

**FACULTE DES SCIENCES EXACTES ET SCIENCES DE LA NATURE ET DE LA VIE
DEPARTEMENT DE MATHÉMATIQUES ET INFORMATIQUE**

Filière : INFORMATIQUE

2^{ème} année Master-option vision artificielle-

Contrôle

Module : Détection et estimation de mouvement

Le 19 /01 / 2025

Exercice N° 1 (14 pts)

1°/Décrire brièvement le principe de la détection de mouvement (1,5pts).

La détection de mouvement dans les séquences d'images (vidéo) a comme objectif de trouver les points (pixels) qui changent de position au cours du temps dans la séquence, c'est à dire consiste d'associer à chaque pixel une étiquette binaire pour modéliser le changement (valeur 1) ou le non changement de position (valeur 0). Le résultat final de la détection de mouvement est donc une carte binaire qui contient des valeurs 1 et 0, ces dernières peuvent être utilisées pour identifier la présence ou l'absence de mouvement dont le but est de distinguer les objets mobiles et statiques dans la scène.

Le processus de détection de mouvement n'est pas toujours possible à cause de plusieurs obstacles, nous pouvons citer :

- La variation de la luminosité et les conditions d'éclairage qui ont des influences sur l'apparence des objets composant la scène.
- Bruit dans la séquence d'images.
- Variabilité de la forme et de la structure des objets de la scène.
- Les zones homogènes où nous ne pouvons pas détecter le mouvement bien que ce dernier existe.

2°/ Ecrire l'algorithme de détection de mouvement avec image de référence (2 pts).

La détection de mouvement avec image de référence consiste à mesurer la différence entre l'image de référence et l'image courante et enfin une opération de seuillage des résultats obtenus précédemment pour réduire le bruit.

$$\zeta(x, y) = |I_t(x, y) - I_R(x, y)|$$

avec $I_R = \sum_{i=t-N}^{t-1} w_i I_t$

avec $w_i = \frac{1}{N}$ où :

I_R et L 'image de référence

I_t : L 'image a l'instant t

N : Le nombre de frame dans ma séquence

w_i Représente le facteur de pondération

3°/Décrire le principe de l'estimation de mouvement (contraintes et hypothèses) (3pts) ?

L'estimation du mouvement consiste à mesurer la projection 2D dans le plan de l'image d'un mouvement réel 3D, dû à la fois au mouvement des objets dans la scène et aux déplacements de la caméra.

L'estimation de mouvement est un problème mal-posé car il n'a pas toujours de solution dans le cas d'une occultation, et s'il en a une, elle n'est pas toujours unique à cause du problème d'ouverture. L'occultation peut être définie comme le recouvrement ou non d'une zone pendant le déplacement causé par un autre objet en mouvement. Le problème de l'ouverture indique que seule la composante normale au déplacement est mesurable, c'est-à-dire seule celle orthogonale au contour local de l'image, orientée dans la direction du gradient spatial de l'intensité, au point considéré

Hypothèses et Contrainte de l'estimation de mouvement

Les hypothèses pour l'estimation du flux optique sont les suivantes :

- La conservation des données qui suppose que l'intensité des pixels reste constante au cours du temps. La contrainte de conservation des données peut être exprimée de la façon suivante :

$$I(x, y, t) = I(x + \delta x, y + \delta y, t + \delta t)$$

Avec $I(x, y, t)$ l'intensité du pixel à la position dans l'image au temps . $[\delta x, \delta y]$ correspond aux déplacements horizontal et vertical dans l'image et un écart de temps faible. Soit $V = [v_x(x, y), v_y(x, y)]^T$ le vecteur vitesse en un point de l'image au temps $t + \delta t$

- La cohérence spatiale qui repose sur la cohérence du mouvement entre deux pixels d'un même voisinage. La contrainte de cohérence spatiale suppose que le flux est constant dans un voisinage donné autour du pixel étudié, appartenant par exemple au même objet.

- Certaines techniques plus évoluées introduisent aussi l'hypothèse de la continuité temporelle, aussi appelée persistance temporelle. L'introduction de nouvelles techniques tenant compte de la continuité du mouvement au cours du temps suppose que les changements d'intensité soient temporellement graduels pour un pixel situé dans un voisinage donné. Autrement dit le support temporel pour l'étude du flux optique est donc supérieur à 2

- Les objets sont des corps rigides c'est-à-dire que la déformation des objets est négligée au moins pour un minimum d'images voisines pour garantir le même mouvement pour tous les pixels du même objet

- Le mouvement des objets est translationnel pour un minimum d'images voisines.

4°/ Décrire les différentes solutions proposées pour la résolution de l'équation du flot optique par les méthodes différentielles (les formules incluses) (2pts).

Les méthodes différentielles globales consiste à minimiser sur le domaine entier de l'image une fonctionnelle prenant en compte l'équation du flot optique ainsi qu'un terme de lissage, c'est à dire en ajoutant une contrainte de régularisation portant sur le gradient, le Laplacien (ou ordre supérieur) du champ de vitesse. La plus connue de ces méthodes est certainement celle proposée par Horn & Schunck.

