

Traitement et Synthèse de l'Image

Examen – 2 heures

L'épreuve comporte trois parties indépendantes. Les documents sont autorisés.

1 Exercice sur le cours de photographie (3 points)

En photographie, on dit généralement que « l'infini commence au-delà de 10 mètres ». On veut savoir si cette affirmation est exacte du point de vue de la profondeur de champ. Si la mise au point est faite « à l'infini », alors les points de l'espace objet dont l'image est nette ne sont pas limités à gauche, mais ils sont limités à droite par un plan Π^+ (notations du cours). On se demande si la distance de Π^+ au centre optique est supérieure à 10m.

1. On utilise un objectif « standard », de distance focale $f' = 50mm$, réglé à l'ouverture $NO = 5$. La mise au point est faite à l'infini. Calculer la taille t de la tache de flou correspondant à l'image d'un point objet P situé sur l'axe optique à $p = -10m$. Il est conseillé de s'aider d'un schéma.
2. On suppose que le récepteur photosensible de l'appareil photographique est un capteur CCD de taille $24mm \times 36mm$ constitué de 10 millions de pixels carrés. En supposant que les pixels couvrent entièrement la surface du CCD, calculer leur côté a . L'image de P est-elle floue ? Conclusion ?

2 Problème sur la partie “vision par ordinateur” (9 points)

On souhaite estimer l'éclairage \mathbf{S} d'une scène 3D, en utilisant un objet de géométrie et de couleurs connues, à savoir un cube à faces colorées (cf. FIGURE ??). Ce problème comporte trois parties qui peuvent être traitées indépendamment, bien qu'il soit conseillé de les traiter dans l'ordre indiqué.

2.1 Détection des sommets du cube et mise en correspondance 2D-3D

On dispose de deux photographies I et I_{app} légèrement différentes d'un même cube dans lesquelles les images de sept sommets du cube apparaissent (cf. FIGURE ??). Dans l'image droite I_{app} (dite d'apprentissage) les sommets ont été étiquetés avec les labels $C_i, i \in 1..7$. Dans l'image gauche I , légèrement différente, les mêmes sommets apparaissent mais ne sont pas labélisés. On veut détecter et mettre en correspondance les sommets de I avec ceux de I_{app} . On utilise pour cela le TP TSI n°2.

1. Le code Matlab de détection des points de Harris donne le résultat présenté dans la figure ?? . Pourquoi C_5 n'est-il pas détecté ? Les sommets du cube ne sont pas les seuls points d'intérêt détectés. Pourquoi la mise en correspondance utilisée dans le TP peut tout de même permettre d'apparier les sommets du cube présents dans I et I_{app} ?
2. Dans le code de détection des points de Harris, on utilise deux convolutions de l'image par les masques directionnels de Prewitt. Rappelez pourquoi. Les deux images suivantes montrent les résultats de ces convolutions. Indiquez le masque associé à l'image gauche (resp. droite). Compte tenu de ce résultat, est-ce qu'une détection des sommets C_i dans l'image I via une détection de contours est envisageable ? Comment opéreriez-vous ? (une réponse succincte suffit).

2.2 Calcul de la position du cube dans le repère de la caméra

On rappelle que la formation d'une photographie est décrite par l'équation de projection centrale

$$\mathbf{u} \sim \mathbf{P}\mathbf{X},$$

où $\mathbf{P} \in \mathbb{R}^{3 \times 4}$ désigne la matrice de projection, $\mathbf{X} \in \mathbb{R}^4$ désigne le vecteur des coordonnées cartésiennes augmentées d'un point dans le repère scène, et $\mathbf{u} \in \mathbb{R}^3$ désigne le vecteur des coordonnées cartésiennes augmentées de son image dans le repère 2D pixélique.

