Exercise 2

Part (a)

The Fourier shift theorem states that translating a function in space domain is equivalent to multiplication by a complex exponential in the Fourier domain. Mathematically, if we translate an image f_1 by (x_0, y_0) to generate another image f_2, that is, $f_2(x,y) = f_1(x-x_0, y-y_0)$, then in the Fourier domain, corresponding transforms F_1 and F_2 are related as:

$$\mathcal{F}(f_2(x,y)) = \mathcal{F}(f_1(x-x_0, y-y_0)) \longrightarrow F_2(\mu,\nu) = e^{-j2\pi(\mu x_0 + \nu y_0)} F_1(\mu,\nu)$$

The cross-power spectrum of two images f_1 and f_2, with Fourier transforms F_1 and F_2, is defined as:

$$\frac{F_1(\mu,\nu)F_2^*(\mu,\nu)}{|F_1(\mu,\nu)F_2(\mu,\nu)|} = e^{j2\pi(\mu x_0 + \nu y_0)}$$

The image registration technique described in this paper leverages the fact that the phase of the cross-power spectrum is equal to the phase difference between the translated and original image. Therefore, an inverse Fourier transform of the cross-power spectrum yields a Dirac delta function at $(-x_0, -y_0)$, which is the translation needed to align the two images with mutual displacement.
The key steps involved in the entire procedure outlined here are two Fourier transforms and one Inverse Fourier Transform to measure the translation. These can be done using two Fast Fourier Transforms ($\mathcal{O}(N\log N)$) and one Inverse Fourier Transform ($\mathcal{O}(N\log N)$). Therefore, the computational complexity of this new method is $\mathcal{O}(N\log N)$. However, the time complexity of pixel-wise image comparison procedure for predicting the translation would be $\mathcal{O}(N^2)$ since there are a total of N^2 translations possible for an $N\times N$ image. Therefore, the time complexity of the method proposed in this paper is lower than that of the brute-force search.

Part (b)

Consider two images f_1 and f_2 where the translation is by (x_0, y_0) and rotation angle is θ_0, then:

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

In the Fourier domain, this can be written as follows using translation and rotation property:

$$F_2(\mu,\nu) = e^{-j2\pi(\mu x_0+\nu y_0)} F_1(\mu\cos\theta_0 + \nu\sin\theta_0 - \mu\sin\theta_0 + \nu\cos\theta_0)$$

The magnitudes of both these quantities are the same, but one is a rotated version of the other. In polar coordinates (ρ,θ):

$$Mag(F_1)(\rho,\theta) = Mag(F_2)(\rho,\theta-\theta_0)$$

Therefore, using the phase correlation technique described in part (a), the rotation angle θ_0 can be computed. Consider two images f_1 and f_2, where f_1 is a scaled version of f_2 by factors of (a,b). Then, in the Fourier domain, we have the following:

$$F_2(\mu,\nu) = \frac{1}{|ab|}F_1(\frac{\mu}{a},\frac{\nu}{b}) \longrightarrow F_2(\log\mu,\log\nu) = F_1(\log\mu - \log a, \log\nu - \log b)$$

where in the right step, we have converted the axes to logarithmic scales and ignored the factor of $\frac{1}{|ab|}$. This is now equivalent to a translation movement, which can be computed using the phase correlation technique described in part (a). In polar coordinates, the problem statement with scaling factors (a,a) and rotation θ_0 can be defined using $\rho_1 = \sqrt{x^2+y^2}$, $\theta_1 = \arctan\frac{y}{x}$, which is being converted to $\rho_2 = \sqrt{\frac{x}{a}^2 + \frac{y}{a}^2} = \frac{\rho_1}{a}$, $\theta_2 = \arctan\frac{y/a}{x/a} = \theta_1$. Therefore, Fourier magnitude spectra of the two images will be related by:

$$M_1(\rho,\theta) = M_1(\frac{\rho}{a},\theta-\theta_0) \longrightarrow M_1(\log\rho,\theta) = M_1(\log\rho - \log a, \theta-\theta_0)$$

Again, this is equivalent to translation movement in $(\log\rho,\theta)$ space. Hence, the displacement can be computed using the phase correlation technique described in part (a).