Horn & Schunck ont proposé d'ajouter, à la contrainte de gradients, un terme de lissage spatial pour contraindre le champ de vitesses estimé (v_x, v_y) . Ce dernier est alors obtenu en minimisant l'Eq suivante :

$$\int_D (\nabla_{xyt} I \cdot \hat{v})^2 + \lambda^2 (\|\nabla_{xy} v_x\|^2 + \|\nabla_{xy} v_y\|^2) dx dy$$

où D est le domaine de définition spatiotemporel, $\| \cdot \|$ le module et λ un terme de pondération de la contrainte de lissage. L'ajout de cette contrainte de lissage revient à considérer que les pixels voisins ont des vitesses proches et que le champ de vitesses varie lentement.

Les méthodes locales consistent à prendre en compte des hypothèses supplémentaires sur un domaine de taille réduite pour particulariser le flot optique. On minimise alors un critère sur un petit domaine, et on obtient ainsi le flot optique de ce petit domaine. La méthode locale la plus célèbre est celle de Lucas & Kanade. Lucas et Kanade proposent d'estimer le flux optique par une résolution par les moindres carrés du système pondéré. On cherche à minimiser :

$$\sum_{\vec{x}=(x,y) \in \Omega} W^2(\vec{x}, t) [\nabla I(\vec{x}, t) \cdot V(\vec{x}, t) + I_t(\vec{x}, t)]^2$$

Avec Ω un voisinage spatial carré de taille $(n \times n)$ et $W(\vec{x}, t)$ contient les coefficients d'une fonction gaussienne 2D qui reflète la mesure de confiance que l'on peut avoir dans la vitesse $V(\vec{x}, t)$ estimée.

5°/ Expliquer le principe de la méthode fréquentielle « corrélation de phase » et donnez son algorithme (2pts)?

Cette technique estime la translation 2D entre chaque paire d'images en prenant, pour chacune d'elles, sa transformée de Fourier. Son principe repose sur la recherche de la différence de phases pour chaque fréquence et le calcul de nouveau de la transformée de Fourier inverse.

Si $I_1 = I(p, t)$ et $I_2 = I(p, t + 1)$ sont les deux images de la scène se recouvrant partiellement et G_1 et G_2 leurs transformées de Fourier respectives, alors la différence de phase entre ces deux images dans le domaine de Fourier est égale à leur spectre de puissance croisé normalisé (SPCN) exprimé comme suit :

$$e^{j(\theta_1 - \theta_2)} = \frac{|G_1| \cdot |G_2| e^{j(\theta_1 - \theta_2)}}{|G_1| \cdot |G_2|} = \frac{|G_1| e^{j\theta_1} \cdot |G_2| e^{-j\theta_2}}{|G_1| \cdot |G_2|} = \frac{G_1 G_2^*}{\|G_1 G_2^*\|}$$

D'autre part, la transformation entre les deux images est une simple translation exprimée par le vecteur $V(u, v)$, soit :

$$I_2(p) = I_1(p + V) = I_1(p) \delta(p - V)$$

δ est la fonction de Dirac .

La transformée de Fourier de cette équation est donnée par :

$$G_2(f) = G_1(f) e^{j2\pi fV}$$

ainsi, il est clair que la différence de phase $e^{j(\theta_1 - \theta_2)}$ entre les deux images n'est autre que $e^{j2\pi fV}$ pour chaque fréquence f . On trouve que la transformée de Fourier inverse de cette différence de phase (SPCN) produit la fonction de Dirac du vecteur de translation $\delta(p - V) = \delta(x - u, y - v)$:

$$d(p) = \delta(p - V) = F^{-1}(e^{j2\pi fV}) = F^{-1}(e^{j(\theta_1 - \theta_2)})$$

Il est évident que pour calculer le déplacement 2D (translations suivant x et y), il suffit de chercher le pic de cette représentation spatiale (La fonction $d(p)$) d présente en général un pic très net en $V = (u, v)$).

L'algorithme de la méthode de corrélation de phase est le suivant :

1. Calculer G_1 et G_2 les TF de I_1 et I_2 .
2. Calculer χ le SPCN de G_1 et G_2

Le spectre de puissance croisé normalisé (SPCN) exprimé par :

$$e^{j(\theta_1 - \theta_2)} = \frac{|G_1| \cdot |G_2| e^{j(\theta_1 - \theta_2)}}{|G_1| \cdot |G_2|}$$

3. Calculer d la TF inverse de χ .
4. Rechercher V , le maximum de d tel que :

$$d(p) = \delta(p - V) = F^{-1}(e^{j2\pi fV}) = F^{-1}(e^{j(\theta_1 - \theta_2)})$$

6°/ Citer les hypothèses pour que le flot optique soit exactement équivalent au mouvement image(1,5pts).

