class: mineral potential, insect infestations, militarily significant activity, environmental releases from industrial facilities, indications of prehistoric land use, habitat for a given species, urban monitoring, fire detection, or flood prediction [2]. The use of remote sensing techniques to map vector-borne diseases has evolved significantly over the past decades [3].

Remote sensing can trace its origins back to the earliest overhead photography, carried out from balloons during the American Civil War. Aerial photography was well established by the end of the First World War, in a decade that also saw the beginning of the science of photogrammetry. Significant developments in the period around the Second World War included the development of technology to make measurements in the IR and microwave regions of the electromagnetic spectrum. These technologies became available for civilian use in the 1950s and 1960s. The first meteorological satellite was launched in 1960, and the term “remote sensing” came into widespread use during this decade. Landsat1, launched in 1972, was the first of a series of Earth resource satellites that today provide repetitive multispectral coverage of the earth in digital format. Hyperspectral instruments capable of recording electromagnetic radiation in tens to hundreds of spectral bands simultaneously were deployed in the 1980s on airborne platforms, and became spaceborne at the end of the century. As more types of coverage become available at increased spatial and spectral resolutions (see Satellite data), almost every natural science – ecology, geology, hydrology, oceanography, atmospheric science – and many areas of economic importance – agriculture, mineral exploration, urban planning, environmental monitoring – are developing new ways to exploit this huge reservoir of data. At the same time, remote sensing and image processing have been constantly evolving in the last decade [4, 5].

Types of Remote Sensing

The overhead perspective mentioned in the working definition of remote sensing given above comes from deploying remote sensing instruments on airplanes or satellites. The former is airborne remote sensing, whereas the latter is known as spaceborne remote sensing. As these platforms move along their flight paths, the instruments scan typically across a swath perpendicular to the direction of motion. The data from a series of such swaths forms a two-dimensional image. Most earth observation satellites have near-polar sun-synchronous orbits, providing repetitive coverage with a period on the order of 2 to dozens of days of each point on earth at the same local time of day. Once launched, there is limited control over data collection by these satellites, although modern instruments have an increasing number of programmable features. Airborne coverage can be targeted more specifically to meet customer demands. Geosynchronous satellites, whose position above a point on the surface of the earth is fixed approximately, are used primarily for meteorological and communications purposes.

The airborne and spaceborne instruments measure electromagnetic radiation, which has been modulated by the surface of the earth and by the atmosphere between the radiation source, the surface, and the sensor. The radiation source may be the sun, materials
Remote sensing

on the surface or in the atmosphere, or the instrument itself. Passive remote sensing employs photography and multispectral scanners that record naturally occurring electromagnetic radiation that is reflected or emitted from the target. Active remote sensing supplies energy from its own sensor/instrument to illuminate the scene, and then record the amount of radiant flux scattered back toward the sensor system.

Some parts of the electromagnetic spectrum are not usable for earth-based remote sensing applications because gases in the atmosphere (primarily ozone, carbon dioxide, and water vapor) absorb most or all of the energy at certain frequencies. Reflected solar energy is measured in the visible/near-IR part of the spectrum, about 0.4–2.5 μm. The thermal signal from sources such as fires is strongest in the mid-IR region, 3–5 μm, while reradiated energy from the sun and cooler atmospheric plumes is measured in the thermal IR region, 8–14 μm. Microwave energy is recorded generally by active sensors such as radar (an acronym for Radio Detection And Ranging). Remote sensing radars emit pulsed coherent radiation at wavelengths between 1 and 30 cm, and measure the reflection of that energy from the earth’s surface or from reflectors in the atmosphere. Energy at these frequencies can penetrate cloud cover and supplement the incomplete observations possible at higher frequencies. Light detection and ranging (lidar) applies the same principles in the optical and near-IR region of the spectrum, and is particularly useful for atmospheric research, measuring forest structure [6] and surveying urban topography [7].

