DIRO
IFT 3205

TRAVAIL PRATIQUE

Estimation du Flot Optique par Intercorrélation dans l'Espace de Fourier

Max Mignotte

DIRO, Département d'Informatique et de Recherche Opérationnelle.
http : //www.iro.umontreal.ca/∼mignotte/ift3205
e-mail : mignotte@iro.umontreal.ca

1 Introduction

L'objectif de ce TP est d'étudier les propriétés de la fonction d'intercorrélation (et sa FFT) et de voir son utilitée dans le cadre de l'analyse et du traitement du signal et l'estimation du float optique. La fonction d'intercorrélation qui est un outil important en traitement du signal car elle permet principalement de mesurer la ressemblance entre deux signaux f et g. En supposant que f et g sont des signaux réels unidimensionels dans le domaine spatial, la fonction d'intercorrelation C_{fg} est défini par :

$$C_{fg}(\tau) \triangleq \int f(x)\, g(x-\tau)\, dx$$

C_{fg} quantifie dans quelle mesure on peut écrire $f(x) = \alpha g(x-\tau) + \beta$.
Remarquer que la convolution est une intercorrelation dans lesquel un des signaux est renversée dans l'espace. Inversement, l'intercorrélation peut donc se noter à partir d'une convolution de la façon suivante

$$C_{fg}(\tau) \triangleq f(\tau) * g(-\tau) \quad \stackrel{\mathcal{F}}{\rightleftharpoons} \quad F(\nu)\, G^{\star}(\nu) \triangleq \mathcal{C}_{fg}(\nu)$$

avec F et G, les TF de f et g et en se souvenant que, avec la correspondance $\tau \stackrel{\mathcal{F}}{\rightleftharpoons} \nu$, on a $f(-\tau) \stackrel{\mathcal{F}}{\rightleftharpoons} F^{\star}(\nu)$. $\mathcal{C}_{fg}(\nu)$, la TF de la fonction d'intercorrélation, désigne la densité interspectrale d'énergie et caractérise la contribution énergétique des différentes contributions fréquentielles.

2 Recalage d'Images ou de Signaux

La fonction d'intercorrélation et la TF sont deux outils puissants utilisés en traitement du signal et des images qui permettent de faire, entre autres, le recalage c'est-à-dire la mise en correspondance de signaux ou d'images, ceci afin de pouvoir comparer ou combiner leurs informations respectives. Cette mise en correspondance se fait par la recherche d'une transformation géométrique (e.g., translation, facteur d'échelle, rotation, etc.) permettant de passer d'une image (ou un signal 1D) à une autre. Cette technique a de nombreuses applications allant de l'imagerie médicale afin par exemple de fusionner plusieurs modalités d'imagerie, au traitement de vidéos comme le suivi de mouvement, la compression, le recalage de profils sismiques, ou encore la création de mosaïques d'images (panoramas).

Considérons deux images $f(x,y)$ et $g(x,y)$ à recaler. Si on considère qu'il existe un décalage, due à une translation (de vecteur (x_0, y_0)) entre ces deux images, alors on peut écrire grâce à une propriété de la TF

$$g(x,y) = f(x-x_0, y-y_0) \quad \overset{\mathcal{F}}{\rightleftharpoons} \quad G(u,\nu) = F(u,\nu)\exp\big[-2\pi j(ux_0+\nu y_0)\big]$$

avec la correspondance $x, y \overset{\mathcal{F}}{\rightleftharpoons} u, \nu$. Ceci signifie que les images f et g ont le même module de Fourier, i.e., $|G(u,\nu)| = |F(u,\nu)|$ tandis que la différence de phase est directement reliée au déplacement spatial existant entre ces deux images. De plus, comme $F(u,\nu) = |F(u,\nu)|\,\exp[-j\phi_f(u,\nu)]$, on trouve

$$\frac{\overbrace{G(u,\nu)\,F^{\star}(u,\nu)}^{\mathcal{C}_{gf}(u,\nu)}}{|F(u,\nu)|^2} = \exp[-2\pi j(ux_0+\nu y_0)] \quad \Leftrightarrow \quad \mathcal{F}^{-1}\left[\frac{\mathcal{C}_{gf}(u,\nu)}{|F(u,\nu)|^2}\right] = \delta(x-x_0, y-y_0) \tag{1}$$

