Question 1: Assignment 5: CS 663, Fall 2024

Amitesh Shekhar
IIT Bombay
22b0014@iitb.ac.in

Anupam Rawat
IIT Bombay
22b3982@iitb.ac.in

Toshan Achintya Golla
IIT Bombay
22b2234@iitb.ac.in

November 07, 2024

1. Read Section 1 of the paper ‘An FFT-Based Technique for Translation, Rotation, and Scale-Invariant Image Registration’ published in the IEEE Transactions on Image Processing in August 1996. A copy of this paper is available in the homework folder.

   (a) Describe the procedure in the paper to determine translation between two given images. What is the time complexity of this procedure to predict translation if the images were of size $N \times N$? How does it compare with the time complexity of pixel-wise image comparison procedure for predicting the translation?

   (b) Also, briefly explain the approach for correcting for rotation between two images, as proposed in this paper in Section II. Write down an equation or two to illustrate your point.

   [10+10=20 points]

Soln:

(a) **Determining the Translation**:-

For determining the translation (x_0, y_0) between two images $f_1(x, y)$ and $f_2(x, y)$ such that $f_2(x, y) = f_1(x - x_0, y - y_0)$, we can proceed as follows:

- Step 1 : Take the Fourier transforms of f_1 and f_2, giving $F_1(\mu, \nu)$ and $F_2(\mu, \nu)$.
- Step 2 : Apply the Fourier Shift theorem to get:

$$F_2(\mu, \nu) = F_1(\mu, \nu) e^{-j\frac{2\pi}{N}(\mu x_0 + \nu y_0)}.$$

  Compute the cross-power spectrum of the two images:

$$C(\mu, \nu) = \frac{F_2^*(\mu, \nu) F_1(\mu, \nu)}{|F_2(\mu, \nu)||F_1(\mu, \nu)|}$$

  where $F_2^*(\mu, \nu)$ denotes the complex conjugate of $F_2(\mu, \nu)$.
  The cross-power spectrum turns out to be:

$$C(\mu, \nu) = e^{-j\frac{2\pi}{N}(\mu x_0 + \nu y_0)}$$

- Step 3 : The inverse Fourier transform of $C(\mu, \nu)$ turns out to be:

$$F^{-1}\left(e^{j\frac{2\pi}{N}(\mu x_0 + \nu y_0)}\right) = \delta(x + x_0, y + y_0),$$

  where $\delta(x + x_0, y + y_0)$ is a delta function that is zero everywhere except at $(-x_0, -y_0)$.

The translation between the two images is then given by inverting the signs of the observed nonzero point. Clearly, we have determined the displacement using the above steps.

Time complexity:-

The time complexity of this procedure is $O(N^2 \log N)$, where $N \times N$ is the size of the images. This is because the Fourier Transform can be computed in $O(N^2 \log N)$ time, and the remaining operations (like element-wise multiplication and normalization used in cross-power spectrum computation) are $O(N^2)$. So, steps 1 and 3 have a time complexity of $O(N^2 \log N)$ each, while step 2 has a time complexity of $O(N^2)$, resulting in an overall time complexity of $O(N^2 \log N)$.

On the other hand, a pixel-wise image comparison to estimate translation involves sliding one image over the other, resulting in a time complexity of $O(N^4)$ for $N \times N$ images. If we use use a window of size $W \times W$ for the pixel comparison, we would have a time complexity of $O(N^2 W^2)$. Therefore, the FFT-based method is much faster, especially for large images.

(b) **Correcting the rotation**:-
Let there be two images $f_1(x, y)$ and $f_2(x, y)$, where $f_2(x, y)$ is a rotated and translated version of $f_1(x, y)$, such that:

$$f_2(x, y) = f_1(x \cos\theta_0 + y \sin\theta_0 - x_0, -x \sin\theta_0 + y \cos\theta_0 - y_0)$$

This means that $f_2(x, y)$ is obtained by rotating $f_1(x, y)$ by an angle θ_0 and then translating it by (x_0, y_0).
Now, let $F_1(u, v)$ and $F_2(u, v)$ represent the Fourier transforms of $f_1(x, y)$ and $f_2(x, y)$, respectively.
According to the Fourier Rotation Theorem, the Fourier transform of the rotated image $f_2(x, y)$ will be a rotated version of the Fourier transform of the original image $f_1(x, y)$. Specifically, we have:

$$F_2(u, v) = F_1(u \cos\theta_0 + v \sin\theta_0, -u \sin\theta_0 + v \cos\theta_0)$$

Translation in the spatial domain corresponds to a phase shift in the frequency domain. The translation by (x_0, y_0) introduces a phase shift of $e^{-i2\pi(ux_0+vy_0)}$ in the Fourier transform. Thus, we can write:

$$F_2(u, v) = F_1(u \cos\theta_0 + v \sin\theta_0, -u \sin\theta_0 + v \cos\theta_0) \cdot e^{-i2\pi(ux_0+vy_0)}$$

So, the magnitudes M_1 and M_2 are related as follows:

$$M_2(u, v) = M_1(u \cos\theta_0 + v \sin\theta_0, -u \sin\theta_0 + v \cos\theta_0)$$

Clearly, the magnitudes of both $F_1(u, v)$ and $F_2(u, v)$ are same, but one of them is a rotated replica of the other. Using polar coordinates, we can write :

$$M_1(\rho, \theta) = M_2(\rho, \theta - \theta_0)$$

A shift in the angle θ by θ_0 in the polar coordinates is equivalent to Rotation by θ_0 in Cartesian coordinates. By applying the shift theorem on M_1 and M_2 and then finding their cross-power spectrum, we will get a peak at $-\theta_0$, and hence the rotation angle is negative of the value of θ at which the peak was obtained.