2010 Paper 8 Question 4

Computer Vision

(a) Explain how each of the following equations or expressions can be used for detecting and estimating visual motion in a spatio-temporal image sequence $I(x, y, t)$. Include in each answer also the name used to describe each of these general classes of motion extraction models:

(i) $$-\frac{\partial I(x,y,t)}{\partial t} = \vec{v} \cdot \vec{\nabla} I(x,y,t)$$ [3 marks]

(ii) $$-\frac{\partial}{\partial t}\left[\nabla^2 G_\sigma(x,y) * I(x,y,t)\right]$$ [3 marks]

(iii) $$\text{argmax} \int_x \int_y \int_t I(x,y,t) \cdot I(x - v_x\tau,\ y - v_y\tau,\ t - \tau)\ dx\ dy\ dt$$ [3 marks]

(iv) $$F(\omega_x, \omega_y, \omega_t) = e^{-i(\omega_x v_x \tau\ +\ \omega_y v_y \tau\ +\ \omega_t \tau)} F(\omega_x, \omega_y, \omega_t)$$
where $$F(\omega_x, \omega_y, \omega_t) = \int_x \int_y \int_t I(x,y,t) e^{-i(\omega_x x + \omega_y y + \omega_t t)}\ dx\ dy\ dt$$
[3 marks]

(b) Explain the key principles underlying the Scale-Invariant Feature Transform (SIFT). What is it used for, what goals does it achieve, and how does it achieve them? [5 marks]

(c) Suppose you were trying to design a machine vision system based as closely as possible upon human vision; for example, perhaps a visual prosthesis for the blind. Would you aim to include, as design goals, the standard geometric visual illusions as well? If such errors of vision emerged as unintended consequences of your design, would you consider them to be features, or bugs, and why? [3 marks]