Tentamen Beeldbewerken
23 Oktober 2012

Geef **korte antwoorden**. Telegram stijl mag! Een goed antwoord kan verdrinken in teveel opschrijven wat er niet toe doet (en daarmee fout gerekend worden). Beantwoord de **vragen in volgorde!**

NB. vragen 1 t/m 3 moeten alleen worden gemaakt als je de deeltoets wilt herkansen (of daaraan niet hebt deelgenomen).

1. **Kleinste Kwadraten Schatter** Van een functie in 1 variabele zijn de meetpunten (x_i, f_i) voor $i = 1, \ldots, n$ gegeven. De meetwaarden f_i zijn verstoord door ruis. We hebben goede reden aan te nemen dat we de functie redelijk goed beschreven kan worden met een functie van de vorm:

$$f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4$$

waarin $a_0, \ldots, a_4$ constante maar onbekende parameters zijn.

   (a) Stel het lineaire model $f = Xp$ op waarin X de modelmatrix is, p de parameter vector en f de vector met functiewaarden. Wat zijn in dit geval de preciese definitie en afmetingen van de matrix X en de vectoren p en f?

   (b) De kleinste kwadraten methode vindt een p die maakt dat $Xp - f$ zo klein mogelijk wordt. Het resultaat is dat het stelsel $X^T X p = X^T f$ opgelost moet worden. Maak duidelijk in 3 zinnen welke stappen daartoe nodig zijn (een wiskundige afleiding is niet nodig maar wel begrip van de te volgen procedure).

2. **Homogene Coordinaten.** Gegeven de transformatie matrices:

$$T_1 = \begin{pmatrix} 1 & 0 & 5 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix}, \quad T_2 = \begin{pmatrix} 1 & 0.5 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad T_3 = \begin{pmatrix} 5 & 2 & 1 \\ 2 & 3 & 1 \\ 1 & 1 & 1 \end{pmatrix}, \quad T_4 = \begin{pmatrix} 2 & -1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & \sqrt{5} \end{pmatrix}$$

   (a) Gegeven het eenheids vierkant met hoekpunten $(0,0)$, $(1,0)$, $(1,1)$ en $(0,1)$. Wat zijn de afbeeldingen van dit vierkant na ieder van de bovenstaande transformaties? Maak voor alle vier de transformaties een schets.

   (b) Geef van alle 4 de transformaties aan of het een translatie, een rotatie, een affiene transformatie dan wel een projectieve transformatie is.

   (c) Gegeven 3 punten (x_1, y_1), (x_2, y_2) en (x_3, y_3) en de corresponderende punten (x'_1, y'_1), (x'_2, y'_2) en (x'_3, y'_3) waarbij gegeven is dat

$$\begin{pmatrix} x'_i \\ y'_i \\ 1 \end{pmatrix} \sim A \begin{pmatrix} x_i \\ y_i \\ 1 \end{pmatrix}$$

   en A een affiene transformatie is. Beschrijf de procedure om A te schatten uit de deze 3 (of mogelijk meer) punt correspondenties.

3. **Locale Structuur I.**

   (a) Leg uit dat:

$$\frac{\partial}{\partial x}(f * G^s) = f * \frac{\partial}{\partial x} G^s$$

   Waarom is deze eigenschap zo belangrijk in de praktijk als we te maken hebben met discrete beelden?

   (b) De Gauss functie in 1 variabele is gegeven als

$$G^s(x) = \frac{1}{s\sqrt{2\pi}} \exp\left(-\frac{x^2}{2s^2}\right)$$

   wat is de Gauss functie in 2 variabelen?

   (c) Wat bedoelen we met het "separeerbaar zijn" van de Gauss functie en waarom is dat van belang bij de convolutie met een Gauss functie?

   (d) Laat zien dat niet alleen de convolutie met de Gauss functie, $f * G^s$, separeerbaar is, maar dat ook de afgeleide $f * \partial_x G^s$ separeerbaar is.

