

**Questions/réponses**  
**Partie Traitement d'Images**  
**Réponses à la classe inversée**  
**portant sur le support de cours distribué**  
**cours du 13 février 2026**

1<sup>er</sup> mai 2026

## 1 Remarques générales

- **Si je n'ai pas répondu à votre question ou si la réponse n'est pas satisfaisante, vous avez le droit de la poser à nouveau pendant le cours ou même après.**
- Si certaines questions n'apparaissent pas c'est simplement parce qu'elles sont redondantes avec d'autres.
- L'ordre d'apparition des questions/réponses correspond au déroulé du cours.

## 2 Questions/Réponses

- (1) *Pages 14–15 : En théorie une image est définie comme une fonction continue, mais en pratique on travaille toujours avec une image discrète après échantillonnage. Est-ce que certaines opérations (comme la convolution ou les dérivées) sont des approximations de leur version continue, ou bien est-ce qu'on raisonne uniquement dans le cadre discret une fois l'image acquise ?*

**Réponse :** Dans la littérature, on formalise très souvent en continu. Mais, en pratique, on travaille en discret. Donc, la représentation informatique est toujours discrète et les outils toujours adaptés au discret.

- (2) *Pages 15, 21-26, 28-35 : Les filtres passe-bas permettent de réduire le bruit mais atténuent aussi les contours, tandis que les filtres passe-haut renforcent les contours mais amplifient le bruit. Existe-t-il un compromis théorique ou pratique qui permette de préserver l'information utile tout en limitant les effets indésirables, ou est-ce toujours dépendant de la scène et de l'application ?*

**Réponse :** Oui, en général, c'est ce qu'on appelle des filtres passe-bande. Si vous souhaitez aller plus loin sur ces aspects, n'hésitez pas à consulter cette ressource :

<https://hal.science/hal-00512280v2/file/CoursFiltrage.pdf>

- (3) Page 20 : comment obtient-on la formule 1.13 ? Car je suis parti du fait que l'on cherchait le moment ou  $P1p1(T)=P2p2(T)$  pour que l'erreur soit minimale et nous trouvons un moins dans la formule au dénominateur de la seconde fraction la ou il y a un plus dans le polycopié.

**Réponse :** Comme vous, j'ai repris le calcul, et vous avez raison, il y avait bien une erreur de signe sur l'équation (1.13) du polycopié, je l'ai corrigé sur la version en ligne. Voici le calcul que j'ai fait :

$$P_2 E_1(T) = P_1 E_2(T) \quad (1)$$

$$\frac{E_2(T)}{E_1(T)} = \frac{P_2}{P_1} \quad (2)$$

$$\frac{\exp(-\frac{(T-\mu_2)^2}{2\sigma^2})}{\exp(-\frac{(T-\mu_1)^2}{2\sigma^2})} = \frac{P_2}{P_1} \quad (3)$$

$$\exp\left(\frac{(T-\mu_1)^2 - (T-\mu_2)^2}{2\sigma^2}\right) = \frac{P_2}{P_1} \quad (4)$$

$$\frac{(T-\mu_1)^2 - (T-\mu_2)^2}{2\sigma^2} = \ln\left(\frac{P_2}{P_1}\right) \quad (5)$$

$$(T-\mu_1)^2 - (T-\mu_2)^2 = 2\sigma^2 \ln\left(\frac{P_2}{P_1}\right) \quad (6)$$

$$T^2 - 2T\mu_1 + \mu_1^2 - T^2 + 2T\mu_2 - \mu_2^2 = 2\sigma^2 \ln\left(\frac{P_2}{P_1}\right) \quad (7)$$

$$2T(\mu_2 - \mu_1) - (\mu_2^2 - \mu_1^2) = 2\sigma^2 \ln\left(\frac{P_2}{P_1}\right) \quad (8)$$

$$2T(\mu_2 - \mu_1) = (\mu_2^2 - \mu_1^2) + 2\sigma^2 \ln\left(\frac{P_2}{P_1}\right) \quad (9)$$

