INTERNATIONAL ATOMIC ENERGY AGENCY
UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION
INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
I.C.T.P., P.O. BOX 586, 34100 TRIESTE, ITALY, CABLE: CENTRATOM TRIESTE

H4.SMR/383 - 03

WORKSHOP ON REMOTE SENSING TECHNIQUES WITH APPLICATIONS TO AGRICULTURE, WATER AND WEATHER RESOURCES

(27 February - 21 March 1989)

DIGITAL IMAGE PROCESSING FUNDAMENTALS

STEFANO BARONTI
IROE - CNR
Via Panciatichi, 64
50127 Firenze
Italy

Digital Image Processing Fundamentals

Stefano Baronti
IROE - CNR
Via Panciatichi 64
50127 Florence, Italy

Introduction

An image can be considered a function f of two spatial independent variable x, y. The value of the function is perceived by our eyes as a grey level which is in general the representation of any physical quantity. In remote sensing applications it is usually a measure of the radiation in a given observation band , depending on the characteristics of the sensor.

From a mathematical point of view it is appropriate to assume that the function f(x,y), representing the image, has a regular behaviour. This is necessary when transformations on f(x,y) have to be applied. We shall also assume that:

$$0 \leq f(x,y) \leq M \qquad (1)$$

where M is a positive constant. The assumption is appropriate since f(x,y) is the representation of a quantity that is related to the energy coming at the sensor and so positive. It is also limited by the dynamic range of the sensor itself which introduces saturation effects when the signal becomes too high.

In the case of natural images, f(x,y) is a continuous function, defined on a continuous domain. Digital image processing deals, instead, with quantity which have to be represented by a computer; these quantities have to be defined on a discrete domain and assume only a limited set of values. This is obtained by sampling and quantizing f(x,y). These operations are described in the following.

We also note that, in remote sensing applications, we often deal not only with one image of the observed area but also with a set of images which refer to several observation bands. In this case, we can think f(x,y) as a vector function. The dimension of the vector (the number of the components) is given by the number of bands which are available.

Sampling

Sampling is the operation by which the domain of f(x,y) is made discrete. The question to be considered is if during this operation some information present in f(x,y) gets lost. This is a

good question since we are passing from an infinite set of values to a finite one: can a finite set of samples maintain the original information present in f(x,y)? The response is given by the sampling theorem.

The sampling theorem states that an analog signal f(x,y) can be recovered by its sampled version f(m,n) if the sampling rate is "sufficiently" high.

Let F_{sx} and F_{sy} the sampling frequencies along the x and y directions, respectively. Sufficiently high means that:

$$F_{xs} \geq 2F_{xmax} \qquad (2)$$

$$F_{ys} \geq 2F_{ymax}$$

where F_{xmax} and F_{ymax}, are the maximum grey level frequency variations along the x, y axes respectively.

In an equivalent way the sampling periods $T_{sx} = 1/F_{xs}$ and $T_{sy} = 1/F_{ys}$, have to be sufficiently small:

$$T_{sx} \leq T_{xmax}/2 \qquad (3)$$

$$T_{sy} \leq T_{ymax}/2$$

where $T_{xmax} = 1/F_{xmax}$ and $T_{ymax} = 1/F_{ymax}$.

The theorem is due to Shannon and can be found in any text as for example[1],[2].

The demonstration also gives a method to reconstruct the original signal: it is sufficient to process the digital signal with an analog low-pass filter tuned on the original signal.

Sampling theorem let us work with digital signals without any trouble; we can act any digital transformation by a computer and then return to the analog representation when we need.

This is what happens, for example, when we look at a color display interfaced with a computer; the signal is read from the memory (or from a graphic board) where it is digitally stored; it is converted to an analog form by the display controller which performs the interpolating low-pass filtering.

We also note at this point that operations performed on an analog signal are very fast, but operations realized on the digital one are more flexible. We cite only two examples. The former refers to linear operations which are defined in the following sections: it is possible to perform perfectly linear operations with a digital system. This is not true with an analog system, particularly for images, since there is always some kind of distortion introduced by the optics (aberrations and non-linearities of the lenses).