(e) Als je Gaussische convolutie in code omzet (gebruik makend van een convolutie functie) moet je de 1D Gauss functie discretiseren. Geef aan hoe u de 1D Gauss functie op schaal s sampled?

(f) Gegeven een klein 10×10 beeld f:

9	4	2	9	4	6	2	10	7	7
5	10	4	4	3	5	6	4	5	$\boxed{8}$
3	1	7	1	3	5	2	8	10	9
5	4	9	5	1	9	1	1	7	10
10	0	7	9	$\boxed{5}$	2	6	9	2	8
1	9	0	10	8	7	1	0	6	6
7	5	6	3	4	0	4	3	3	9
7	1	4	$\boxed{3}$	5	3	2	2	9	5
10	4	3	5	3	3	8	3	5	2
6	8	5	0	0	10	8	8	0	7

wat is de waarde van de convolutie $f * w$ op de omkaderde punten in het bovenstaande beeld? Hierbij is de convolutie weeg functie

$$w = \left\{ \begin{array}{ccc} 1 & \underline{1} & 1 \\ & 1 & \end{array} \right\}$$

waarin het onderstreepte element de oorsprong aangeeft.

4. **Locale Structuur II.**

(a) Geef de definitie van de gradient vector? En wat is de geometrische interpretatie van de gradient vector?

(b) Gegeven het onderstaande beeld, geef op de gemarkeerde posities de assen van het gradient coordinaten systeem weer. Maak een schets op je antwoordvel (niet nodig om het grijswaarden landschap te schetsen.

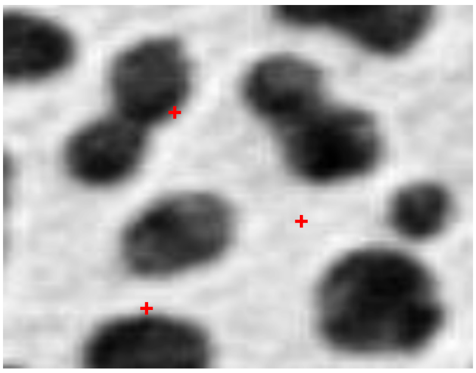

Figure 1: Sketch the gradient vector at the three indicated positions.

(c) De Canny edge detector wordt bepaald door 2 condities:

$$f_w \gg 0, \qquad f_{ww} \approx 0$$

Leg uit dat met deze condities inderdaad edges in 2D beelden worden gekarakteriseerd.

(d) Beschrijf kort hoe u de Canny edge detector heeft geimplementeerd. Daarbij is gegeven:

$$f_{ww} = \frac{1}{f_x^2 + f_y^2} \left(f_x^2 f_{xx} + 2 f_x f_y f_{xy} + f_y^2 f_{yy} \right)$$

(e) Als het beeld wordt geroteerd over hoek ϕ dan veranderen natuurlijk ook de afgeleiden in x en y richting. Wat heeft dat voor gevolg voor de gradient vector?

5. **Scale Space.**

   (a) Beschouw een scale space van het beeld f_0 bestaande uit de n beelden:

   $$f(x, y, s_i) = (f_0 * G^{s_i})(x, y)$$

   voor $i = 1, \dots, n$ waarbij $s_1 < s_2 < \cdots < c_n$. Waarom kunnen we het beeld op schaal s_i ook berekenen door een Gaussische convolutie van het beeld op schaal s_{i-1}?

   (b) Voor welke waarde van t geldt dat $f_0 * G^{s_i} = (f * G^{s_{i-1}}) * G^t$?

   (c) Wat is een verstandige keuze voor de scale sampling, dus hoe kies je de getallen $s_1, \dots, s_n$?

   (d) Wat zijn scale normalized derivatives en waarom zijn ze nodig?

   (e) Als je de Canny edge detector wilt normaliseren naar de schaal op welke wijze veranderen dan de twee condities $f_w \gg 0$ en $f_{ww} \approx 0$? Kunt u een reden bedenken waarom je de Canny edge detector zou willen normaliseren naar de schaal?

