Field Guide to

Hyperspectral/ Multispectral Image Processing

Xiuping Jia

SPIE Field Guides
Volume FG52

J. Scott Tyo, Series Editor

SPIE PRESS
Bellingham, Washington USA

Library of Congress Cataloging-in-Publication Data

Names: Jia, Xiuping, author.
Title: Field guide to hyperspectral/multispectral image processing / Xiuping Jia.
Description: Bellingham, Washington : SPIE Press, [2022] | Includes bibliographical references and index.
Identifiers: LCCN 2022007877 | ISBN 9781510652149 (spiral bound) | ISBN 9781510652156 (pdf)
Subjects: LCSH: Hyperspectral imaging. | Multispectral imaging. | Image processing–Digital techniques. | Remote sensing.
Classification: LCC TR267.73 J53 2022 | DDC 771–dc23/eng/20220408
LC record available at https://lccn.loc.gov/2022007877

Published by

SPIE
P.O. Box 10
Bellingham, Washington 98227-0010 USA
Phone: +1 360.676.3290
Fax: +1 360.647.1445
Email: books@spie.org
Web: http://spie.org

Printed in the United States of America.
First printing 2022.
For updates to this book, visit http://spie.org and type "FG52" in the search field.

Introduction to the Series

2022 is a landmark year for the SPIE *Field Guide* series. It is our 19th year of publishing *Field Guides*, which now includes more than 50 volumes, and we expect another four titles to publish this year. The series was conceived, created, and shaped by Professor John Greivenkamp from the University of Arizona. John came up with the idea of a 100-page handy reference guide for scientists and engineers. He wanted these books to be the type of reference that professionals would keep in their briefcases, on their lab bench, or even on their bedside table. The format of the series is unique: spiral-bound in a 5″ by 8″ format, the book lies flat on any page while you refer to it.

John was the author of the first volume, the seminal *Field Guide to Geometrical Optics*. This book has been an astounding success, with nearly 8000 copies sold and more than 72,000 downloads from the SPIE Digital Library. It continues to be one of the strongest selling titles in the SPIE catalog, and it is the all-time best-selling book from SPIE Press. The subsequent several *Field Guides* were in key optical areas such as atmospheric optics, adaptive optics, lithography, and spectroscopy. As time went on, the series explored more specialized areas such as optomechanics, interferometry, and colorimetry. In 2019, John created a sub-series, the *Field Guide to Physics*, with volumes on solid state physics, quantum mechanics, and optoelectronics and photonics, and a fourth volume on electromagnetics to be published this year. All told, the series has generated more than $1.5 million in print sales and nearly 1 million downloads since eBooks were made available on the SPIE Digital Library in 2011.

John's impact on the profession through the *Field Guide* series is immense. Rival publishers speak to SPIE Press with envy over this golden nugget that we have, and this is all thanks to him. John was taken from us all to early, and to honor his contribution to the profession through this series, he is commemorated in the 2022 Field Guides.

We will miss John very much, but his legacy will go on for decades to come.

Vale John Greivenkamp!

J. Scott Tyo
Series Editor, SPIE Field Guides
Melbourne, Australia, March 2022

Related Titles from SPIE Press

Keep information at your fingertips with these other *SPIE Field Guides*:

Image Processing, Khan M. Iftekharuddin and Abdul Awwal, Vol. FG25

Other related titles:

Fourier Transforms Using Mathematica®, Joseph. W. Goodman, Vol. PM322

Multispectral Image Fusion and Colorization, Yufeng Zheng, Erik Blasch, and Zheng Liu, Vol. PM285

Optical Interference Filters Using MATLAB®, Scott W. Teare, Vol. PM299

Optimal Signal Processing Under Uncertainty, Edward R. Dougherty, Vol. PM287

Regularization in Hyperspectral Unmixing, Jignesh S. Bhatt and Manjunath V. Joshi, Vol. SL25

Table of Contents

Table of Contents

Table of Contents

Table of Contents

Preface

Hyper/multispectral imagery in optical remote sensing is an extension of color RGB pictures. The utilized wavelength range is beyond the visible, up to the reflective shortwave infrared. Hyperspectral imaging offers higher spectral resolution, leading to many wavebands. The spectral profiles recorded reveal reflected solar radiation from Earth-surface materials when the sensor is mounted on an airborne or spaceborne platform. An inverse process using machine-learning approaches is conducted for target detection, material identification, and associated environmental applications, which is the main purpose of remote sensing.

