Fundamentals of Image Processing

What is image processing?

Image processing may be regarded as an application of a particular porcess to an image in order to obtain another image.

What is an image and why do we need to process an image in order to obtain another?

Fundamentally, an image is a pictorial representation of a particular domain.

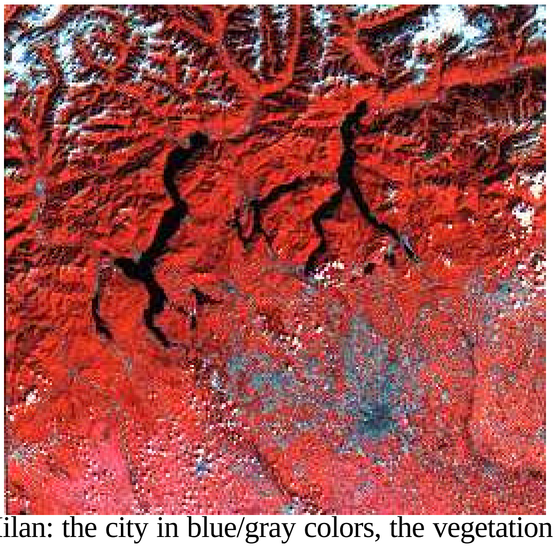
Examples are:

- Medicine (x-rays);
- Biology (electron micrographs of biological material);
- Satellite date (pictures of the Earth and its environment);
- Commercial documents (legal and insurance records);
- Archaeological data (Pictures of Relics);
- Forensic and video data (fingerprints, identification shots);
- Industrial processing (quality and assembly control data);
- Any camera negative.

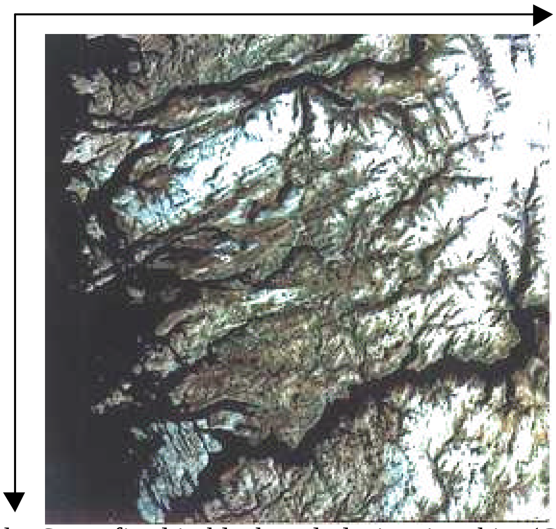
The purpose in processing images is to enhance the pictorial information for subsequent human or computer interpretation.

IMAGE REPRESENTATION

As an example of image representation, consider a simple satellite scanning photographic image in one of the ranges:



The lakes north to Milan: the city in blue/gray colors, the vegetation and the Alps in red.
(Cordis Focus, EU, No. 20, 1999)



Norwegian landscape with the Sognefjord in black and glaciers in white (Cordis Focus, EU, No.20, 1999)

On closer inspection, the image shows that it is made up of ***variations in light and dark*** at different parts of the negative according to the density of the photographic emulsion at any chosen point.

In fact it's possible to quantify the variation in intensity across the image to give us what is commonly known as the ***gray level distribution*** of the image.

Since the image is planar, we represent this distribution by a function ***f(x,y)*** where ***f(x,y)*** is simply the intensity at a selected point having coordinates ***x,y***.

To establish the ***co-ordinate structure*** we adopt the convention that, the ***top left hand corner*** of the image corresponds to ***the origin*** and the x and y axes are the corresponding vertical and horizontal directions drawn from the origin (as in picture above). Each point ***x,y*** is called a '***pixel***' which stands for 'picture element'.

IMAGE ACQUISITION

Devices to capture images:

the microdensitometer, which is frequently used in the electron microscopy field (an electron micrograph consists of a film transparency obtained by recording the distribution of electrons scattered by a specimen in the microscope; the transparency is subsequently placed in the microdensitometer and the intensity distribution is measured by focusing a beam of light on the transparency at different points and recording, using a photodetector, the amount of light either transmitted or reflected at each point).

CCD (Charge Coupled Device) camera or sensor (this type of sensor can be configured either as a line scan sensor or as an area sensor. ***A line scan sensor*** consists of a row of silicon imaging elements that have a voltage output proportional to the intensity of the incidental light. This output can then be converted into digital form for subsequent input into the computer. The scanner operates by scanning the two-dimensional image on a successive line basis starting at the top of the image and working down. ***An area sensor*** consists of a matrix of imaging elements and is capable of acquiring an image in a way similar to a video device).