Le calcul de la TF inverse de la densité interspectrale entre ces deux fonctions permet de trouver le vecteur de translation (x_0, y_0) existant entre ces deux images. On peut ensuite utiliser cette information pour recaler g sur f ou pour d'autres applications. Dans la suite on va considérer que l'image UdM1.pgm est l'image f et que l'image UdM2.pgm est l'image g. On vous demande donc dans ce TP de

Fig. 1 – *De gauche à droite. Les deux images f (UdM1.pgm) et g (UdM2.pgm) considérées et image de l'interspectre de ces deux images ; on voit le $\delta(x-x_0, y-y_0)$ (point blanc) de l'Équation (1).*

Calculer et afficher à l'écran, l'estimation du vecteur de translation entre l'image f et l'image précédente en utilisant l'Équation (1) (cf. image 1.(c)).

3 Estimation du Flot Optique

En vision par ordinateur, l'estimation du flot optique consiste à estimer, à partir d'une séquence d'images[1], le mouvement apparent des objets composant une scène tridimensionnelle (ou plus formellement, à recouvrer la projection des différents mouvements tridimensionnels de la scène sur le plan image). Concrètement, pour une image $I(t)$ extraite d'une séquence d'image $\{I(t=0), I(t=1), \ldots, I(t), I(t+1), \ldots\}$. Le calcul du mouvement apparent consiste à associer à chaque pixel (x,y) de cette image, un vecteur (v_x, v_y) représentant la vitesse apparente du pixel estimé à cet instant t.

Ce mouvement apparent est estimé uniquement à partir des changements de la distribution spatiale des intensités lumineuses (ou niveaux de gris des pixels) existant entre deux images consécutives de la séquence. Ce mouvement est dit appelé "apparent" car il ne correspond pas nécessairement à la réalité. Dans le cas, où il existe un mouvement mais aussi aucune variation d'intensité existant dans les deux images (cas d'une sphère de couleur homogène, non texturée, tournant sur elle -même), les vitesses apparentes seront estimées comme

étant nulle ! C'est ce qui distingue le calcul du flot optique (mouvement apparent reflété par une variation d'intensité lumineuse) et le mouvement apparent réel. Dans le cas, où les objets de la scène possèdent des bords et quelques textures, le mouvement apparent réel peut raisonnablement être estimé par le flot optique, *i.e.*, sur la base des variations de la fonction de luminance.

Dans ce TP, le calcul du vecteur de translation pour chaque pixel de l'image (représentant le flot optique ou les vitesses apparentes de chaque pixel estimé à cet instant t) est calculé pour chaque imagette de taille 32×32 centré sur le pixel ([4]) par la méthode décrite dans la section 2. Dans ce cas, puisque l'imagette est petite et pour contrer les effet de bord créés par la FFT (qui va implicitement périodiser l'image) et donc créeer des pics de corrélation factices, on vous conseille de pondérer chaque imagette avec une fenêtre de Hanning : $h(i)\,h^T(j)$ où $0 \leq i \leq \text{length}$ et $0 \leq j \leq \text{width}$ et $h(n)$ définie en 1D par :

$$h(n) = 0.5\left\{1 - \cos\Big(\frac{2\,\pi\,n}{N-1}\Big)\right\}$$

De plus, à la fin, le flot optique est filtré avec un filtre moyenneur de taille 33×33. Tracer la solution du flot optique pour l'image RUBIKSEQ, pour l'image SALESMAN er pour l'image CARFRONT (cette solution du flot optique pourra être ensuite visualisée sur l'écran de l'ordinateur[5] (cf. Figs 2 et 3).