On associe à la scène un repère orthonormé (repère scène) tel que les huit coins du cube aient pour coordonnées cartésiennes $(\pm 1, \pm 1, \pm 1)$. On désignera par $\mathbf{C}_j \in \mathbb{R}^4$, $j \in \{1..8\}$ les vecteurs de ces coordonnées cartésiennes augmentées et par $\mathbf{c}_i \in \mathbb{R}^3$ les vecteurs des coordonnées pixéliques augmentées de leurs images.

1. Formulez le problème de *resection* consistant à estimer \mathbf{P} à partir des correspondances $\mathbf{c}_j \leftrightarrow \mathbf{C}_j$, en donnant, sous sa forme la plus adaptée au cas du cube, l'équation homogène linéaire en les éléments de \mathbf{P} décrivant la relation entre \mathbf{P} , \mathbf{c}_j et \mathbf{C}_j .

On rappelle que, pour toute paire de vecteurs $\mathbf{m} \in \mathbb{R}^3$, $\mathbf{n} \in \mathbb{R}^3$, on a l'égalité suivante :

$$\mathbf{m} \wedge \mathbf{n} = [\mathbf{m}]_{\wedge} \mathbf{n}$$

où

$$[\mathbf{m}]_{\wedge} = \begin{pmatrix} 0 & -m_3 & m_2 \\ m_3 & 0 & -m_1 \\ -m_2 & m_1 & 0 \end{pmatrix}$$

2. Donnez une définition à la notion de « points de fuite associés aux directions des arêtes ».
3. Que représentent les colonnes de \mathbf{P} ?
4. Précisez de quelle connaissance (à savoir une matrice inconnue notée \mathbf{G}) faudrait-il disposer, en plus de \mathbf{P} , pour calculer « directement », dans le repère scène, les normales des trois faces du cube visibles dans l'image ? Détaillez le calcul de ces normales que l'on supposera unitaires.
5. Expliquez comment il serait possible d'estimer cette matrice \mathbf{G} à partir de \mathbf{P} ?

2.3 Calcul de la direction d'éclairage dans le repère de la caméra

On souhaite maintenant calculer la direction d'éclairage \mathbf{S} dans le repère scène. On suppose qu'aucune face du cube ne se trouve dans l'ombre.

1. Montrez que la couleur est uniforme sur chaque face du cube, si l'éclairage est supposé parfaitement parallèle, si le matériau est supposé parfaitement lambertien et si on néglige le vignettage.
2. Montrez que l'estimation de \mathbf{S} dans un repère 3D lié au cube n'est possible que si on connaît les albédos des différentes faces du cube dans les canaux R, V, B (rouge, vert et bleu).
3. En supposant que l'albédo dans le canal R est égal à 1 pour toutes les faces, montrez que l'estimation de \mathbf{S} est très simple dans un repère 3D lié au cube, et donc dans le repère lié à la caméra.
4. Comparez cette "mire" de calibrage de l'éclairage avec celle qui a été utilisée en TP. En particulier, que peut-on dire de la présence de taches brillantes dues à la "réflexion spéculaire" ?

3 Problème sur la partie "synthèse d'images" (8 points)

3.1 Questions (4 points)

Il est demandé de répondre à chaque question de manière claire et concise (10 lignes maximum par réponse) :

1. Comment modifier la transformation 3D vers 2D pour faire une projection orthogonale ? Qu'est-ce que cela changerait au reste du rendu ?
2. Éclairage de Gouraud :
 - (a) Ce modèle d'éclairage simule une surface lisse : pourquoi ?
 - (b) Expliquez pourquoi une réflexion spéculaire ne peut pas arriver au milieu d'une face avec ce modèle d'éclairage.
 - (c) Comment pourriez-vous modifier les données géométriques et le calcul de l'éclairage pour pouvoir modéliser les arêtes vives (cf. FIGURE 2) ?
3. Expliquez le problème de la perspective quand on texture avec des coordonnées barycentriques calculées sur l'image.