Radiometers (on satellites: visible and infrared spin scan radiometers, and various modifications of them, like Return Beam Vidicon, Multispectral Scanner, Thematic Mapper, Visible Imaging Instrument, Heat Capacity Mapping Radiometer, etc.)

The ***resolution or extent*** to which details of the image can be recorded vary with the choice of scanner and can typically extend from 100 to 1000 dots per inch (dpi).

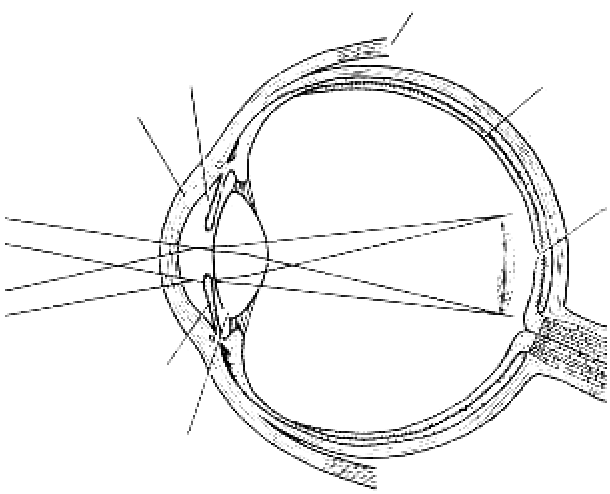
The digital image (8 by 8 pixels, in range between 0 and 7) of let's say letter " ***T*** " is like this:

0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	7	7	7	7	7	7	0
0	7	7	7	7	7	7	0
0	0	0	7	7	0	0	0
0	0	0	7	7	0	0	0
0	0	0	7	7	0	0	0
0	0	0	0	0	0	0	0

THE HUMAN EYE AND RECOGNITION

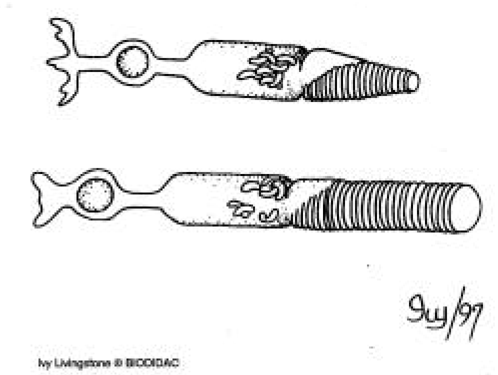
The eye, due to its flexibility, is generally regarded as significantly ***superior*** to any camera system developed to date.

The range of intensities to which the eye can adapt is of the order ***10^{10}*** i.e. from the lowest visible light to the highest bearable glare. The basic components of the eye consist of a lens, a retina and an iris:



The central part of the retina called the fovea, contains between six and seven million ***cones*** which are sensitive to colour and are connected directly with the brain via ***individual nerves***. An image that is projected on to the retina is converted into an electrical impulse by the cones and then transmitted by the nerves into the brain.

Another crucial part in the structure of the eye are the ***rods***. These are distributed across the surface of the retina and number between 75 and 150 million. The rods share nerve endings, and are not involved in colour vision, but are sensitive to light and dark. In fact the ***cones do not function in dim*** light. Rods & cones:



The eye is flexible in that, unlike a fixed camera, it is able to adapt to different situations of luminosity by changing its structure. For example, the iris may contract or expand to control the amount of light that enters the eye.

Despite the remarkable capabilities of the eye, in form we receive most image material, the eye is often ***unable to distinguish*** the relevant components from the image and that these components are only recognisable after the image has been processed.

USE OF COLOUR

In general, the eye is ***less able to discern*** gradual changes in brightness compared to gradual changes of colour.

It is thought that the eye is only capable of differentiating between approximately sixteen distinct gray levels although some research suggests that up to fifty gray levels can be identified on most *monochrome* displays.

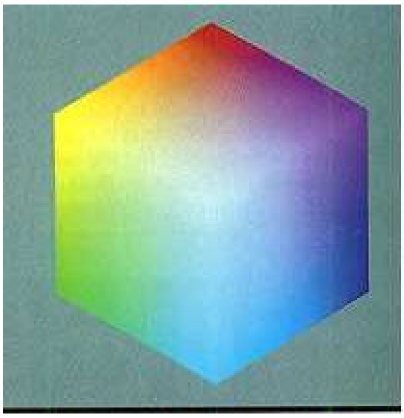
With colour, however, **the** eye is capable of distinguishing around two hundred colour shades.

The Milan picture above illustrates the effectiveness of using colour in distinguishing the different parts of the environs of the city. Similarly, Norway picture shows the effectiveness of assigning different colours to differentiate the structural components.