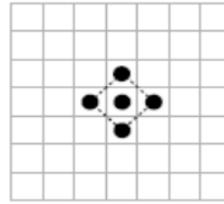
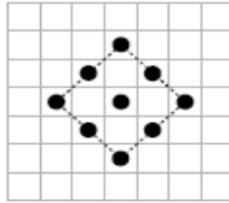
Le flot optique est exactement équivalent au mouvement image si les hypothèses suivantes sont respectées :

- éclairage uniforme
- surface à réflexivité Lambertienne
- mouvement de translation pure, parallèle au plan de l'image

7°/ Expliquer le principe de la méthode de mise en correspondance de blocs « Diamond search » (2pts)

Le voisinage dans l'algorithme de DS prend le forme d'un diamant et non pas un carré comme les algorithmes décrits précédemment et que le nombre des étapes est illimite. L'algorithme DS utilise deux modèles de recherche:

- LDSP : Large Diamant Search Pattern (voisinage d'ordre 1) avec neuf points de recherche.
- SDSP : Small Diamant Search Pattern (voisinage d'ordre 2) avec cinq points de recherche.



À la première étape de cette stratégie de recherche, le modèle LDSP est sélectionné, si la meilleure ressemblance n'est pas située au centre du diamant, le point central du diamant est alors remplacé par le nouvel emplacement trouvé. La recherche se poursuit en utilisant le modèle de recherche LDSP jusqu'à ce que l'emplacement qui minimise le critère de ressemblance soit au centre du diamant. Ensuite, on passe au modèle SDSP pour effectuer un processus de raffinement. Si le point qui minimise le critère de ressemblance est situé aux quatre points de recherche à l'entourage du point central du petit diamant, alors le point minimum sera le nouveau point central. Un sous-ensemble doit être vérifié à toute nouvelle étape, car le nouveau modèle est toujours partiellement superposé à l'ancien et l'algorithme s'arrête lorsque le meilleur point est situé au centre du SDSP.

Exercice N°2 (6 pts)

1°/ Ecrire l'algorithme d'estimation de mouvement dans une séquence de 2 images en utilisant la méthode mise en correspondance de blocs « Full search » (4pts).

```
function [u,v] = FS(curimg, refimg, mbSize, N)
[row col] = size(refimg);
vectors = zeros(2,row*col/mbSize^2);
costs = ones(2*N + 1, 2*N + 1) * 65537;
mbCount = 1;
for i = 1 : mbSize : row-mbSize+1
    for j = 1 : mbSize : col-mbSize+1
        for m = -N : N
            for n = -N : N
                refBlkVer = i + m;
                refBlkHor = j + n;
                if ( refBlkVer < 1 || refBlkVer+mbSize-1 > row ...
                    || refBlkHor < 1 || refBlkHor+mbSize-1 > col)
                    continue;
                end
                costs(m+N+1,n+N+1)= costFuncMAD(curimg(i:i+mbSize-1,j:j+mbSize-1),
                ...
                    refimg(refBlkVer:refBlkVer+mbSize-1,
refBlkHor:refBlkHor+mbSize-1), mbSize);

            end
        end
        [dx, dy, min] = minCost(costs);
        vectors(1,mbCount) = dy-N-1;
        vectors(2,mbCount) = dx-N-1;
        mbCount = mbCount + 1;
        costs = ones(2*N + 1, 2*N + 1) * 65537;
    end
end
```

```

motionVect = vectors;
a=1;b=1;c=1;
for i = 1 : mbSize : row-mbSize+1
    b=1;
    for j = 1 : mbSize : col-mbSize+1
        u(a,b)=motionVect(1,c);
        v(a,b)=motionVect(2,c);
        b=b+1;
        c=c+1;
    end
    a=a+1;
end
avec
function [dx, dy, min] = minCost(costs)
[ row, col ] = size(costs);
min = 65537;
for i = 1:row
    for j = 1:col
        if (costs(i,j) < min)
            min = costs(i,j);
            dx = j; dy = i;
        end
    end
end
end
et
function cost = costFuncMAD(currentBlk,refBlk, n)
err = 0;
for i = 1:n
    for j = 1:n
        err = err + abs((currentBlk(i,j) - refBlk(i,j)));
    end
end
end
cost = err / (n*n);
2°/ Ecrire l'algorithme d'estimation de mouvement « Full search Pyramidal » (2pts)

```

- Compute 'simple FS optical flow at highest level
- At level i
 - Take flow u_{i-1}, v_{i-1} from level $i-1$
 - bilinear interpolate it to create u_i^*, v_i^* matrices of twice resolution for level i
 - multiply u_i^*, v_i^* by 2
 - compute f_t, f_x, f_y using masks centered at (x,y) and $(x+u_i^*, y+v_i^*)$
 - Apply FS to get $u_i'(x, y), v_i'(x, y)$ (the correction in flow)
 - Add corrections $u_i', v_i', i.e. u_i = u_i^* + u_i', v_i = v_i^* + v_i'$.

Bon courage