Thematic Mapping

The ultimate goal of remote sensing is to extract information from remotely sensed data about the material properties of the earth’s surface and of the atmosphere, together with their geographical relationships. The first and most common product of data analysis is a thematic map, that is, a classification of the areas and features in the scene. Examples of the types of classes of interest include soil and rock types, classification of crops and forests, delineation of ice and snow cover, and identification of land use in urban and suburban areas. The task of thematic mapping is the classification of pixels or regions in the scene into information classes that are meaningful for the task at hand. Thematic mapping can be accomplished by using supervised or unsupervised approaches [1, 8, 9], and either of these has its advantages and disadvantages [10].

Supervised classification requires that a training dataset be available, consisting of pixels whose classification is known. As the number of dimensions in the data space (i.e., the number of spectral bands) increases, the number of training samples must also increase. Rules of thumb proposed by various authors indicate that the number of training samples per class should be 10 to 100 times the number of discriminating variables, otherwise, the overall performance of a classifier can actually degrade (this is referred to as the Hughes phenomenon, after Hughes [11]). Landgrebe and coworkers [12] have experimented with several methods for alleviating this problem, including manual identification of training samples from spectral information alone and the addition of unlabeled samples using an expectation–maximization (EM) algorithm. Other problems with training sets for remote sensing data include the difficulty of finding unmixed pixels and of ensuring that all of the classes of interest are represented. Mixed pixels can be removed from the training set [13], but this exacerbates the problem of obtaining a sufficiently large training set. The presence of untrained classes may degrade the performance of a classifier (e.g., Ref. 14).

Statistical algorithms include methods based on maximum likelihood or nearest-neighbor decision rules and classification trees (see Regression trees). Maximum likelihood methods may require the estimation of separate covariance matrices for each class or may use a pooled covariance matrix (see Covariance matrix estimation (estimated parameters)). A particularly simple version that assumes the covariance matrix is the identity is called the minimum distance classifier in the remote sensing literature. Other methods include penalized discriminant analysis that takes the high level of correlation among the variables into account [15], a Bayesian interpretation of maximum likelihood that allows updating prior class probabilities and obtaining unbiased estimates of class coverages, and classification tree algorithms. While some of the literature suggests that neural nets may be able to learn on smaller datasets without dimension reduction, training sets must still be representative in order to obtain good results, and feature extraction may increase the interpretability of results and shorten the time required to train the network. In terms of accuracy, however, neural
networks have been shown to perform favorably in comparison with most statistical methods in several studies (e.g., Ref. 16).

At the spatial scale of most remote sensing systems, a pixel contains a mixture of materials or land cover types. Statistical and neural net classifiers produce typically hard classification results, that is, one pixel is assigned to one class. Both types of methods could provide more information: posterior classification probabilities for each class are available from many statistical classifiers, and activation weights for each output node can be observed in a neural net. However, it is not always clear how to relate uncertainty in classification to the fraction of land cover present in a given pixel. Fuzzy classifiers that allocate every pixel to every class with varying grades of membership have also been proposed to address this problem more explicitly. Foody and Arora [17] point out that the mixed pixel problem needs to be accommodated not only in reporting classification results but also in training and error estimation.

Unsupervised classification approaches do not utilize training data as the basis for classification. The basic premise of unsupervised classification is that spectral values within a given land cover type should be close together in the feature space, whereas data values in different classes should be comparatively well separated [1]. One of the most commonly used algorithms for creating spectral classes is the iterative self-organizing data analysis technique algorithm (ISODATA), which uses the Euclidean distance in the feature space to assign every pixel to a cluster through a number of iterations [8]. Conventional unsupervised classification is easier to use than supervised classification but this approach has limitations for creating desired, accurate thematic maps. Because there is often no one-to-one correspondence between spectral and information classes, human errors are inevitable in labeling some of spectral classes into information classes. Lang et al. [18] reduced the human errors by using a data-driven method in labeling spectral classes.