3.2 Problème : “Environment Mapping” (4 points)

On dispose d’une image omnidirectionnelle d’une scène (l’“environnement”, représenté par exemple par une photographie panoramique), et on souhaite faire le rendu d’objets en simulant leur immersion dans cette scène à la position du point de vue de l’image d’environnement. On suppose qu’on dispose d’une fonction `getPixelAtAngle(image, theta, phi)` qui prend en paramètres l’image et une direction depuis le centre de la scène décrite par les angles `theta` et `phi` et renvoie la couleur du pixel de l’image omnidirectionnelle correspondant à cette direction.

1. Décrivez un algorithme permettant de simuler la réflexion de l’environnement décrit par l’image omnidirectionnelle sur un objet considéré comme un miroir parfait. Distinguez en particulier les traitements réalisés au niveau des sommets de ceux réalisés au niveau des pixels.
2. Quelles modifications apporter à cet algorithme pour des objets dont la surface est lambertienne ?

Traitement et Synthèse de l'Image

Examen – 2h – Notes de cours autorisées

L'épreuve comporte trois exercices indépendants.

1 Exercice 1 : reconstruction 3D par stéréophotométrie (6 points)

En stéréophotométrie (cf. TP1), on peut considérer la surface d'un objet comme le graphe $z = u(x, y)$ d'une fonction $u : \mathbb{R}^2 \rightarrow \mathbb{R}$, qui à tout point image (x, y) associe une « profondeur » $u(x, y)$, dans un repère 3D tel que Oxy coïncide avec le plan image et Oz avec l'axe optique de l'appareil photographique. La normale unitaire $\mathbf{N}(x, y)$ s'exprime alors en fonction du gradient $\nabla u(x, y) = [\partial_x u(x, y), \partial_y u(x, y)]^\top$:

$$\mathbf{N}(x, y) = \frac{1}{\sqrt{1 + \|\nabla u(x, y)\|^2}} [-\partial_x u(x, y), -\partial_y u(x, y), 1]^\top \quad (1)$$

On veut reconstruire le 3D d'un objet lambertien, d'albédo égal à 1, par stéréophotométrie. On dispose de n images de cet objet vu sous n éclairages différents. La loi de Lambert donne l'expression du niveau de gris $I_k(x, y)$ au point image (x, y) , sous l'éclairage \mathbf{S}_k (cf. TP1) :

$$I_k(x, y) = \max\{\mathbf{S}_k^\top \mathbf{N}(x, y), 0\} \quad k \in [1, n] \quad (2)$$

En l'absence d'ombre, le système (2) peut donc être réécrit sous forme matricielle :

$$\mathbf{I}(x, y) = \mathbf{T} \mathbf{N}(x, y) \quad (3)$$

Le vecteur $\mathbf{I}(x, y) = [I_1(x, y), \dots, I_n(x, y)]^\top$ et la matrice $\mathbf{T} = [\mathbf{S}_1, \dots, \mathbf{S}_n]^\top$ sont supposés connus.

1. Précisez les dimensions de la matrice \mathbf{T} . À quelle condition sur les vecteurs \mathbf{S}_k , $k \in [1, n]$, la pseudo-inverse de \mathbf{T} peut-elle permettre de résoudre le système (3) ? Que peut-on dire, en particulier, du cas où $n = 2$?
2. On suppose dorénavant que $n = 2$. En utilisant (1), réécrivez le système (3) sous la forme d'un système Σ de deux équations aux dérivées partielles (EDP) en $u(x, y)$.
3. On s'intéresse maintenant au cas où $u(x, y) = (x^2 + y^2)/2$. On suppose que les $n = 2$ éclairages valent $\mathbf{S}_1 = [s, s, c]^\top$ et $\mathbf{S}_2 = [-s, -s, c]^\top$, où $s = \sqrt{2}/2 \sin \alpha$, $c = \cos \alpha$ et $\alpha \in]0, \pi/2[$. Calculez la normale $\mathbf{N}(x, y)$, puis les niveaux de gris $I_1(x, y)$ et $I_2(x, y)$ des deux images (on suppose qu'il n'y a pas d'ombre).
4. Montrez que, dans ce cas, le système Σ implique les deux égalités suivantes :

$$\begin{cases} \|\nabla u(x, y)\|^2 = x^2 + y^2 \\ \partial_x u(x, y) + \partial_y u(x, y) = x + y \end{cases} \quad (4)$$