The ***assignment of colour*** can be done in two ways:

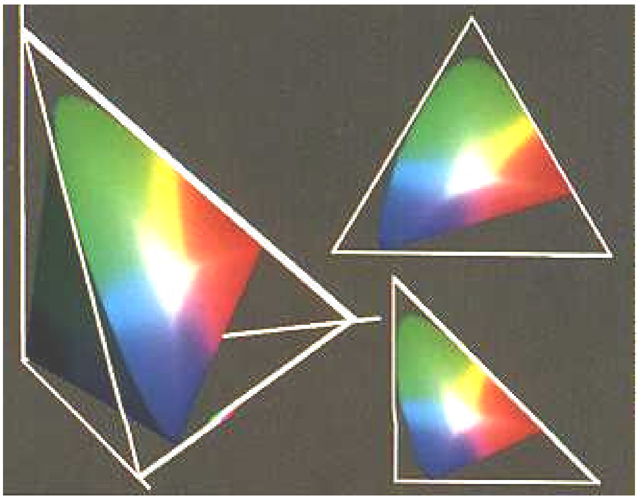
- we can use a colour scanner such as a colour TV camera,
- we can assign a particular colour to every gray level intensity value on a graphics display of the image.

The latter approach approach assumes that we have to use some modelling procedure of colours. The popular ***model is RGB***, assuming primary colours, red, green and blue:



Other models are:

- CIE:



- and also CMY, CMYK, YIQ, HSV, HLS, etc.

The RGB model assumes that the three RGB components comprise the axes of a unit cube. The diagonal of the cube represents the range of gray level intensities. The origin of the cube corresponds to the white end of the gray scale and has RGB component values (0, 0, 0). The opposite end of the diagonal represents black and has RGB value (1, 1, 1). Intermediate values on the gray level intensity diagonal can be assigned proportionate values of the RGB components.

IMAGE SAMPLING

Before digitising an image we need to know how many ***pixel gray*** levels we should sample so that ***a subsequent display*** of the levels shows the image in sufficient detail.

If we ***over-sample*** then we have a situation in which the image contains redundant information implying that we would be taking unnecessary storage.

If we ***under-sample***, a subsequent display of the image will show insufficient detail for us to arrive at a correct interpretation of the image.

original photo

continuous-tone photo with 4 intensity levels

continuous-tone photo with 8 intensity levels

continuous-tone photo with 16 intensity levels

continuous-tone photo with 32 intensity levels

continuous-tone photo with 64 intensity levels

For a given image, we now define some of the basic steps which represent the overall procedures that combine to represent an image processing system:

- acquire the image in a digital form using a suitable digitising device;
- display all or part of the image depending upon the size of the image and the nature of the problem domain (the display may be in either monochrome or colour);
- select an appropriate transformation process in order to enhance the image or to isolate regions specific to subsequent analysis;
- analyse and interpret the image and store the result together with the enhanced image.

THE FOURIER TRANSFORM

Before beginning mathematical section, it is useful first to describe the situation in words.

If a problem is difficult, we try to "look at in another way" or to "turn it round" so that we can see it differently. The techniques of transforming a situation are basic to mathematics and there are very many different kinds of transformation. A picture, ***a distribution of density in a space*** of one, two or three dimensions may be transformed and represented in a new way in ***another space***. One of the most important ways is a Fourier transformation (called after Jean Baptiste Joseph Fourier who first described the method in 1812 when dealing with problems in the flow of heat).

We may set out quantitatively the properties of the Fourier transform as follows:

(1) a density distribution can be built up by the linear superposition of a number of ***sine waves*** of different frequencies. If the density distribution is periodic, of frequency ***f***, then these sine waves have only the frequencies ***f, 2f, 3f ... nf*** ... , but if the density distribution is non-repeating, like a single pulse, then an infinite band of frequencies is required. Resolving a complex signal, into its Fourier components (the sine waves) is known as ***Fourier analysis***. Combining the sine waves together again to give the general signal is correspondingly called ***Fourier synthesis***;

(2) for each constituent sine wave we have to state its ***frequency***, its ***amplitude*** and its ***phase***. If the wave is in more than one dimension then we must also state its ***direction***. Complex numbers provide an appropriate mathematical representation of amplitude and phase;

(3) in many physical systems, such as for radio waves in space or for ***light in a camera***, the principle of ***linear superposition*** applies. This means that wave trains can cross each other without getting mixed up. If linear superposition does not apply then two waves crossmodulate each other and new sum and different frequencies arise. ***Non-linear devices*** have their important place. If we have moving waves then it is vital to know whether waves of all frequencies travel with the same velocity (as is the case for electromagnetic waves in free space, but it is not so for a medium such as glass). We have to be clear on the properties of the particular system with which we are dealing;