The use of hybrid supervised/unsupervised classification methods reduces dimension effectively, and thus both the training set requirements and computation times for the supervised portion of the algorithm. In these methods, unsupervised clustering is followed by labeling of the spectral classes (for example, by supervised majority voting). Then the entire image is classified into the refined set of spectral classes by any of the standard supervised methods. Another type of hybrid algorithm classifies the regions of an oversegmented image (sample) instead of individual pixels [12]. Contextual classification methods make use of information from neighboring pixels as well as the spectral information associated with the pixel being classified. One large class of contextual methods performs postclassification revision of results from a first-pass classification based on examination of the classification of neighboring pixels [19]. Revision methods range from simple voting algorithms within windows of fixed size to iterative spectral- and class-specific procedures [20]. Probabilistic label relaxation methods start with a vector of probabilities for each pixel, such as might be produced by maximum likelihood discriminant analysis, and modify these iteratively based on the likelihood that pixels of two different classes will be found in the same neighborhood [21]. Neighborhood information is used directly in classification by the iterated conditional modes algorithm (e.g., [22]). Classification experiments suggest that a sampled-based algorithm called extraction and classification of homogeneous objects (ECHO) [12] can easily result in thematic maps with higher accuracy than pixel-based algorithms [23].

Rule-based approaches are now broadly used for classifying segmented images. An image object is defined as adjacent pixels with similar spectral values [24] and each image object could possess an intrinsic size, shape, and geographic relationship with the real-world scene component it models. So-called object-oriented classification procedures were incorporated into an image analysis software system eCognition by a German company DEFNiENS [25]. Segmented image objects can be classified by either simple algorithms such as nearest neighbor classifier or sophisticated algorithms that combine shape and spatial information, ancillary data, as well as spectral responses of image objects characteristics [15, 26]. Object-oriented classification approaches are particularly effective while using high-spatial-resolution data to generate accurate thematic maps (e.g., Ref. 27). By using the object-oriented classification approach, detailed thematic maps can be produced even from natural-color digital aerial orthophotography with limited spectral power [28]. However, some classification rule sets and workflows have become too complex, and raise as many research questions as they resolve [29].
example, the selection of scale is very important and continues to be a hot research topic in object-based classifications for remote sensing.

Quantitative Estimation

For some applications, a thematic map may be less useful than estimates of the fraction of each cover type in each pixel. Linear mixture models are suitable for spectral unmixing of individual pixels if the contributing components form a mosaic within the pixel, but not if they are in such intimate association that electromagnetic radiation interacts with more than one endmember as it is scattered from the surface. The linear mixture model takes the form $Z_i = M\pi_i + e_i$, where $M$ is an $n \times c$ matrix whose columns are the (known) $n$-dimensional spectra of the $c$ pure cover types or endmembers. For a true mixture model, the $\pi_i$ are constrained to be positive and sum to one, but when the endmember spectra are not well known, it may be useful to relax these constraints.

The endmember spectra may be derived from laboratory spectra [30] or ground-level field measurements. In practice, it may be difficult to adjust these spectra to match the signal received at a remote sensor after passing through the intervening atmosphere, so other techniques have been devised to refine the endmembers based on the data. Purely empirical methods such as convex hull analysis have been used [31], but there is no guarantee that these will provide physically interpretable results. Two-step or iterative methods that solve for both $M$ and the $\pi_i$, incorporating a priori knowledge about both the spectra and cover type abundances in the starting model and in iteration constraints, are proposed by Tompkins et al. [32] and van der Meer [33]. Maselli [34] and Roberts et al. [35] customize the selection of endmembers from a large number of candidates on a pixel-by-pixel basis. Support vector machines (SVMs) have become a promising machine learning methodology with limited training samples in remote sensing applications [36].