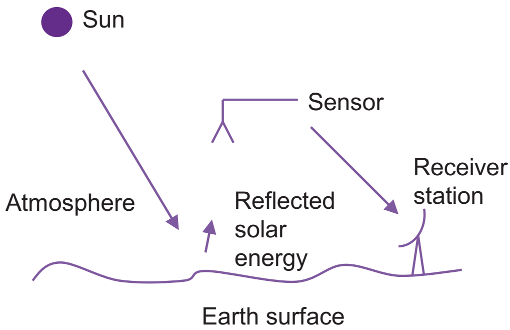
This Field Guide covers three areas: the fundamentals of remote sensing imaging for image understanding; image processing for correction and quality improvement; and image analysis for information extraction at subpixel, pixel, superpixel, and image levels, including feature mining and reduction. Basic concepts and fundamental understanding are emphasized to prepare the reader for exploring advanced methods.

I owe thanks to Professor John Richards, who introduced me to the remote sensing field and supervised my Ph.D. study on hyperspectral image classification. Over the years I learned a lot from John about critical thinking and thorough presentation.

I dedicate this Field Guide to my husband, Xichuan, my son, James, and my daughter, Jessica, who have shared and accompanied me on my learning journey in this subject for so many years and make me a proud career wife and mom.

Xiuping Jia
The University of New South Wales,
Canberra, Australia
May 2022

Spectral Coverage of Optical Remote Sensing



The wavelength used for **optical remote sensing** ranges from 0.4 to 2.5 μm. In this range, the radiation of the Sun (at a temperature of 6000 K) is strong, according to **Planck's black body radiation law**:

$$M(T,\lambda) = \frac{C_1}{\lambda^5\left[\exp\left(\frac{C_2}{\lambda T} - 1\right)\right]} \,(\mathrm{W\ m^{-2}\mu m^{-1}})$$

C_1: first radiation constant $= 3.7413 \times 10^8$ W $\mu m^4/m^2$,
C_2: second radiation constant $= 1.4388 \times 10^4$ μm K,
T: absolute radiant temperature (K), and
λ: radiation wavelength (μm).

T and λ are the two variables. Even after the attenuation is considered due to the large distance between the Sun and the Earth, the received solar radiation in the optical reflective range by the Earth is much stronger than the emission from itself (300 K). The signals collected by an optical sensor on board a satellite or aircraft are then the reflected solar radiation from the Earth surface, thanks also to the generally high atmospheric transmittance in the optical range.

When $\lambda > 3$ μm, emission from hot bodies dominates the upwelling radiation from the Earth. This phenomenon forms **thermal infrared remote sensing** (3 μm $< \lambda <$ 14 μm).

Spectral Resolution

Different materials have different spectral reflectance characteristics with different absorption bands of solar radiation, which make them discriminable from each other. Optical sensors are designed to capture images in multiple wavebands with good **spectral resolution** so that the image data can then be used for material identification and recognition. The table shows the band designation of the Landsat-8 OLI (Operational Land Imager) as an example. Landsat-9 OLI-2's bands are similar.

Spectral Band	Wavelength (nm)	Spectral Resolution (nm)
Band 1 Coastal/Aerosol	435–451	16
Band 2 Blue	452–512	60
Band 3 Green	533–590	57
Band 4 Red	636–673	37
Band 5 Near Infrared	851–879	28
Band 6 SWIR	1566–1651	85
Band 7 SWIR	2107–2294	187
Band 8 Panchromatic	503–676	173
Band 9 Cirrus	1363–1384	21

Note that Band 8 provides a **panchromatic image** with higher spatial resolution (15 m) than other **multispectral bands** (30 m) with a lower spectral resolution. There is a trade-off between the two to meet a required signal-to-noise ratio.

Hyperspectral sensors such as A**irborne Visible/Infrared Imaging Spectrometer (AVIRIS)** and **Hyperion** provide higher spectral resolution on the order of 10 nm and collect data for the complete optical spectral range from 400 to 2450 nm, leading to more than 200 bands. The rich spectral information is important for fine material and condition monitoring. **HyMap** is another hyperspectral sensor that provides a spectral resolution around 20 nm, with a better spatial resolution of 3 m, compared with the 20 m offered by AVIRIS.