(4) the spectrum of sine waves which make up a particular density distribution is know as its ***Fourier transform***. In representing the transform we have to make provision for noting the amplitude, frequency, phase and direction of every constituent wave. We can plot out the transform in ***transform space***. If the scales used in the ***direct space or real space*** in which we plot our original density distribution are, for example centimetres, then the scales used in the transform space will be reciprocal, for example per centimetre, which we may call reciprocal centimetres;

(5) when we buy an audio amplifier we ask about ***its frequency response*** and hope that all frequencies of notes in the complex sound signal going in, will come out "undistorted". The curve which describes the response of the system to different frequencies is called the "***contrast transfer function***". Asking about the contrast transfer function of an optical lens, an electron microscope, or of some other system has revolutionised our understanding of these instruments.

Many problems, then, ***are simplified*** if we first resolve a complex density distribution into sine waves and ask what happens to each sine wave separately as it passes through the system. We then reassemble the somewhat changed sine waves to give a resultant density distribution. Waves may be changed in amplitude, phase or direction. We may be considering the properties of a physical system, such as a microscope, or we may be using a computer as a kind of generalised imaging system, where we can apply any contrast transfer function, physically realisable or not.

(6) we have to build up a kind of ***library or dictionary*** of typical density distributions and their transforms. Such lists, in mathematical form, are to be found in handbooks. Visually, a picture in real space corresponds to another picture in transform space. It is important to acquire a qualitative pictorial appreciation of a range of such relationships to understand what is happening mathematically. The dictionary is reciprocal and works equally in both directions - transforming a transform takes us back to the original distribution;

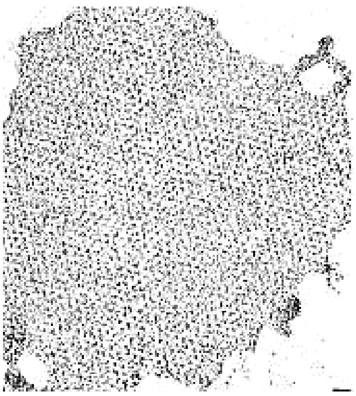
(7) there is an important theorem or principle, the ***convolution theorem***, which helps us to build up this library. It tells us bow to ***think about*** the transforms of complex patterns in terms of the transforms of simpler objects. If we have two density distributions ***A*** and ***B*** (***two pictures*** for example), then the convolution of ***A*** and ***B*** means repeating the whole of ***A*** at every point of ***B***, or vice versa (it does not matter which way we go). For example, if ***A*** is a circle and ***B*** is a lattice of points, then the convolution of ***A*** and ***B*** (which we may write as ***conv (A, B))***, is a lattice of circles;

The convolution theorem then states that if we have a distribution in real space which is the point-by-point product of the distributions ***A*** and ***B***, then the transform of this distribution is the convolution of ***transform (A)*** with ***transform (B)***. The theorem works in either direction. If in real space the distribution ***C*** is the convolution of the distributions ***A*** and ***B***, the ***transform (C)*** is the point-by-point product of the transform of ***A*** and the transform of ***B***;

(8) the convolution theorem enables us to do many remarkable things. During the ***Apollo 13*** space flight the astronauts took a photograph of their damaged spacecraft, but it was out of focus. It is possible by image processing methods to put such an out-of-focus picture back into focus and thus to clarify it.

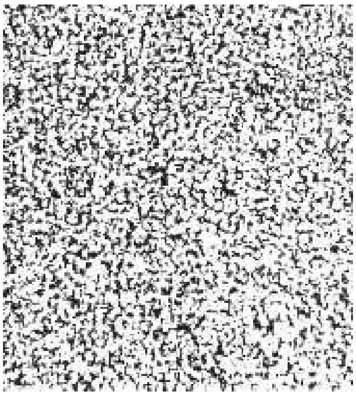
Before formulating the mathematics of the Fourier transform an example is given to illustrate the power of Fourier processing methods applied to images containing a periodic substructure.

One of the most widely used techniques for the study of biological materials, such as viruses or components of isolated cells, is electron microscopy. This process involves exposing the biological material to a beam of electrons which are then scattered by the material. The subsequent pattern of the scattered electrons is recorded photographically and it is this pattern which provides details about the structure of the material. The photograph is referred to as an electron micrograph and can be automatically converted into a set of digital density values by assigning numbers to the different levels of intensity within the micrograph. ***Figure a*** shows an electron micrograph of material comprising one of the surface proteins which forms part of the structure of the influenza virus:



