Determining How Well Two Pictures Match

Given images I_1 and I_2:

1. Sum magnitude of differences between pixels: $\Sigma_x \Sigma_y |I_1 - I_2|$.
2. Sum squares of differences: $\Sigma_x \Sigma_y |I_1 - I_2|^2$.
3. Based on Schwartz Inequality: For any two integrable functions (real-valued, non-negative, not identically zero):

   $$(\int f \cdot g)^2 \leq \int f^2 \cdot \int g^2$$

   over any domain of integration for which integrals are all defined. The two are equal if and only if $g = cf$ for constant c (i.e. g is a scalar multiple of f),

Corollary 1

$$\frac{\int f \cdot g}{\sqrt{f^2 \cdot g^2}} \leq 1$$

Once again, the equality holds if $g = cf$.
Corollary 2

Define *cross-correlation:*

$$f \otimes g = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(u, v) g(u + x, v + y)\, dudv$$

Normalized cross-correlation operator:

$$\frac{f \otimes g}{\sqrt{\int \int f^2 \cdot \int \int g^2}} \leq 1$$

Cross-correlation

The **cross-correlation**[1] operation is defined as

$$f \otimes g \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(u,v)\, g(x+u, y+v)\, du\, dv$$

which takes in f and g and returns a function of variables x and y.
Corollary

This corollary follows from the **Schwarz inequality**:

$$\frac{f \otimes g}{\sqrt{\iint f^2}\sqrt{\iint g^2}} \begin{cases} = 1 & \text{at positions } (x,\, y) \text{ for which } g = cf, \\ \leq 1 & \text{otherwise.} \end{cases}$$

Interpretation 1: One possible interpretation of the above corollary is to use the following analogy.
Consider the notion of the angle between 2 vectors $\vec{x}$ and $\vec{y}$

$$\cos\theta = \frac{\vec{x} \cdot \vec{y}}{|\vec{x}||\vec{y}|}$$

where θ is the *angle* between the 2 vectors, and the denominator is normalizing the dot product of $\vec{x}$ and $\vec{y}$.
We can thus think of the above corollary as an operation to normalize the correlation of 2 functions, and to compute the "cosine of the angle" between the 2 functions being compared.

[1] Often called simply **correlation**.

Interpretation 2:

Suppose the domain of f where $f > 0$ is small, compared to the domain of g where $g > 0$. We can view f as a *template* (often called a *filter*) of some pattern.
The normalized cross-correlation of f and g is equal to 1 only at positions (x, y) in the domain of g where f and g match *exactly*[2] up to a scale factor.

[2] In reality, we may never achieve an exact match because of *noise*.

Now let's consider the *discrete* version of function f.
Suppose $F(X, Y)$ is an $M \times M$ "digital filter" (a digital image extended to include the possibility that certain pixel values may be negative.), where M is an odd number.
Suppose that $I(X, Y)$ is an $N \times N$ image, where $N \gg M$.

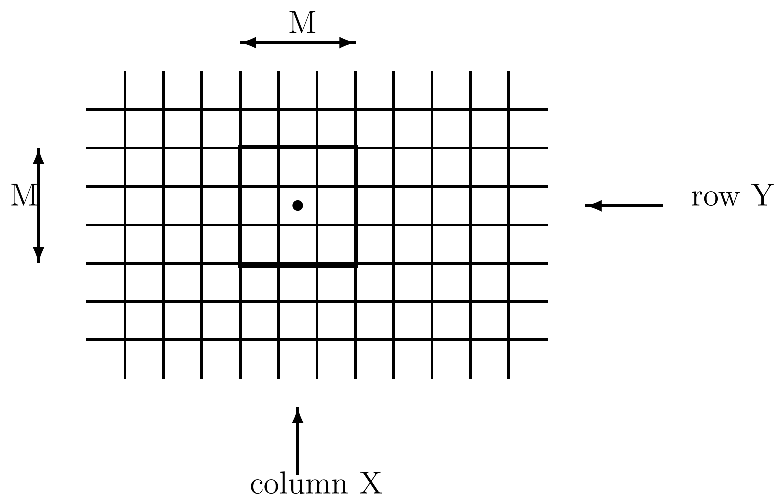

Figure 1: Discrete Filter

Suppose (X, Y) is the point we are looking at, the filter will be covering (with respect to image I) the range $[X - \lfloor M/2 \rfloor,\ X + \lfloor M/2 \rfloor]$ and $[Y - \lfloor M/2 \rfloor,\ Y + \lfloor M/2 \rfloor]$. (Figure 1 illustrates the case $M = 3$.)