5. Montrez que le système (4) a deux solutions en $\partial_x u(x, y)$ et $\partial_y u(x, y)$, et que ces deux solutions vérifient la « condition d'intégrabilité » :

$$\frac{\partial}{\partial y} (\partial_x u) = \frac{\partial}{\partial x} (\partial_y u) \quad (5)$$

6. Terminez la résolution en calculant $u(x, y)$. Que peut-on dire du nombre de solutions de la stéréophotométrie lorsque $n = 2$?

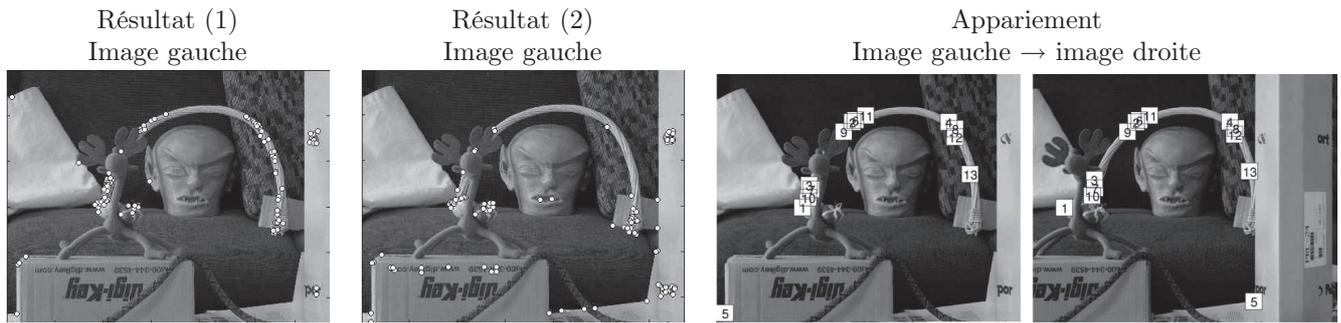


FIG. 1 – Détection de points d'intérêt (pastilles blanches) et appariement (rectangles numérotés).

2 Exercice 2 : Détection et appariement de points d'intérêt (6 points)

Pour détecter des points d'intérêt, généralement, les opérateurs proposés s'appuient sur une détection des contours (cf. TP2). Nous souhaitons travailler avec une paire d'images présentée dans la figure 1.

1. Quels types de contours peuvent être détectés dans cette image (appuyez vos propos en utilisant des exemples)? En utilisant le calcul de la dérivée première de cette image, comment peut-on détecter ces types de contours? (La réponse à ces 2 questions doit être succincte, elle ne doit pas dépasser 4/5 lignes).
2. Sachant qu'une fonction gaussienne à deux variables, d'écart type σ , en supposant une moyenne nulle en x et en y , est donnée par :

$$g(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}, \quad (6)$$

exprimez les dérivées partielles secondes, notées $g_{xx}(x, y)$, $g_{yy}(x, y)$ et $g_{xy}(x, y)$.