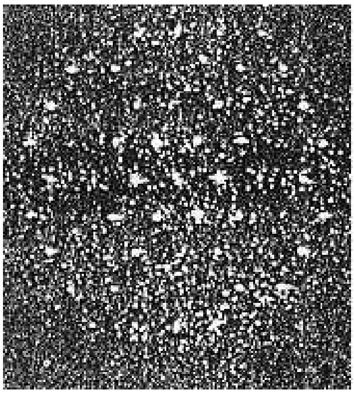
(a)

Figure b shows an enlarged display of the digitised micrograph:



(b)

This digitised set of densities is stored in the computer, the Fourier transform applied and the corresponding frequency components displayed (figure c):

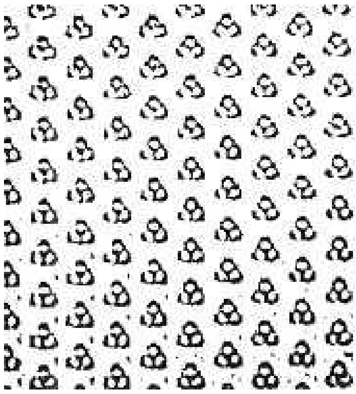


(c)

The figure, however, shows two types of distribution; one which consists of larger spots with a regular spacing between them, and the other which consists of smaller spots with no apparent regular spatial features. The former set of spots correspond to the ***basic frequencies*** of the material and the latter arise from the "noise" components which occur during the process of obtaining the electron micrograph.

The "noise" constitutes an unwanted part of the micrograph, since it obscures much of the structural detail and we must therefore try and remove it. We achieve this by retaining only those spots in the frequency display which correspond to the basic structural frequencies:

and then apply the inverse Fourier transform to arrive at a “noise free” image:



The most striking feature of this process, called ***digital spatial filtering***, can be seen in the comparison between ***figures b*** and ***c***, where the structural details of the image can be clearly seen, whereas they were previously obscured by “noise”.

FORMULATION OF THE FOURIER TRANSFORM

The discrete one-dimensional Fourier transform of the function ***f(x)***, evaluated over N points, may be written as

$$F(u) = \frac{1}{N}\sum_{x=0}^{N-1} f(x)\exp(-2\pi iux/N)$$

$$\text{for } u = 0, 1, 2, \ldots, N-1 \qquad \text{with } i = \sqrt{-1}$$

This transform has an inverse representation which may be correspondingly written as:

$$f(x) = \sum_{u=0}^{N-1} F(u)\exp(+2\pi iux/N)$$

$$\text{for } x = 0, 1, 2, \ldots, N-1$$

In applying the transform in practical situations we frequently need to use both the transform and its inverse form, particularly in situations where modifications are made to the transform values and we require to see how these modifications may affect our original values ***f(x)***.

For the application of Fourier transforms to actual images we require the implementation of the two-dimensional form. This we can do with no loss of generality although we must remember that our two-dimensional function ***f(x, y)*** now represents the distribution of gray levels:

The two-dimensional representation of the Fourier transform taken over an image whose x and y dimensions are each of value N is:

$$F(u,v) = \frac{1}{N^2} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y) \exp\left[-\frac{2\pi i}{N}(ux + vy)\right]$$

$$\text{for } u = 0, 1, 2, \ldots, N-1; \qquad v = 0, 1, 2, \ldots, N-1$$

The corresponding inverse relationship is:

$$f(x,y) = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} F(u,v) \exp\left[+\frac{2\pi i}{N}(ux + vy)\right]$$

$$\text{for } x = 0, 1, 2, \ldots, N-1; \qquad y = 0, 1, 2, \ldots, N-1$$

In order to obtain a representation shown in ***figure (c)***, we must evaluate what is called the Fourier spectrum.

Using the expansion for exp in sin and cos, the components of the above expression can be written as:

$$R(u) = \frac{1}{N} \sum_{x=0}^{N-1} f(x) \cos 2\pi ux/N$$

$$I(u) = \frac{1}{N} \sum_{x=0}^{N-1} f(x) \sin 2\pi ux/N$$

$$\text{for } u = 0, 1, 2, \ldots, N-1$$

R(u) and **I(u)** are said to represent the real and imaginary parts.

The Fourier spectrum is then defined by the formula

$$|F(u)| = [R^2(u) + I^2(u)]^{\frac{1}{2}}$$

$$\text{for } u = 0, 1, 2, \ldots, N-1$$

In two dimensions the corresponding formulation is

$$|F(u,v)| = [R^2(u,v) + I^2(u,v)]^{\frac{1}{2}}$$

for $u = 0, 1, 2, \ldots, N-1$; $v = 0, 1, 2, \ldots, N-1$

Where:

$$R(u,v) = \frac{1}{N^2}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y)\cos\frac{2\pi}{N}(ux + vy)$$

$$I(u,v) = \frac{1}{N^2}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y)\sin\frac{2\pi}{N}(ux + vy)$$

For each value of ***u*** and ***v*** we determine our corresponding value **| *F(u, v)* |** which gives the spectral representation of the type shown in figure (c). In general images containing periodic or regular features whether visible to the eye or not, the spectral representation provides us with a highly effective means of representation of those features for subsequent analysis.