$$T = \frac{(\mu_2^2 - \mu_1^2)}{2(\mu_2 - \mu_1)} + \frac{\sigma^2}{\mu_2 - \mu_1} \ln\left(\frac{P_2}{P_1}\right) \quad (10)$$

$$T = \frac{\mu_2 + \mu_1}{2} + \frac{\sigma^2}{\mu_2 - \mu_1} \ln\left(\frac{P_2}{P_1}\right) \quad (11)$$

$$(12)$$

- (4) Page 40 : Dans le cours, on suppose que l'on possède une segmentation de référence, ce qui nous permet d'évaluer la précision, la sensibilité et le coefficient de similarité. Dans un contexte où aucune vérité terrain n'est disponible, existe-t-il des stratégies d'évaluation qui permettent de juger objectivement la qualité d'une segmentation d'images ? En pratique, comment un ingénieur peut-il décider si une segmentation est "assez bonne" pour être utilisée dans une chaîne de traitement ?

**Réponse :** Sans évaluation quantitative, cela ne peut être que du qualitatif. Ce qui est questionnable. L'évaluation sera visuelle, soit par l'ingénieur, soit par des utilisateurs. Il est éventuellement possible d'évaluer si cette segmentation fait partie d'une chaîne de traitement et que les éléments suivants sont évaluables.

- (5) *Pages 40-41 : Le cours mentionne que le Deep Learning a révolutionné la vision par ordinateur. Dans ce contexte, les méthodes de segmentation classiques du chapitre 3 (comme le Watershed ou les contours actifs) sont-elles aujourd'hui rendues obsolètes, ou sont-elles utilisées en pré-traitement (ou post-traitement) pour hybrider et affiner les résultats des réseaux de neurones ?*

**Réponse :** La réponse est oui et non. Ces méthodes sont devenues obsolètes dans les cas où les contours sont difficiles à extraire. Si on possède le jeu de données annotées pertinent pour la tâche visée, la puissance de calcul pour l'apprentissage du modèle ou le transfert learning, ainsi que la ressource nécessaire pour l'inférence, alors, utiliser de l'apprentissage profond est plus adapté. Cependant, dans des cas où les contours sont suffisamment contrastés, ce type d'approche classique est moins coûteux (en mémoire, en exécution, en données) et reste plus adapté.

- (6) *Page 42 : Pour la convergence du contour actif/snake : comme on utilise le gradient pour minimiser l'énergie, est ce que le contour n'est pas trop sensible aux bruits ou aux changements de texture/couleurs en restant bloqué dans un minimum local avant d'atteindre le vrai contour de l'objet ? Est ce qu'il ne faudrait pas ajouter une sorte de seuil pour éviter les minimums locaux ? Le risque augmente aussi si l'initialisation est loin de l'objet donc comment bien choisir cette initialisation ?*

**Réponse :** Tout va dépendre de la méthode d'optimisation utilisée pour faire converger le contour. Il faut choisir une approche la plus robuste possible aux minima locaux. Dans la littérature, il y a énormément de publications sur ce sujet et c'est la raison pour laquelle on peut trouver des approches basées sur les coupures de graphes ou encore les courbes de niveaux. C'est aussi l'origine des approches basées géodésiques que j'évoque rapidement dans le cours. Pour la robustesse à l'initialisation, il y a également de nombreuses possibilités, évoquées dans le cours : initialisation manuelle par un.e expert.e ou un.e utilisateur.ice, utilisation des gradients, utilisation d'une autre méthode pour initialiser un contour. Ce qu'il faut retenir c'est que, souvent, les contours actifs sont utilisés dans des images où l'on est déjà très centré et focalisé sur l'objet à détourner.

- (7) *Pages 43 : Dans les approches par contours avec énergie à minimiser, quelle est la différence entre flexibilité et élasticité et que représentent concrètement les forces externes ?*

**Réponse :** La flexibilité va permettre d'autoriser une courbe au contour irrégulier alors que l'élasticité va permettre d'autoriser la courbe / le contour à s'aggrandir ou à rétrécir.