The latter example refers to the opportunity of storing the signal after it has been digitized. It is possible to consider operations in which various samples are involved in various positions of the stored frame; this is very immediate with the digital signal and nearly impossible with the analog one.

From a practical point of view, after sampling, the function f(m,n) can be thought and represented as a matrix of convenient dimensions.

Just as an example, fig. 1 shows a Landsat thematic Mapper image whose dimensions are 512x512 pixels (picture elements). Varying sampling rate have been considered and the effect is evident. The pixel resolution is 30 m.

Quantization

Quantization is the process by which an image, that originally assumes a continuous range of values, is reduced to assume a discrete set of values. The resulting set depends on various factors, the main being the signal to noise ratio of the sensor.

Practical considerations also limit the number of representation levels. For example, it is particularly convenient to represent data in a byte-aligned form. In this case, each grey level is represented with 8 bits and the range of possible values is [0 255]. In general with N bits the resulting range of values is $[0 \; 2^N-1]$.

Quantization is usually performed in a linear way: the grey-level interval is divided into 2^N steps of fixed amplitude. The signal is given level 0 when it falls in the first interval and in general is given level l-1 when it falls in l^{th} interval.

This is not the best way to quantize a signal. In fact steps should be finer where the signal values are more likely. However, this is not taken into account in most acquisition systems and in particular on satellites. For this reason a-posteriori corrections are often considered as for example histogram equalization, whose purpose is to redistribute grey level in a better way.

In fig. 2, as an example, different quantizations have been considered to the image already presented in fig. 1. The quantization effect appears and it is perceived as a contrast variation.

Digital system properties

A digital system can be defined as an operator (or a filter) which transforms an input sequence f(m,n) into an output sequence g(m,n).

$$g(m,n) = T[f(m,n)] \qquad (4)$$

A digital system is said to be linear if:

$$T[Af'(m,n)+Bf''(m,n)] = AT[f'(m,n)]+BT[f''(m,n)] = \qquad (5)$$
$$= Ag'(m,n)+Bg''(m,n)$$

where A,B are real constants, f'(m,n) and f"(m,n) are input sequences and g'(m,n) and g"(m,n) are output sequences.

Linear systems are very important because of their properties, as it will be evidentiate in the following. Eq. 5 states that for linear systems the superposition principle holds.

Impulse response

A linear system is completely characterized by its impulse response; this is the output of the system when an unitary impulse d(n,m) is applied in input.

$$d(n,m) = \begin{cases} 1 \text{ if } m,n = 0 \\ 0 \text{ otherwise} \end{cases} \qquad (6)$$

For images, the inpulse response is often indicated as Point Spread Function (PSF) due to the nature of the input source used for the definition: just a spot of light in the input plane.

Let us consider a linear system whose PSF is h(m,n). Let f(m,n) the input image. The output image g(m,n) can be expressed in the following way:

$$g(m,n) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} h(i,j)f(m,i;n,j) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} f(i,j)h(m,i;n,j) \qquad (7)$$

This expression can be simplified if the shift-invariance property is assumed. Shift invariance is defined by requesting that eq. 7 becomes:

$$g(m,n) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} h(i,j)f(m-i,n-j) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} f(i,j)h(m-i,n-j) \qquad (8)$$

From a physical point of view, shift-invariance means that the response $g(m,n)$ of the system to $f(m,n)$ has to be the same, a part a translation, when $f(m,n)$ is translated in the input plane.

Convolution is often indicated in a concise form:

$$g(m,n) = f(m,n)*h(m,n) \qquad (9)$$

Correlation

Correlation is formally similar to convolution as it regards the definition. We usually refer to autocorrelation when considering a function correlated with itself and cross-correlation when considering two different functions. Correlation is defined as:

$$C_{ff}(m,n) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} f(i,j)f(m+i,n+j) \qquad (10)$$

In the particular case that $m,n = 0$ eq.(10) becomes:

$$C_{ff}(0,0) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} f(i,j)f(i,j) \qquad (11)$$

Eq. (11) expresses the energy of $f(m,n)$.