THE FAST FOURIER TRANSFORM (FFT)

Computation of the Fourier transform using a standard approach requires of order N^2 operations for N data points. This results from the fact that for every single value of ***u*** in:

$$F(u) = \frac{1}{N}\sum_{x=0}^{N-1} f(x)\exp(-2\pi iux/N)$$

for $u = 0, 1, 2, \ldots, N-1$ with $i = \sqrt{-1}$

we have to evaluate the summation over N data values, which gives a computational dependence of around N^2 in terms of the number of operations required.

However, the computational dependence could be reduced from order N^2 to order $N\log_2 N$. Clearly for large N this yields a drastic reduction in terms of the number of operations and hence corresponding computer time. For example, for N = 256, N^2 = 65536, whereas $N\log_2 N$ = 2048.

The method is based on the fact that the Fourier transform may be written in the form:

$$F(u) = \frac{1}{2}\left[\frac{1}{M}\sum_{x=0}^{M-1} f(2x)W_M^{ux} + \frac{1}{M}\sum_{x=0}^{M-1} f(2x+1)W_M^{ux}W_{2M}^{u}\right]$$

for $u = 0, 1, 2, \ldots, M-1$

with $W_N = \exp(-2\pi i/N)$

N is of the form $N = 2^n$ where n is a positive integer. Based on this, N can be expressed as N = 2M where M is also a positive integer.

If we now further define:

$$F_{even}(u) = \frac{1}{M}\sum_{x=0}^{M-1} f(2x)W_M^{ux}$$

for u = 0, 1, 2, . . . , M-1 and

$$F_{odd}(u) = \frac{1}{M}\sum_{x=0}^{M-1} f(2x+1)W_M^{ux}$$

for u = 0, 1, 2, . . . , M–1

equation above may be written as:

$$F(u) = \tfrac{1}{2}\{ F_{even}(u) + F_{odd}(u)\, W^{u}_{2M}\}$$

Also , since $\mathbf{W_M^{u+M} = W^u_M}$ **and** $\mathbf{W_{2M}^{u+M} = -W^u_{2M}}$ it follows from equations above

$$F(u + M) = \tfrac{1}{2}\{ F_{even}(u) - F_{odd}(u)\, W^{u}_{2M} \}$$

Consider the above equations in relation to a transform consisting of two data values i.e. M = 1. Then the above equations may be written:

$$F_{even}(0) = f(0)$$
$$F_{odd}(0) = f(1)$$
$$F(0) = \frac{1}{2}\{F_{even}(0) + F_{odd}(0)W_2^0\} = \frac{1}{2}\{f(0) + f(1)W_2^0\}$$
$$F(1) = \frac{1}{2}\{F_{even}(0) - F_{odd}(0)W_2^0\} = \frac{1}{2}\{f(0) - f(1)W_2^0\}$$

Computationally therefore, the two Fourier coefficients F(0) and F(1) are simply determined by firstly calculating $F_{even}(0)$ and $F_{odd}(0)$ which are just the data values f(0) and f(l) thus requiring no multiplications or additions. One multiplication of $F_{odd}(0)$ by W^0_2 and one addition yields the coefficient F(0).

Similarly F(l) can be determined by a simple subtraction (computationally the equivalent operation to addition) of $F_{odd}(0)\ W^0{}_2$ from $F_{even}(0)$.

Therefore the total number of operations required for a two point transform consists of one multiplication and two additions. It may be shown, using the principle of induction, that for $N = 2^n$ data values the number of complex multiplications and additions is ½ N^n and N^n respectively.

The importance of the above illustrations is that it is theoretically possible by splitting our set of N data values into two point pairs, repeated application of equations enables us to determine the total transform in a time proportional to $N\log_2 N$. Before application of the equations we need to restructure our original set of data values into appropriate pairs using a method known as “Successive Doubling”.

To illustrate the successive doubling approach consider the function f(x) defined over the range of values for which x = 0, 1, . . . , 7 to yield a set of values:

f(0), f(l), f(2), f(3), f(4), f(5), f(6), f(7)

The first step is to divide the array into so-called even and odd parts as follows:

even parts : f(0), f(2), f(4), f(6)

odd parts: f(1), f(3), f(5), f(7)

Each of the above represents an individual array which may be further subdivided into even and odd parts to give for the first array:

Even part f(0), f(4)

Odd part f(2), f(6)

and for the second array:

Even part f(l), f(5)

Odd part f(3), f(7)

Input into the above equations now consists of the set of two point transforms:

f(0), f(4); f(2), f(6);

f(l), f(5); f(3), f(7);

A constraint of the method is that the total number of data values must be expressible as a power of two. In practice this may not always be the case so in order to satisfy the constraint we simply 'pad' our data set to the next power of two with zeros.