- (8) *Page 46 : dans les approches par croissance de régions, quelle est la différence entre propagation séquentielle et parallèle et en quoi la propagation parallèle permet de ne pas privilégier une région d'une autre ?*

**Réponse :** Une propagation séquentielle signifie que l'on va prendre chaque germes les uns après les autres. Lorsqu'on travaille sur un germe, on va le propager tant que c'est possible. Lorsque ce ne sera plus possible, on étudie le germe suivant et ainsi de suite. Une propagation en parallèle, pour chaque itération, tous les germes sont propagés dans leur voisinage 1-connexe. Puis, à l'itération suivante ce sont tous les nouveaux pixels étiquetés qui sont propagés dans leur 1-voisinage, et ainsi de suite. Dans la deuxième solution, chaque germe est propagé en même temps, on ne favorise donc pas un germe par rapport aux autres.

- (9) Page 48, Section 3.6.1 "k-moyennes" et Figure 3.10. Si on décide d'appliquer les k-moyennes non seulement sur la couleur, mais également sur la distance des pixels, faut-il calculer la distance euclidienne sur un vecteur 5 dimensions  $(r,g,b,x,y)$  ? si oui comment donner plus ou moins de poids à la couleur ou la distance entre les points ?

**Réponse :** Oui, c'est exactement cela et si on veut pondérer, on réalise exactement l'algorithme SLIC étudié en TP.

- (10) Pages 50 et 51 : Dans la méthode du Mean Shift, lorsqu'on parle de  $n$  observations pour estimer la fonction de densité, que représentent exactement ces  $x_i$  ? Correspondent-ils uniquement à la position spatiale des pixels dans l'image (par exemple en dimension 2 pour une image 2D), ou bien incluent-ils également l'information de couleur du pixel (par exemple dimension 3 pour une image en niveaux de gris et dimension 5 pour une image couleur) ?

**Réponse :** Il s'agit bien d'une dimension 5, en théorie. Mais, comme expliqué dans le cours un peu plus loin, page 51, en pratique, on sépare le traitement entre la position et la couleur.

- (11) Page 51 : Dans l'algorithme, on observe une contrainte basée sur un seuil colorimétrique  $h_c$  plutôt que sur un seuil spatial. Cela signifie-t-il que la classification des objets de l'image prend en compte la couleur même si les pixels sont spatialement éloignés ? Si oui, comment la contrainte spatiale est-elle prise en compte dans l'algorithme ?

**Réponse :** Lorsque j'ai écrit, page 51, en dessous de l'algorithme : « à l'intérieur d'une fenêtre de taille  $(2h_s + 1) \times (2h_s + 1)$  centrée en  $\mathbf{x}$ , nous considérons les pixels  $\mathbf{x}_i$  ayant une couleur proche de celle de  $\mathbf{x}$ , c'est-à-dire tels que  $\|I(\mathbf{x}_i) - I(\mathbf{x})\| \leq h_c$ . », cela signifie que pour qu'un pixel soit considéré dans le calcul de la moyenne, il faut qu'il soit conjointement, dans la zone spatiale définie autour du pixel et proche colorimétriquement. Dans l'algorithme de la page 51, j'indique que l'on utilise la formule (3.7) et dans cette formule, nous considérons bien une zone (spatiale et colorimétrique) autour du pixel étudié ( $S_h$ ).

- (12) Page 56-58 : Comment choisir entre une méthode qui génère des superpixels très réguliers (SLIC) et une autre qui est plus réactive aux contours mais moins régulière (Felzenszwalb) pour la segmentation sémantique ou la détection d'objets dans des scénarios complexes ? À part la vitesse de calcul, quels critères utiliser pour décider, par exemple si la régularité aide vraiment pour la suite ou si c'est la précision des contours qui compte le plus ?