Cross correlation is defined as:

$$C_{fg}(m,n) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} f(i,j)g(m+i,n+j) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} g(i,j)f(m+i,n+j) \qquad (12)$$

As it can be easily shown $C_{fg}(m,n)$ expresses a similarity measure between $f(i,j)$ and $g(i,j)$. In fact, if our aim is to measure the difference between $f(i,j)$ and $g(i,j)$, when the two images are considered with some shifting, we can evaluate the square error as a function of the shifting. In the following formula, m, n are the two shifts.

$$E_{fg}(m,n) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} (f(i,j) - g(m+i,n+j))^2 \qquad (13)$$

And squaring the terms:

$$E_{fg}(m,n) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} (f(i,j))^2 + \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} (g(m+i,n+j))^2 + \qquad (14)$$

$$- \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} f(i,j)g(m+i,n+j)$$

The error is minimum when eq. (13) assumes the minimum or in an equivalent way the third term of the second member of eq. (14) assume its maximum. But this last term is just $C_{fg}(m,n)$.

The problem of finding corresponding points between images of the same scene is very common in remote sensing. One way, to solve it, is to consider $C_{fg}(m,n)$ and to find its maximum. The position m', n' of the maximum represents the degree of shifting between the two images.

Since $C_{fg}(m,n)$ depends on the energy of $f(i,j)$ and $g(i,j)$ it is necessary to consider a normalization. The final expression which can be implemented on a computer is so the following:

$$C_{fgn}(m,n) = C_{fg}(m,n) \Big/ \Big(\sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} (f(i,j))^2 \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} (g(i,j))^2 \Big)^{1/2} \qquad (15)$$

In this way $0 \le C_{fgn}(m,n) \le 1$ and $C_{fgn}(m,n) = 1$ only when $f(i,j)$ is equal to $pg(i,j)$, where p is a real constant.

Convolution theorem

As in the case of continuous signals, Fourier transform can be defined also for digital signals. The result is a Discrete Fourier Transform (DFT) whose properties are similar to those which are defined for analog Fourier transform.

Let $f(m,n)$, $g(m,n)$, $h(m,n)$ the input image, the output image, and the PSF of a linear system respectively. Let $F(u,v)$, $G(u,v)$, $H(u,v)$ the correspondent Fourier transforms.

The convolution theorems states that the DFT of

$$g(m,n) = f(n,m)*h(m,n)$$

is given by:

$$G(u,v)=F(u,v)H(u,v) \qquad (16)$$

This theorem is very useful since some operations can be conveniently realized by considering the DFT of an image by passing in the transformed domain, then multiplying it by a suitable transfer function H(u,v) and finally returning to the original spatial domain by an inverse transformation. Besides, an operation performed in the spatial domain can be analyzed by examining the Fourier transform of the PSF and of the resulting image.

Separability

A system is said to be separable if:

$$H(m,n) = H_1(m)H_2(n) \tag{17}$$

This property is very important from an implementation point of view. In fact, for such systems, two-dimensional convolution is evaluated by applying two one-dimensional convolutions. The number of sums and products necessary for the operation are in this way proportional to the dimension of the PSF and not to its square.

Statistical properties of images

Some basic concepts of statistics are necessary to analyze remote sensing data.

Let p(x) the probability that an event x occurs. The event in which we are interested is that a point in the image assumes a given grey level, or, if we are dealing with multispectral data, that in $band_1,\ldots,band_N$, the grey levels are $i_1,\ldots,i_N$.

We will associate a probability to this event; p(i) in the former case and $p(i_1,\ldots,i_N)$ in the latter.