That is if our set corresponded to values for f(x) from x = 1, 2, . . . , 450 for example, we would simply generate values for f(x) all of zero magnitude for x = 451, 452,. . ., 512.

COMPUTING THE FOURIER TRANSFORM OF AN IMAGE

Since most images consist of a two-dimensional array of gray levels and we have only concerned ourselves so far with the one-dimensional determination of the Fourier Transform, we now address the question of extending our analysis into two-dimensions.

The form for the Fourier Transform in two dimensions for a square or two dimensional array is:

$$F(u,v) = \frac{1}{N^2}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y)\exp[(-2\pi i/N)(ux+vy)]$$

$$u = 0, 1, 2, \ldots, N-1; \quad v = 0, 1, 2, \ldots, N-1$$

However, since ***exp(ux + vy)*** may be written as ***exp (ux)exp(vy)*** the above equation can be re-written as:

$$F(u,v) = \frac{1}{N^2}\sum_{x=0}^{N-1} \exp(-2\pi iux/N)\sum_{y=0}^{N-1} f(x,y)\exp(-2\pi ivy/N)$$

Notice that the expression in the summation over y values is simply a one-dimensional transform for a fixed value of x over the frequency values v = 0, 1, . . . , N-1.

This generates a set of Fourier coefficients ***F(x, v)***. The expression for ***F(u, v)*** may now be written as:

$$F(u,v) = \frac{1}{N^2}\sum_{x=0}^{N-1} F(x,v)\exp(-2\pi iux/N)$$

This in turn corresponds to a one-dimensional transform over the frequency values:

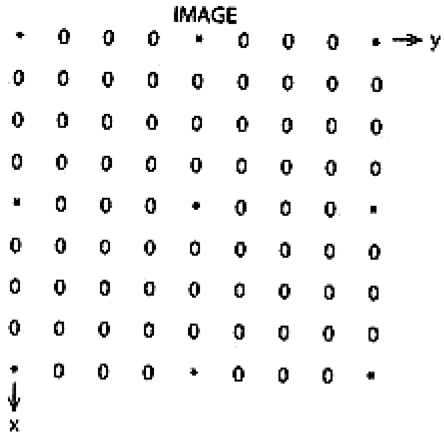
u = 0, *N - 1*

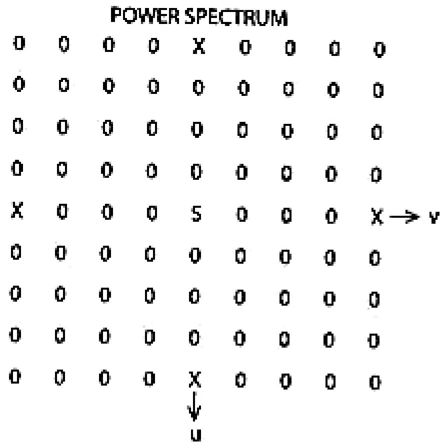
Therefore, we can regard our two-dimensional transform as consisting of a series of one-dimensional transforms which we would evaluate using the FFT algorithm.

Assuming the formerly defined ***axis convention*** for the image, the above representation means that the two-dimensional transform may be determined by firstly evaluating the coefficients ***F(x, v)*** from the transform of the image rows and then subsequently applying the transform of this (complex) set of coefficients along the image columns.

Note however that due to a translational property of the Fourier Transform, if we wish to display our spectrum of values with the origin at the centre so that the values are symmetrically disposed, we must multiply each gray scale value f(x, y) by $(-1)^{x+y}$ before invoking the transform process.

As a further illustration of the use of the Transform and its relation to the set of spectral values, consider the image:





The (*) symbol denotes a point in the image for which the gray scale value is greater than the background level (denoted as zero). The corresponding central part of the Fourier or Power spectrum obtained following the procedure outlined, would appear in the form as above.

S denotes the sum of all the pixel or gray levels in the image and corresponds to the 1^{st} coefficient, often called the origin term, in the spectrum.

X denotes values within the spectrum (greater than zero) occurring for values of u and v and correspond in position to the reciprocal spacings between the periodic non-zero elements (*) in the original image.

Application of the inverse transform to the components used to construct the spectrum would obviously yield the original image.