The correlation process involves moving the filter F over a finite image I a pixel (a square grid in the above diagram) at a time, and to compute a new image $I'(X,Y)$, where

$$I'(X,Y) = \sum_{j=-\lfloor M/2 \rfloor}^{\lfloor M/2 \rfloor} \sum_{i=-\lfloor M/2 \rfloor}^{\lfloor M/2 \rfloor} F(i,j)\; I(X+i, Y+j)$$

For each point (X,Y) in image I, there are M^2 multiplications, and M^2-1 additions. Since there are N^2 points in image I, if $M \approx N$, then we will have $O(N^4)$ complexity. This is quite an expensive operation, and we would definitely be interested in trying to speed up the operation. We will see later how to speed up cross-correlation using convolution and the Fourier transform.

If part of the filter $F(X,Y)$ is out of the boundary of image I, then we may end up having *artifacts* in the resulting image I'.
This is known as the *boundary effect* and there are several ways to handle this problem:

1. use only a portion of the $N \times N$ image I defined inside the a particular region for correlation, such as $\lfloor M/2 \rfloor \leq X, Y \leq N - \lfloor M/2 \rfloor$.
2. assume that the magnitude of I outside its boundary is 0.
3. use the *wrap-around* method to define magnitude of I outside its boundary.

Convolution

The **convolution** operation is defined as

$$f * g \stackrel{\mathrm{def}}{=} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(u, v)\; g(x-u, y-v)\, du\, dv$$

The above equation can be rewritten as

$$f * g = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(-u, -v)\; g(x+u, y+v)\, du\, dv$$

It is obvious then that the convolution of $f(x, y)$ and $g(x, y)$ is identical to the correlation of $f(-x, -y)$ and $g(x, y)$, and the correlation of $f(x, y)$ and $g(x, y)$ is identical to the convolution of $f(-x, -y)$ and $g(x, y)$.

If f is an even function, then convolution and correlation are equivalent operations. Otherwise, we can think of convolution as a process which correlates $g(x, y)$ with a *flipped* version of the function f, that is, by replacing $f(x, y)$ with $f(-x, -y)$.

Here are some useful identities:

1. $f \otimes g = g \otimes f$
2. $f * g = g * f$
3. $(f * g) * h = f * (g * h)$
4. $f(x, y) * g(x, y) = f(-x, -y) \otimes g(x, y)$

Convolution Theorem

Let $f(x, y)$ and $g(x, y)$ be two functions with Fourier transforms $F(\omega_x, \omega_y)$ and $G(\omega_x, \omega_y)$ respectively. Let

$$h(x, y) = f(x, y) * g(x, y)$$

then

$$H(\omega_x, \omega_y) = F(\omega_x, \omega_y)\ G(\omega_x, \omega_y)$$

Fourier tranform has reduced convolution in the spatial domain to multiplication in the frequency domain. Therefore, the convolution theorem gives us the possibility to speed up convolution (and also correlation) provided that the Fourier transforms exist and can be computed efficiently.

$$\left.\begin{array}{l} f(x, y) \xrightarrow{\text{F.T.}} F(\omega_x, \omega_y) \\ g(x, y) \xrightarrow{\text{F.T.}} G(\omega_x, \omega_y) \end{array}\right\} \quad \xrightarrow{complexmultiplication} \quad H(\omega_x, \omega_y) \quad \xrightarrow{\text{I.F.T.}} \quad h(x,$$

The complexity can potentially be reduced from $O(N^4)$ to $O(N^2)$ if the transforms are sufficiently easy to compute.

Now, let's consider a system that takes in $f(x,y)$ and outputs $g(x,y)$.

Definition of Linear System

A system is *linear* if

$$\alpha\, f_1(x,y) + \beta\, f_2(x,y) \longrightarrow \alpha\, g_1(x,y) + \beta\, g_2(x,y)$$

where $g_1(x,y)$ is the output of $f_1(x,y)$ and $g_2(x,y)$ is the output of $f_2(x,y)$.

Definition of Shift-Invariant System

A system is *shift-invariant* if

$$f(x-\alpha, y-\beta) \longrightarrow g(x-\alpha, y-\beta)$$

for arbitrary α and β.

Linearity and shift-invariance are 2 main properties in image processing that are of particular interests to us.

Characterization Theorem
Any linear, shift-invariant operation can be expressed in terms of a convolution.
Convolution is a linear and shift-invariant operation.

Spatial Filtering

Suppose $F(\omega_x, \omega_y)$ is the Fourier transform of $f(x, y)$.
Suppose we multiply $F(\omega_x, \omega_y)$ by some function $H(\omega_x, \omega_y)$, H defines a *linear spatial filter* where H is called a *transfer function* of the filter.

$$F'(\omega_x, \omega_y) = H(\omega_x, \omega_y)\ F(\omega_x, \omega_y)$$

Let $f'(x, y)$ be the function whose Fourier transform is $F'(\omega_x, \omega_y)$, then

$$f'(x, y) = h(x, y) * f(x, y)$$

where $h(x, y)$ is function whose Fourier transform is $H(\omega_x, \omega_y)$.
$h(x, y)$ is called the *impulse response* or *point spread function* of the filter.