Radiometric Errors Due to Atmosphere

The presence of atmosphere can significantly modify the received signal from the Earth's surface as a result of scattering and absorption by the particles in the atmosphere.

Absorption:

O_2, CO_2, O_3, H_2O, etc.

These molecules are wavelength selective.

Scattering:

By air molecules ($\propto \lambda^{-4}$): Rayleigh scattering.

By large particles (0.1 λ to 10 λ), such as those associated with smoke, haze, and fumes: Mie scattering.

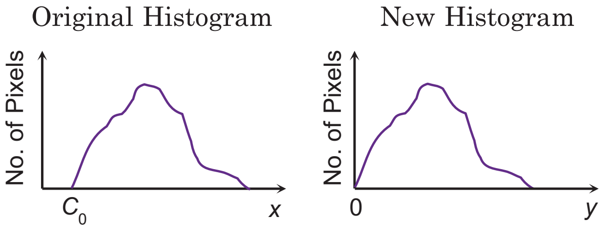
Fine detail in image data will be obscured due to the atmospheric effect.

Bulk Correction of Atmosphere Effects:

Assumption: There are some pixels in the image with brightness of about zero if there is no atmospheric effect.

Find the (average) lowest brightness value in each band (C_0).

$$\boxed{\text{New value } y = \text{Original value } x - C_0}$$



Detailed corrections can be achieved by modeling the absorption and scattering with defined parameters.

Mapping Functions and Ground Control Points

Mapping functions describe the location relationship between the image with error and the reference image (or map). Their coordinates can be denoted as (u, v) and (x, y).

For example,

$$u = 0.5x$$

$$v = y + 5$$

gives the relationship of a compression by half of the horizontal direction, and an offset by 5 in the vertical direction.

Often, the relationship is unknown due to several error sources. A polynomial form is often assumed. The second degree of polynomial is given by

$$u = a_0 + a_1 x + a_2 y + a_3 xy + a_4 x^2 + a_5 y^2$$

$$v = b_0 + b_1 x + b_2 y + b_3 xy + b_4 x^2 + b_5 y^2$$

where the coefficients are estimated from some known samples, which are called **ground control points (GCPs)**.

GCPs are those small, distinguished spatial features, such as a road intersection, corner of a building, or a tower, that can be identified manually or automatically in both images. Their locations (u_i, v_i) and (x_i, y_i) are used as training data to find the model parameters of a and b in the mapping functions.

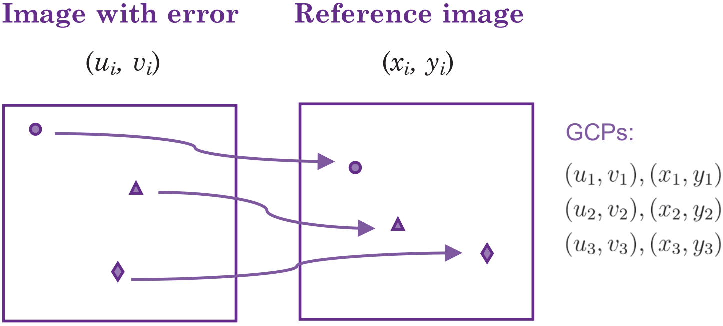
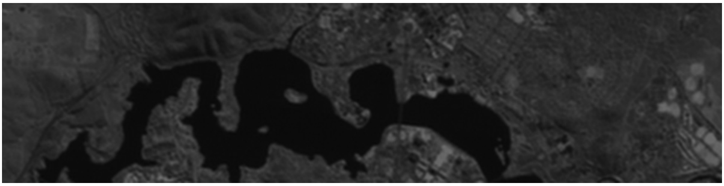