CONVOLUTION AND CORRELATION

If we have two sequences f(x) and g(x), then the convolution of these sequences is expressed as f(x)*g(x) where:

$$f(x) * g(x) = \sum_{x=0}^{K-1} f(x')g(x - x')$$

where *g(x-x')* corresponds to a translation *x'* of the function *g(x).*

It can be shown mathematically that the above summation can be determined through the use of Fourier transforms, which is computationally more efficient than a straightforward evaluation, given the availability of the fast Fourier transform.

To achieve this requires the use of the convolution theorem, which states that the convolution of two functions, is simply the inverse transformation of the product of the Fourier transforms of the two functions.

That is, if ***f(x)*** has the Fourier transform ***F(u)*** and ***g(x)*** has the Fourier transform ***G(u)***, then the convolution of the two functions may be written as:

$$T^{-1} \{ F(u)\, G(u) \}$$

where T^{-1} denotes the inverse transform.

Convolution, as already mentioned, plays an important part in image processing applications. For example, if we have an image with graylevel distribution f(x, y) and a function h(x, y) which is designed to highlight some specific features of the image, we may write using the convolution theorem:

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

where $H(u, v)$ and $F(u, v)$ are respective Fourier transforms of $h(x, y)$, and $f(x, y)$.

Then our new "highlighted" image is simply:

$$g(x, y) = T^{-1}\ [G(u, v)]$$

Similarly, it is known that in certain images, degradation of the image takes place due to "noise" or unwanted values which may result in blurring of the original image. If the "noise" can be quantified and in certain image applications this happens to be the case, then we can say that our resultant image or the image that is actually recorded, can be represented as a convolution of the original image with a "noise" function.

Mathematically, if the Fourier transform of our "noise" function is $N(u, v)$ say and $F(u, v)$ corresponds to the Fourier transform of our original image then:

$$g(x, y) = T^{-1}\ \{\ F(u, v)N(u, v)\ \}$$

What we have actually recorded is $g(x, y)$ and ideally we want to see the "noise free" image $f(x, y)$. Therefore we may write:

$$\mathrm{F(u,\ v)} = \frac{\mathrm{T\{g(x,\ y)\}}}{\mathrm{N(u,\ v)}} = \frac{G(u,v)}{N(u,v)}$$

Then the "noise free" image ie ***f(x, y)***, would result from the inverse transform of ***F(u, v)***.

The above process which involves the division of Fourier transform coefficients is referred to as deconvolution.

To implement the above processes computationally requires some ***pre-processing*** of the data. This arises from the fact that in forming the Fourier products above for two arrays of different dimensions, errors in the higher order coefficients will result. This is due to the Fourier transform itself being a periodic function.

To overcome the problem requires that the value of K is chosen to be consistent with the periodicity of the transform function. This requires that for two functions f(x), g(x) each of length N_1 and N_2 say, K should be chosen to be of a value such that:

$$K \geq N_1 + N_2 - 1$$

For practical purposes we generally choose K for which:

$$K = N_1 + N_2 - 1$$

Although the above analysis has been illustrated in relation to the one dimensional situation, the extension to two-dimensions means that instead of forming products between the Fourier coefficients for F(u) and G(u) we simply determine the products for F(u, v) and G(u, v).

EVALUATION OF THE INVERSE TRANSFORM

Modifications to the Fourier spectrum of any coefficients in the Fourier transform of the image, e.g. convolution of the image transform with a specified function, means that we will require to see the effect of such a modification on the original image.

This requires the use of the inverse Fourier transform to take us back from the domain of reciprocal space to image or real space.

This transformation as defined above may be written:

$$f(x) = \sum_{u=0}^{N-1} F(u) \exp(-2piux / N)$$

Taking the complex conjugate of the above and dividing both sides by N gives:

$$\frac{1}{N} f^*(x) = \frac{1}{N} \sum_{u=0}^{N-1} F^*(u) \exp(-2piux / N)$$

(*denotes complex conjugate).

Compare this result with the expression for the Fourier transform ie:

$$F(u) = \frac{1}{N} \sum_{x=0}^{N-1} f(x) ep(-2pux / N)$$

The above comparison shows that the input of a set of values ***F*(u)*** into an algorithm designed to compute the forward transform yields the quantity ***1/N f*(x).*** Multiplication by N and taking the corresponding complex conjugate give the required values for f(x).

SAMPLING

The periodicity of the Fourier transform ***imposes a constraint*** on our sampling interval of f(x) if we are to successfully resolve the required information in the transformation process.

This sampling criteria is known as the ***Whittaker-Shannon sampling theorem*** and states that if we require to resolve information within the function spaced at units distance 'd' apart, then we must sample our function at a minimum interval of 'd/2' units.

In practice however, our sampling is generally chosen to be of a finer interval than the above, which is based on a “noise-free” distribution of data values f(x), as most data contains some noise due to imperfections which may result during recording of the original data.