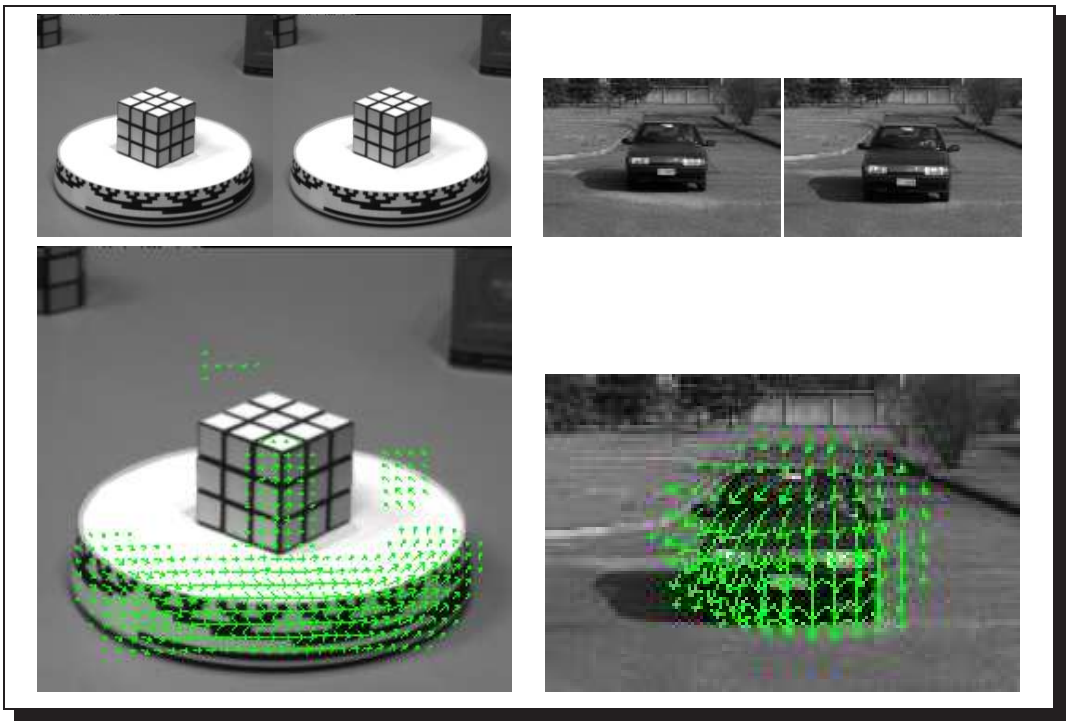

FIG. 2 – Estimation du flot optique obtenu sur l'image RUBIKSEQ et l'image CARFRONT.

[4] Pour gérer les bords de cet algorithme, on prendra un modèle d'image toroidale, c'est-à-dire repliée sur elle-même. Dans ce modèle d'image, ou l'image est de taille length×width, les pixels auront des coordonnées allant de 0 à length−1 en hauteur et allant de 0 à width−1 en largeur. Ainsi, le pixel de coordonnée $(-3,-1)$, qui sort de l'image, correspond (lorsque l'image est toroidale) au pixel de coordonné $(-3+\text{length}, -1+\text{width})$. De même, si length= 256 et width= 256, alors le pixel de coordonnée $(256, 266)$, qui sort de l'image, aura pour coordonnée $(256 - \text{length} = 0, 266 - \text{width} = 10)$.

[5] Utiliser le programme que je vous donne sur ma page web qui affichera à l'écran le flot optique avec des flêches.

[6]

FIG. 3 – Estimation du flot optique obtenu sur l'image SALESMAN.

Remise & Rapport

Vous devez rendre physiquement au démonstrateur le rapport et électroniquement le(s) programme(s) fait en C avant la date de remise spécifiée dans le fichier *barême* situé sur la page web du cours. Pour la remise électronique, utilisez le programme *remise* (*man remise* pour plus de détails) pour remettre votre code dans le répertoire TP<Numéro du Tp>. N'oubliez pas d'inscrire vos noms, courrier électronique en commentaire en haut du fichier .c remis. Les noms des programmes à remettre devront avoir le format suivant *TpIFT3205– <Numéro du Tp>-<Numéro de la question>.c*. Les programmes devront se compiler et s'executer sur Linux tel qu'indiqué dans le barême.

Le rapport (si celui-ci est demandé dans le fichier barême) doit être concis, clair et devra contenir (en plus des éléments et des discussions que vous jugerez importants), la description brève du problème et de la technique utilisée (lorsque cela est nécesaire ou demandé), les réponses aux questions posées, les images résultats et les commentaires que l'on vous demande et que vous jugerez utile.