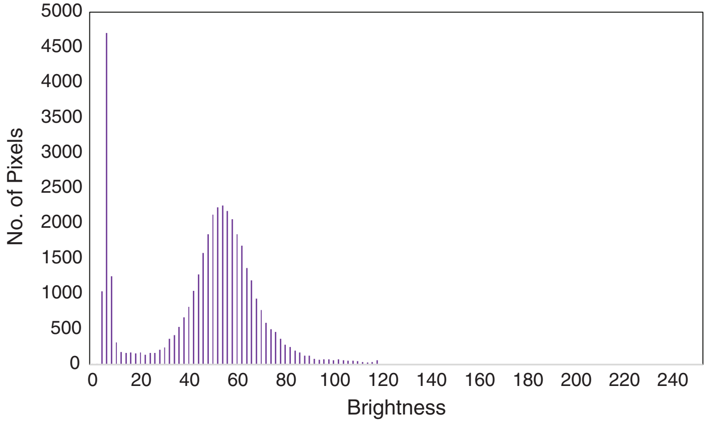
Image Histogram

Let X denote an image data set, $X \in \mathbb{R}^{M \times N}$, where M and N are the number of rows and columns of the image. The intensity of each pixel is denoted as $x(m, n)$, which is in the range of 0 to $2^B - 1$, and B is the number of bits used. For an 8-bit image, $0 \leq x(m, n) \leq 255$.

An **image histogram** is a graph showing the number of pixels in an image whose intensity value is x, $x = 0, 1, 2, \ldots, 2^B - 1$.

An example of the image and its histogram is given below.





It is important to note that an image has a corresponding histogram, but a given histogram is not associated with a unique image.

Spatial Filtering

An image's geometric content can be enhanced via **spatial filtering**. For example, noise and details may be ignored to generate a smoothed image, or edges may be detected to highlight object boundaries. These goals can be achieved by designing a **kernel** for the purpose and performing a **running window operation**, which is also called a **convolution**.

An M by N kernel is $M = 2a + 1$ and $N = 2b + 1$. For example, if $a = 1$ and $b = 1$, the kernel has 9 entries.

This kernel is applied to each pixel $f(x, y)$ and its neighbors on an image to generate a new value of $g(x, y)$:

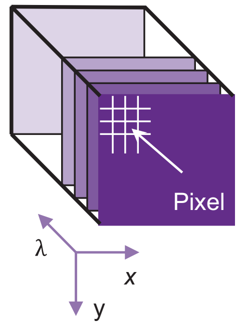
$$g(x, y) = \sum_{s=-a}^{a} \sum_{t=-b}^{b} w(s, t) f(x + s, y + t)$$

$w(-1, -1)$	$w(-1, 0)$	$w(-1, +1)$
$w(0, -1)$	$w(0, 0)$	$w(0, +1)$
$w(+1, -1)$	$w(+1, 0)$	$w(+1, +1)$

$f(x-1, y-1)$	$f(x-1, y)$	$f(x-1, y+1)$
$f(x, y-1)$	$f(x, y)$	$f(x, y+1)$
$f(x+1, y-1)$	$f(x+1, y)$	$f(x+1, y+1)$

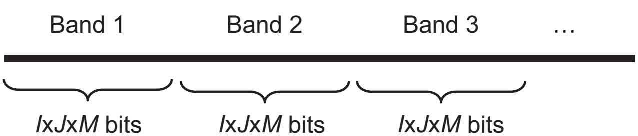
Image Data Cube Files

Number of rows: I
Number of columns: J
Number of bands: N
Number of bits of each data: M

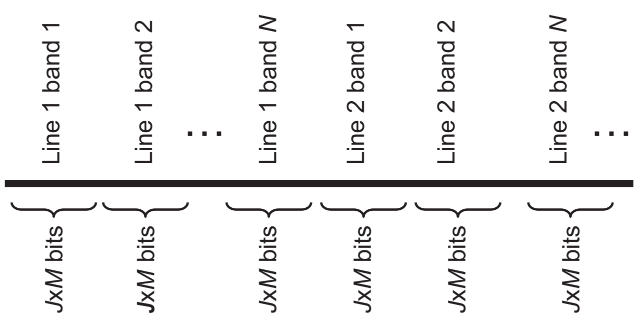


Three data-file formats:

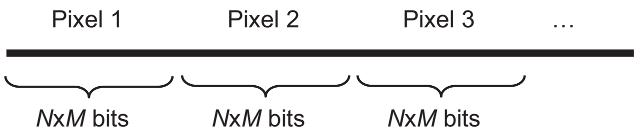
Band sequential (BSQ)



Band-interleaved-by-line (BIL)



Band-interleaved-by-pixel (BIP)