An estimation of p(i) is obtained by counting the number of times, n(i), that a grey level appears in an image:

$$p(i) = n(i)/N_p \tag{18}$$

where N_p indicates the total number of pixels of the image and:

$$\sum_{i=0}^{L} p(i) = 1 \tag{19}$$

where $L=2^N-1$ represents the number of grey level in the range [0 L] and N indicates the number of bits adopted for the representation. In the same way:

$$p(i_1,\ldots,i_N) = n(i_1,\ldots,i_N)/N_p \tag{20}$$

$$\sum_{i1=0}^{L} \ldots \sum_{iN=0}^{L} p(i_1,\ldots,i_N) = 1 \tag{21}$$

$p(i_1,\ldots,i_N)$ is called joined probability of the event $i_1,\ldots,i_N$.

p(i) is referred as one-dimensional histogram of grey levels and is particularly used in the characterization of sensors, and in image analysis.

$p(i_1,\ldots,i_N)$ is referred as n-dimensional histogram of grey levels and is particularly used in segmentation and classification algorithms.

p(i) is formally a probability density function; a Cumulative Density Function (CDF) can be associated to it:

$$c(i) = \sum_{j=0}^{i} p(j) \tag{22}$$

c(i) estimates the probability that a point in the image assumes a level which is less or equal than i. In a similar way the CDF can be defined for a multispectral image.

Often in remote sensing, events are considered whose probability density function is assumed to be normal; in this case

$$p(x) = (2\pi)^{-1/2}\sigma^{-1}\exp(-1/2(x-m)^2/\sigma^2) \tag{23}$$

where m indicates the mean value of x and σ is its standard deviation. An estimation of m and σ is given by:

$$m = 1/q \sum_{i=1}^{q} x_i \tag{24}$$

$$\sigma^2 = 1/q \sum_{i=1}^{q} (x_i-m)^2 \tag{25}$$

where q is the number of experiments considered.

In the multivariate case eq. (23) can be extended by considering that x is now a vector of N rows and 1 column. Symbol $()^T$ indicates the transposition operation. Eq. (23) becomes the following:

$$p(x) = (2\pi)^{-N/2}|\Gamma|^{-1/2}\exp(-1/2(x-m)^T\Gamma^{-1}(x-m)) \tag{26}$$

where

$$\Gamma = E[(x-m)(x-m)^T] = 1/q \sum_{i=1}^{q} (x_i-m)(x_i-m)^T \tag{27}$$

$$m = 1/q \sum_{i=1}^{q} x_i \tag{28}$$

and $|\Gamma|$ indicates the determinant of Γ; Γ is indicated as covariance matrix.

Conditional probability

Conditional probability indicates the probability that an event x occurs conditional upon an other event y. The notation is $p(x|y)$. If we consider the join probability $p(x,y)$, we have that:

$$p(x,y) = p(x|y)p(y) = p(y|x)p(x) \qquad (29)$$

from which

$$p(x|y) = p(y|x)p(x)/p(y) \qquad (30)$$

Eq. (30) expresses the Bayes theorem.

These definitions are extensible also to the multispectral case in which x and y are vectors.

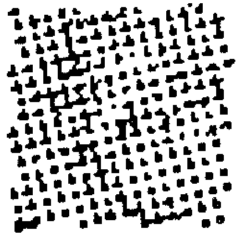
Bibliography

A. V. Oppenheim, R. W. Schafer, Digital Signal Processing (Prentice Hall, Englewood Cliffs, 1975)

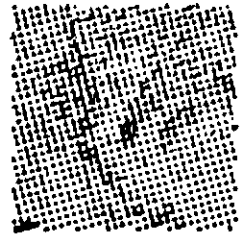
W. K. Pratt, Digital Image Processing (Wiley, New York, 1978)

J. A. Richards, Remote Sensing Digital Image Analysis (Springer-Verlag, 1986)

FIG. 1



Subsampling 8



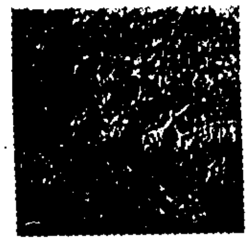
Subsampling 4

Subsampling 2

Original

FIG. 2

16 grey levels



8 grey levels