**Réponse :** Avant de répondre à la question de la régularité, il faut également garder en mémoire que le tableau 4.1 doit vous aider à faire votre choix. Ensuite, pour la régularité, cela peut dépendre également de ce que vous souhaitez faire avec ces superpixels. Notamment, si vous souhaitez avoir un raisonnement en considérant un voisinage régulier ou des métriques (qui peuvent être biaisées par le manque de régularité), il vaut mieux favoriser la régularité. Notamment, et c'est évoqué dans le cours, tous les travaux qui souhaitent appliquer des champs de Markov, auront plutôt besoin d'une structure régulière. En revanche, dans le cadre de la segmentation, comme on l'a fait en TP, la régularité n'était pas du tout nécessaire.

- (13) *Page 56, équation (4.1) : en quoi cette équation minimise-t-elle la similiarité inter-régions puisque qu'on minimise une somme sur les écarts-types au sein d'unions de superpixels et en quoi maximise-t-elle la similarité des pixels intra-région puisque les termes de cette somme sont inversements proportionnels aux écarts-types intra-superpixels ?*

**Réponse :** Merci pour cette question qui m'a permis d'identifier une erreur dans le texte. Ce qu'il faut comprendre, c'est bien que l'on souhaite maximiser la variance inter-régions (les couleurs entre deux régions distinctes doivent être très différentes) et minimiser la variance intra-région (les couleurs d'une même région doivent être très proches). Dans l'équation (4.1), on calcule la moyenne sur l'ensemble des superpixels d'un critère intra/inter-région. Ainsi, dans la somme permettant de calculer cette moyenne, soit, pour chaque superpixel, le numérateur calcule la moyenne des variances inter-pixels, pour chaque superpixel dans son voisinage. On souhaite donc que cette moyenne soit la plus élevée possible. En revanche, le dénominateur calcule la variance du superpixel et on souhaite que celle-ci soit la plus basse possible. En conclusion, plus le rapport entre les deux est élevé, plus le critère est satisfait. L'erreur était donc d'avoir écrit qu'il faut minimiser ce critère alors que c'est l'inverse, ce critère doit être maximisé.

- (14) *Page 58, équation (4.4) : Dans l'algorithme SLIC, comment le réglage des paramètres  $m$  et de  $V$  impacte le résultat final ? Puisqu'ils sont liés dans la formule, doit-on forcément ajuster les deux en même temps pour que les superpixels collent bien aux contours ? / on s'interroge sur le rôle exact du paramètre  $m$  qui pondère la distance spatiale. Concrètement, si on augmente fortement cette valeur, quel comportement privilégie-t-on ? Est-ce que cela va forcer les superpixels à garder une forme très géométrique et régulière, quitte à moins bien "coller" aux contours réels des objets de l'image ?*

**Réponse :** L'objectif de pondérer en utilisant  $\frac{m}{V}$  est que le poids s'ajuste de lui même en fonction du nombre de superpixels, donc, fixer  $m$ , doit suffire. Effectivement, le rôle de  $m$  est uniquement de travailler / d'ajuster la compacité des superpixels. Plus les superpixels seront compacts et plus ils auront de mal à suivre correctement les contours colorimétriques. C'est donc à ajuster en fonction des images traitées et vous avez pu largement observé le choix  $m$  sur la forme des superpixels.

- (15) *Page 65 : Le cours présente le modèle HSI (page 65) comme un système permettant de séparer l'information de chrominance de la luminance. Notre interrogation est la suivante : quand on est dans des zones de gris ou des zones très sombres, les valeurs  $R$ ,  $G$ ,  $B$  sont presque identiques, ce qui fait que le dénominateur tend vers zéro. Mathématiquement, la teinte devient alors indéfinie ou hyper sensible au moindre petit bruit numérique (on peut passer d'un angle de  $0^\circ$  à  $180^\circ$  à cause de faible variation). Étant donné que la teinte devient mathématiquement instable dans les zones peu saturées, est-ce que ces pixels ne risquent pas de fausser le calcul des centroïdes lors d'une segmentation par  $k$ -means (page 48), et ne serait-il pas alors nécessaire d'instaurer un seuil de saturation minimal en dessous duquel on ignorerait la teinte pour éviter que l'algorithme ne diverge ?*

**Réponse :** Effectivement, si on a des valeurs  $R$ ,  $G$  et  $B$  quasi identiques, cela signifie, par construction, qu'on a une nuance de gris donc, la teinte n'est pas une information exploitable. Dans le cas d'une teinte de gris, la formulation de la transformation le considère et attribue systématiquement une valeur par défaut, la même. Une solution peut donc être d'ajouter un epsilon sur cette égalité. Ma proposition est équivalente à ce que vous proposez, mais, on évite d'ignorer des teintes et d'avoir un traitement différencié.