Pixel Vector, Image Matrix, and Data Set Tensor

Feature-space plots are often used to display two bands of data as an illustration. For high-dimensional cases, the full measurements of N bands cannot be plotted and visualized. They are denoted as a **vector** $\mathbf{x}$ for each pixel:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_N \end{bmatrix}$$

A **matrix** can be used to denote an image of a given band n:

$$\mathbf{X}_n = \begin{bmatrix} x_{1,1,n} & x_{1,2,n} & \cdots & x_{1,J,n} \\ x_{2,1,n} & x_{2,2,n} & \cdots & x_{2,J,n} \\ \cdots & \cdots & \cdots & \cdots \\ x_{I,1,n} & x_{I,2,n} & \cdots & x_{I,J,n} \end{bmatrix}$$

where I and J are the number of rows and columns of the image, respectively.

A **tensor** is used to denote the data set:

$$\tau = \{x_{i,j,n}\} \in \mathbb{R}^{I \times J \times N}$$

The operations of vector, matrix, and tensor are used to implement pattern recognition and deep-learning algorithms.

Otsu's Method

Otsu's method is a simple and automatic process for two-class gray-image clustering. The method assumes that the image contains two classes. It has good performance if the histogram is close to a bimodal distribution.

Otsu's method gives an intensity threshold that separates the pixels of a grayscale image into two classes. The threshold is optimal in the sense that it maximizes the between-class variance, a well-known measure used in statistical discriminant analysis.

Steps of Otsu's method

Step 1: Give an initial threshold value of $k = 0$, label the data into Class 1 ($\leq k$) or Class 2.

Step 2: Find the numbers of pixels in Class 1 and Class 2, n_1 and n_2.

Step 3: Calculate the mean values of Class 1 and Class 2, m_1 and m_2.

Step 4: Calculate the between-class variance:

$$\sigma_b(k) = \frac{n_1 n_2}{(n_1 + n_2)/2}(m_1 - m_2)^2$$

Step 5: $k = k + 1$, repeat Steps 1 to 4, until $k = L - 1$ (the maximum intensity value).

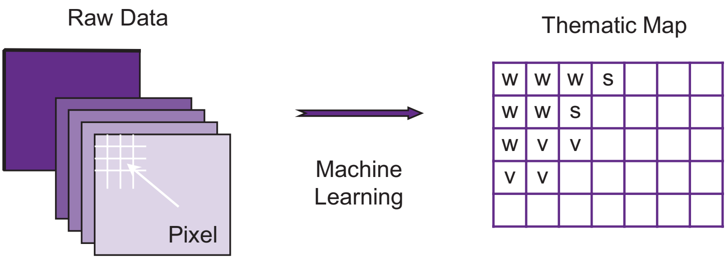
Step 6: Obtain the Otsu threshold k^*, if

$$\sigma_b(k^*) > \sigma_b(k) \forall k \neq k^*.$$

Otsu's method can be extended to more than two classes. This method is too time consuming for an image with multiple bands.

Supervised Classification Procedure

The advantage of a digital remote sensing image is it can help us understand the landcover and perform quantitative Earth observation at the pixel level. To do so, **pattern-recognition** algorithms via **machine learning** need to be applied to label each pixel with an information class, for example, 'w' as water and 's' as soil. The classes are often color coded. This process is referred to as **image classification** or **pixel classification**. A **thematic map** is then generated from the raw data to provide the user-required information.



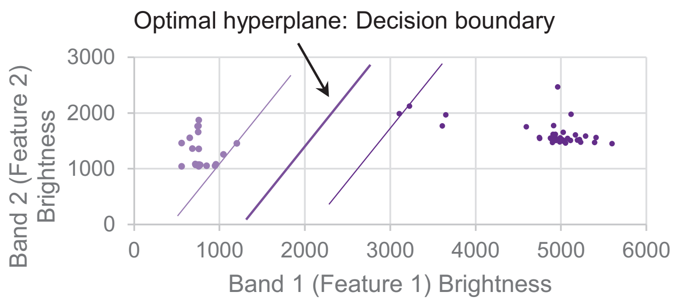
Supervised classification relies on three inputs from an analyzer:

1. Identified and defined **information classes** in the image scene to classify, such as, grassland, bare soil, and water. An exhaustive list of classes is expected, or some pixels in the image may be left undetermined.
2. Prototype pixels in the image based on partial ground truth learned from site visits, photointerpretation, or auxiliary information sources to represent each class. These pixels are used as **training samples** to train a classifier.
3. An appropriate model to describe the difference between the classes that is associated with which **classifier** to use, for example, **support vector machines**. Then, corresponding **class signatures**, often in the form of statistics, or **discriminant functions**, are estimated using training data.