Hyperspectral instruments open up the possibility of identifying not only broad land cover types but also specific chemicals in a pixel or scene, using techniques that are based on the principles of reflectance spectroscopy [1, 37]. Typically, these methods require spectral libraries, which are available for minerals, soils, and vegetation types. Most spectral matching algorithms begin by approximating the continuum of the spectrum, often by the convex hull of the measurements, and normalizing the measurements by the resulting function of wavelength [38]. This isolates the absorption features from the overall reflectance properties of other components in the signal. Techniques for matching continuum-corrected spectra use binary coded data [39], angular information (the spectral angle mapper algorithm), or the complete spectrum via the cross-correlation algorithm of van der Meer and Bakker [40].

Many studies are designed to estimate changes in land cover proportions from remote sensing data [41]. Swamy and Brivio [42] combine estimates of snow-covered areas with digital elevation models to obtain input for hydrological models of seasonal and real-time runoff in the Italian Alps. Piwowar and Le-Drew [43] evaluate the potential of remote sensing data to address trends in sea ice extent. The use of remote sensing data for change detection (see Change, detecting) in these and other applications requires accurate image registration. Radiometric matching methods have been developed to make a pair of images appear to have been collected under the same atmospheric and illumination conditions. Collins and Woodcock [10] suggest that simpler image-based normalization is adequate for commonly used linear change detection techniques. For post-classification change detection, Shao and Wu [44] suggest that the consistency in resolution and analysis methods between temporal thematic mapping processes can help reduce bias in change detection.

Semiempirical models form an integral component of recent methods for using remote sensing data in the estimation of geo- and biophysical parameters such as land and sea surface temperatures, photosynthesis, and evapotranspiration on regional and global scales [45-47]. Olioso et al. [48] review a number of soil-vegetation-atmospheric transfer (SVAT) models and methods for driving them by using remotely sensed data. Such models must be designed carefully to make accurate and optimal use of remote sensing information, as illustrated for example by Gastelli-Etchegorry and Trichon [49]. Goetz et al. [50] assess production efficiency models and show that surface variables recovered from satellite observations using such models are in good agreement with field measurements across a number of ecosystems.

Encyclopedia of Environmetrics, Online © 2006 John Wiley & Sons, Ltd.
This article is © 2013 John Wiley & Sons, Ltd.
DOI: 10.1002/9780470057339.var033.pub2
Accuracy Considerations

Any thematic maps derived with remote sensing methods contain errors, which can be assessed with reference data. Congalton and Green [51] suggest commonly used accuracy assessment methods with a variety of measures or statistics, including overall accuracy as well as producer’s and user’s accuracy. The assessment of thematic maps is important because they are not always the final products from user’s perspective. Where mapping is not an end in itself, classification is usually a prerequisite to further analysis. If thematic maps can be used for various instances of decision making, errors in them can mislead decision making, causing economic and ecological losses. While thematic maps can be made easily with remote sensing technology, their misuse is found in both academic [20] and nonacademic [52] fields. Shao and Wu [44] suggest that remote sensing analysts should select appropriate classification techniques, select imagery of appropriate scales, and balance producer’s and user’s accuracy when they prepare thematic maps. Classification accuracy alone is insufficient for determining the reliability of thematic maps because there is no universally applicable standard based on which the adequacy of classification accuracy can be quantified. Any two thematic maps with the same classification accuracy do not necessarily have 100% pixel-to-pixel agreement. Users need to be aware of the sensitivity of classification accuracy to the applications of thematic maps in their specific fields. For example, the accuracy of landscape indices often has an exponential relationship with classification accuracy [53], meaning that a small difference in classification accuracy can cause a big difference in landscape index values. From this point of view, it is necessary to increase classification accuracy as much as possible.

References


Encyclopedia of Environmetrics, Online © 2006 John Wiley & Sons, Ltd.
This article is © 2013 John Wiley & Sons, Ltd.
This article was published in Encyclopedia of Environmetrics Second Edition in 2012 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470057339.var033.pub2
Remote sensing


Further Reading


(See also Classification; Machine learning; Ecosystem monitoring)

GUOFAN SHAO