6. **SIFT.**

   (a) The scale invariant feature transform (SIFT) kent de volgende stappen in de berekeningen:

      i. key point localisatie

      ii. orientatie bepaling

      iii. descriptor berekening (zonder de normalisatie t.b.v. de invariantie onder veranderingen van de gemiddelde brightness van een beeld, dat komt in een later onderdeel aan de orde).

   Geef van al deze stappen een zeer korte uitleg.

   (b) Leg uit waarom SIFT schaal invariant en rotatie invariant is.

   (c) SIFT is ook invariant onder vermenigvulding van de beeldfunctie met een constante. Deze invariantie houdt in dat als we beeld af bewerken (met a een constante factor) we dezelfde SIFT resultaten krijgen als zouden we krijgen door beeld f te bewerken. Lowe legt in zijn artikel deze laatste invariantie als volgt uit:

   > Finally, the feature vector is modified to reduce the effects of illumination change. First, the vector is normalized to unit length. A change in image contrast in which each pixel value is multiplied by a constant will multiply gradients by the same constant, so this contrast change will be canceled by vector normalization. A brightness change in which a constant is added to each image pixel will not affect the gradient values, as they are computed from pixel differences. Therefore, the descriptor is invariant to affine changes in illumination. However, non-linear illumination changes can also occur due to camera saturation or due to illumination changes that affect 3D surfaces with differing orientations by different amounts. These effects can cause a large change in relative magnitudes for some gradients, but are less likely to affect the gradient orientations. Therefore, we reduce the influence of large gradient magnitudes by thresholding the values in the unit feature vector to each be no larger than 0.2, and then renormalizing to unit length. This means that matching the magnitudes for large gradients is no longer as important, and that the distribution of orientations has greater emphasis.

   Neem nu $\mathbf{v}$ als een feature vector (d.w.z. een descriptor) voordat de normalisaties zijn uitgevoerd. Lowe onderscheidt twee stappen in het normalisatie process. Welke zijn dat? Wat is de resulterende feature vector $\mathbf{v}'$ na deze twee stappen? Waarom is de volgorde van beide stappen essentieel?

   (d) (*extra points*) Lowe gebruikt in zijn SIFT niet de schaal genormaliseerde Laplaciaan:

   $$\ell(x, y, s) = s^2 (f_0 * \nabla^2 G^s)(x, y)$$

   om de keypoints te localiseren maar de difference of Gaussians:

   $$f_0 * G^{ks} - f_0 * G^s$$

   Laat zien dat dit een benadering is van $\ell(x, y, s)$.

7. **Pinhole Camera.** De projectie van een 3D punt (X, Y, Z) op het punt (pixel) (x, y) wordt vastgelegd met:

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \sim \begin{pmatrix} f_x & f_y & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} R^T & -R^T t \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

(a) Benoem de twee matrices in bovenstaande vergelijking en geef ook de betekenis van alle symbolen (f_x, f_y, α, u_0, v_o, R en t).

(b) Beschouw de camera waarvan gegeven is $f_x = f_y = 5$, $u_0 = v_0 = 0$, $\alpha = 0$, $R = I$ (de identiteits matrix) en $t = 0$. Gegeven 4 punten op het oppervlak $Y = -160$ met coordinaten $(-50, -160, 400)$, $(50, -160, 400)$, $(50, -160, 500)$ en $(-50, -160, 500)$. Bereken de projecties (x, y) op de retina van deze 4 gegeven punten. Maak een schets en verklaar wat de camera 'ziet'.

(c) Alle 4 de punten met $Y = -160$ komen op de retina terecht op een punt met negatieve y coordinaat. Waarom is dat zo?

(d) Neem beide matrices bijeen in 1 matrix P zodat:

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \sim P \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

Beschrijf hoe je de matrix P schat uit een aantal punt correspondenties, d.w.z. voorbeelden van (x_i, y_i) als geprojecteerd beeld van (X_i, Y_i, Z_i).