The computer then labels every pixel in the image based on the classification rule of the classifier selected.

Support Vector Machines

The **support vector machine** (**SVM**) is another supervised classification method, which focuses on the training samples around the class boundaries. Those samples are called **support vectors** and are used to define class boundaries (the two thin lines in the image). The **optimal separating hyperplane** (a line in the 2D case) is in the middle of the two class boundaries (the thicker line) and used as a decision boundary for classification.



This decision hyperplane is defined as

$$\mathbf{w}^t\mathbf{x} + b = 0$$

where $\mathbf{w}$ is a weight vector, and b is a bias; these parameters are obtained by solving a quadratic programming problem using the training samples as input. $\mathbf{x}$ is the input vector:

$$\mathbf{x} \in \omega_r, \text{if } \mathbf{w}^t\mathbf{x} + b < 0$$

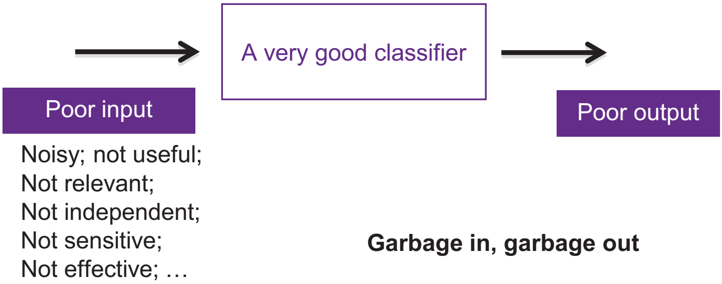
$$\mathbf{x} \in \omega_l, \text{if } \mathbf{w}^t\mathbf{x} + b > 0$$

The SVM is a binary classifier for the class on the left ω_l and the class on the right ω_r in the feature space. For three or more classes, a SVM is performed on each class pair (**one-vs-one**) or on each class against the rest of the classes (**one-vs-rest**). A consensus is used to label each pixel vector.

The Need for Feature Reduction

Features that describe an object or class, for example, reflectance values at RGB bands (color), near- or mid-infrared bands, or textures, are used as input to a classifier for pattern recognition and classification. When the dimensionality of the input is high, class modeling can be biased with limited training samples. This situation leads to the **Hughes phenomenon**, i.e., **overtraining**. The trained model has low generalization capacity; it works for the training data but not for the new data. This over-training is also called the **curse of dimensionality**.

It is important to provide good input features.



Feature mining is an important task to find useful and effective features for class discrimination. It includes

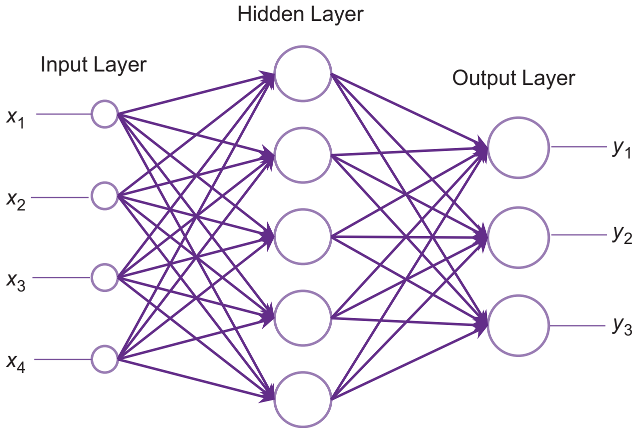
- **Feature selection**
- **Feature generation**
- **Feature extraction**
- **Feature learning**

The approaches can be **knowledge based** or **data driven**. They can be **supervised** to address class separability or **unsupervised** in terms of statistics on data quality.

Feature learning by a deep neural network is an alternative approach to find discriminant features for a given classification or target detection task.

Artificial Neural Networks: Structure

Artificial neural networks (**ANNs**) mimic the structure of the human brain. An example ANN is shown below.



Input layer: The number of neurons equals the dimensions of the input vector, which is 4 in the example given above.