- (16) *Chapitre 5 : Représentation des couleurs : Pourquoi le système RGB n'autorise-t-il que des valeurs positives, alors que d'autres systèmes de couleurs comme les systèmes antagonistes acceptent des valeurs négatives, alors qu'il ne s'agit mathématiquement que d'un changement de base via une matrice de passage entre des systèmes pourtant équivalents ?*

**Réponse :** Je n'ai pas de réponse à cette question s'appuyant sur des éléments de la littérature. Cependant, mon avis, c'est qu'informatiquement, c'est beaucoup plus simple à coder et gérer.

- (17) *Chapitre 6 : Étude de la texture : Comment peut-on choisir de manière automatique et robuste le seuil  $S$  pour détecter les zones texturées dans une image, plutôt que de le fixer arbitrairement ?*

**Réponse :** Vous pouvez faire appel aux techniques proposées dans le cours pour binariser une image : méthode d'Otsu, seuillage par mélange de gaussienne ou par hystérésis.

- (18) *Question d'ouverture : L'égalisation d'histogramme améliore le contraste global, mais elle modifie la distribution relative des niveaux de gris entre les objets. Est-ce qu'on peut dire que cette transformation améliore toujours l'analyse de la scène, ou peut-elle au contraire nuire à certaines tâches comme la segmentation ?*

**Réponse :** Effectivement, aucune des opérations proposées est 100% positive ou négative. Parfois, réaliser une égalisation ou un étalement de la dynamique peut être négatif pour la segmentation. Cependant, il faut quand même retenir que, globalement, l'étalement va plutôt être utile pour améliorer le confort visuel (et donc parfois être négatif pour la segmentation) alors qu'à l'inverse l'égalisation va surtout être utile pour la segmentation (et plutôt négative pour le confort de visualisation).

- (19) *Question d'ouverture : serait-il possible de décomposer les images en plusieurs sous images ? C'est-à-dire plusieurs sous problèmes, afin de minimiser les inconvénients de nos méthodes, et enfin recomposer nos résultats pour récupérer l'image d'origine ?*

**Réponse :** Oui, bien sûr, on peut décomposer l'image en imagerie pour réduire la complexité du problème. Et, on peut donc, aussi utiliser des superpixels. C'était le message du cours.

- (20) *Question d'ouverture : Dans le cours, il n'est jamais question d'utiliser d'autres parties du spectre lumineux, serait-il dans les cas où d'autres images en infra-rouge ou ultra-violet par exemple sont accessibles possible de les utiliser pour mieux analyser l'image initiale ? Par exemple, est-ce qu'utiliser un domaine non visible pourrait permettre de faciliter le traitement d'une image où le fond est complexe ?*

**Réponse :** Oui, bien sûr, cela peut être intéressant mais il faut avoir accès aux capteurs qui permettent ce type d'acquisition.

- (21) *Question d'ouverture : Pourrait-on imaginer une approche hybride pour les contours actifs où l'on utiliserait d'abord une sur-segmentation rapide (type SLIC, Chapitre 4) pour générer un graphe de régions, dont la fusion fournirait une initialisation grossière mais correcte pour le Snake ? Cela permettrait-il de s'affranchir du placement manuel des points de contrôle tout en garantissant une meilleure convergence de l'énergie ?*

**Réponse :** Oui, c'est une très bonne idée. En fait, l'état de l'art regorge de méthodes hybrides qui combinent des méthodes comme vous le proposez. L'objectif du cours était de vous proposer des briques de base pour vous permettre d'imaginer et de comprendre le plus d'algorithmes possibles.