3. Donnez les masques de taille 3×3 correspondants. Seule l'expression du masque en fonction de g_{xx} , g_{yy} et g_{xy} est souhaitée, en faisant apparaître les simplifications possibles.
4. En reprenant le même schéma que la fonction `gaussmask` (cf. TP2), complétez ce code (4 instructions maximum) :

```

fonction [Lissage,DerivI,DerivJ,DerivII,DerivJJ,DerivIJ]=gaussmask2(TailleFenetre)
    n=floor(TailleFenetre/2); Sigma=TailleFenetre/4; [J,I]=meshgrid(-n:n);
    Lissage=exp(-(I.^2+J.^2)/(2*Sigma^2)); Lissage=Lissage/sum(sum(Lissage));
    DerivI=-I.*exp(-(I.^2+J.^2)/(2*Sigma^2));
    DerivI=DerivI/sum(sum(DerivI(1:n,:)));
    DerivJ=DerivI';

```

5. Le détecteur de Kitchen et Rosenfeld s'appuie sur l'expression suivante :

$$KR(x, y) = \frac{I_{xx}(x, y)I_y^2(x, y) - 2I_{xy}(x, y)I_x(x, y)I_y(x, y) + I_{yy}(x, y)I_x^2(x, y)}{I_x^2(x, y) + I_y^2(x, y)}. \quad (7)$$

fonction `[XY,Res]=harris(Im,TailleFenetre,NbPoints,k)` (cf. TP2) suit les étapes suivantes :

- (a) Calcul des filtres utilisés pour le calcul de la réponse du détecteur;
- (b) Convolution de l'image originale pour obtenir les images des dérivées;
- (c) Calcul de la réponse du détecteur;
- (d) Suppression des non-maxima locaux;
- (e) Sélection des `NbPoints` en fonction des "`NbPoints`" réponses les plus fortes.

Que faut-il modifier dans le profil de la fonction et les étapes pour écrire la fonction `rosenfeldKitchen` permettant la détection de points d'intérêt suivant Kitchen et Rosenfeld?

6. Dans la figure 1 :
 - (a) À quels détecteurs correspondent les résultats (1) et (2)?
 - (b) Pourquoi le nombre d'appariements est moins important que le nombre de points détectés au départ?
 - (c) Quelle contrainte *a priori* peut on ajouter pour éviter l'erreur d'appariement sur le point 5?

3 Exercice 3 : Recalage affine de deux images (3 points)

Comme le suggère l'exercice précédent, des appariements entre points d'intérêt peuvent être imprécis, peu fiables. Ici on suppose que l'on dispose de deux images à recaler et de n couples de points d'intérêt appariés $(u_i, v_i)_{i=1..n}$. Chaque appariement est doté d'une mesure de confiance $w_i \in [0, 1]$. Plus la mesure de confiance est forte et proche de 1, plus l'appariement est fiable et peut donc contribuer significativement au recalage. Grâce à ces appariements, on veut opérer un recalage affine de nos deux images. On vous demande de modéliser le problème d'estimation des 6 paramètres de la transformation affine (il faut donc $n \geq 3$) avec prise en compte des mesures de confiance qui pondèrent l'influence des n appariements. Le recalage affine utilisant des points d'intérêt que nous avons discuté en cours est linéaire. Est-ce que votre modèle reste linéaire ? Avez-vous toujours une solution analytique au problème ?

4 Exercice 4 : Synthèse d'images (5 points)

Question 1 (1 point)

Justifiez pourquoi l'algèbre linéaire de \mathbb{R}^3 n'est pas suffisante pour modéliser les transformations nécessaires en synthèse d'images.

Question 2 (2 points)

Pour l'éclairage, vous avez implémenté une méthode appelée *Gouraud* (de son inventeur, un informaticien français né en 1944) et vous avez vu en cours la méthode de *Phong* (informaticien vietnamien, 1942-1975). Expliquez et comparez ces deux méthodes du point de vue algorithmique, mais aussi du point de vue des résultats obtenus. Comparez aussi leur complexité en temps et en place.

Question 3 (2 points)

Le renderer implémenté en TP propose une gestion naïve des parties non visibles ou cachées. Dans une scène complexe ayant énormément de primitives graphiques (polygones), on souhaite limiter le nombre de primitives à traiter. Proposez et expliquez deux ou trois stratégies pour pouvoir limiter les calculs inutiles, en particulier pour les primitives hors du champ de vue, ou occultées.