Hidden layers: One or more hidden layers is often used, and each has several neurons that achieve nonlinear decision regions. More than two hidden layers are later called a **deep network**. If there are no hidden layers, a linear classifier is implemented, and a hyperplane is established as the decision surface.

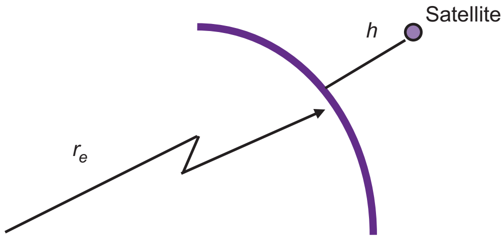
Output layer: This layer has the same number of neurons as the number of classes to classify. The input vector $\mathbf{x}$ is recognized as belonging to the i^{th} class if the neuron y_i has the highest output.

Arrows: Each arrow indicates a data path with a given weight obtained after training.

Neurons: Each neuron (except those at the input layer) is a classifier as explained on the next page. This neuron is also called a **node**.

Satellite Orbit Period

The period of a satellite orbiting the Earth T is determined by the altitude of the satellite above the Earth surface h and is independent of the size and the weight of the satellite.



$$T = 2\pi\sqrt{(r_e + h)^3/\mu}\ (\mathrm{s})$$

where
$\mu = 3.986 \times 10^{14}\ \mathrm{m^3 s^{-2}}$ (Earth gravitational constant)
$r_e \approx 6.378 \times 10^6$ m (Earth radius)

When $h = 35.8$ Mm, $T = 24$ hr, which is the same period as the Earth's rotation. The **geosynchronous equatorial orbit** (**GEO**) is located at this height, where satellites rotate with the Earth; these are called **geostationary satellites**.

For **low Earth orbit** (**LEO**) observation, the altitude is normally below 1000 km so that spatial resolution is reasonable and revisit time is good. However, it is impractical to set the altitude below 160 km due to strong **atmospheric drag** (atmospheric force against the motion of an object). At this altitude, T is much less than 24 hr.

The NASA Aqua satellite is located 705 km above the Earth surface, where $T = 99$ min. The orbit altitude of ESA QuickBird-2 is 450 km, where $T = 93$ min. The period is weakly affected by the orbit altitude.

Error Matrix for One-Class Mapping

An **error matrix** lists how reference data are recognized as defined classes. For the one-class case, this matrix shows how accurately the target is detected against the background.

Error Matrix		Reference Data	
		Target	Background
Results	Target	TP	FP
	Background	FN	TN

TP: True Positive TN: True Negative
FN: False Negative FP: False Positive

The accuracy of the target detection can be assessed by

Recall = TP/(TP + FN)
Precision = TP/(TP + FP)
Overall accuracy = (TP + TN)/(TP + TN + FN + FP)

Recall gives the percentage of the reference data that are correctly recognized. Precision gives the percentage of the labeled data that are correct. High recall implies a low miss rate. High precision implies a low false alarm rate.

Example:

Error Matrix		Reference Data	
		Target 100	Background 200
Results	Target 188	98	90
	Background 112	2	110

Recall: 98/100 = 98% for target
110/200 = 55% for background

Precision: 98/188 = 52% for target
110/112 = 98% for background

Overall accuracy: (98 + 110)/(100 + 200) = 70%

Xiuping Jia received a B.Eng. degree from the Beijing University of Posts and Telecommunications, Beijing, China, in January 1982. Later she received a Ph.D. degree in Electrical Engineering and a Graduate Certificate in Higher Education from The University of New South Wales, Australia, in 1996 and 2005, respectively, via part-time study. She has had a lifelong academic career in higher education, for which she has continued passion. She is currently an Associate Professor with the School of Engineering and Information Technology, The University of New South Wales at Canberra, Australia. Her research interests include remote sensing, hyperspectral image processing, and spatial data analysis. Dr. Jia has authored or co-authored more than 280 referred papers, including over 170 journal papers addressing various topics, including data correction, feature reduction, and image classification using machine learning techniques. She co-authored the remote sensing textbook *Remote Sensing Digital Image Analysis* [Springer-Verlag, 3rd (1999) and 4th eds. (2006)]. These publications are highly cited in the remote sensing and image processing communities. She is an IEEE Fellow.