

**TRAVAUX PRATIQUES DE VISION PAR ORDINATEUR,
VISION INDUSTRIELLE ET VISION ROBOTIQUE
CREATION D'OPERATEURS DE DETECTION DE MOUVEMENT :
APPLICATION A LA VIDEO SURVEILLANCE ET A LA POURSUITE DE CIBLE**

Patrick Bonnin
patrick.bonnin@gmail.com

I / Introduction :

Le but de ces Travaux Pratiques est la Réalisation d'Opérateurs de Détection de Mouvement, dans deux cadres applicatifs :

- La vidéo surveillance : caméra fixe, détection des objets mobiles grâce au mouvement à partir de deux techniques simples : l'une de différence d'images, l'autre d'ajustement par rapport à un gabarit.
- La Poursuite de Cible : pas de restriction relative au mouvement de la caméra, à partir de technique de corrélation. La position initiale de la cible est donnée. La détection automatique de la cible n'est pas l'objet de ce TP.

II / Schéma Algorithmique :

II. 1°) Principe :

Quel que soit le cas des trois précédemment énoncés, le schéma algorithmique est relativement similaire. Nous allons donc commencer par l'étudier et le programmer.

Le schéma algorithmique comporte deux étapes :

- L'acquisition de la référence : à partir d'une ou de plusieurs images (ie. le gabarit),
- Le Traitement de la Suite de la Séquence : Image par Image, le Traitement de l'Image Courante remettant à jour ou non la Référence.

II. 2°) Codage de la Séquence Vidéo :

La Séquence Vidéo est fournie sous forme d'images couleurs codées en RVB en format .ppm. Les images sont numérotées avec un incrément de une unité entre deux images successives. Pour pouvoir visualiser la séquence en dynamique avec le logiciel Irfan View, le numéro est codé sur 4 chiffres, soit 0000, 0001, 0002 etc... Le nom de l'image courante est donc par exemple « **NomGenerique0002.ppm** » pour l'image de numéro 2. Ce codage permet le traitement de séquences pouvant aller jusqu'à 10 000 images.

Il est demandé de pouvoir commencer le traitement de la séquence à une image de numéro donné, ce numéro sera passé en paramètre.

II. 3°) Cahier des Charges :

Ecrire le squelette algorithmique suivant :

```

* Constitution du nom, ouverture du fichier de la première image, Lecture
de l'Entête,
* Création des Structures de Données : Image de Référence, Image Courante,
Image Résultat,
* Lecture de la première image : constitution de la Référence (ou du
gabarit, à partir de plusieurs images dans le second cas)
* Fermeture du Fichier
* Traitement de la Séquence : Boucle Infinie
  - Constitution du nom de l'image courante et de l'image résultat
  - Ouverture du Fichier de l'Image Courante
  - Si Impossibilité : Fin de Séquence : Sortie de Boucle
  - Lecture de l'Entête, puis Lecture de l'Image Courante
  - Traitement proprement dit
  - Enregistrement de l'image résultat
  - Fermeture des Fichiers
    
```

III / Différence Simple du Premier Ordre :

III. 1°) Principe

La différence d'image s'effectue pixel à pixel, dans les 3 plans RVB, entre l'image courante et l'image de référence (qui n'est pas nécessairement la première image de la séquence).

III. 2°) Détection de Mouvement

Le mouvement n'est pas détecté au pixel de coordonnées (x,y) de l'image courante I(n) si et seulement si :

```

| I(n)_R(x,y) - I(ref)_R(x,y) | <= Seuil ET
| I(n)_V(x,y) - I(ref)_V(x,y) | <= Seuil  ET
| I(n)_B(x,y) - I(ref)_B(x,y) | <= Seuil
    
```

Où $I(n)_{R(x,y)}$, $I(n)_{V(x,y)}$ et $I(n)_{B(x,y)}$ sont respectivement les composantes rouge, verte et bleue du pixel de coordonnées (x,y) dans l'image courante I(n) et Seuil une valeur de seuil fixée a priori.

III. 3°) Visualisation des Résultats

L'image des résultats est également une image couleur RVB codée ensuite .ppm. Pour un pixel donné : si le mouvement est détecté : le pixel est recopié de l'image courante (dans les 3 plans : R,V et B). Si pas de mouvement détecté, le pixel est visualisé en noir. Cette visualisation sera également utilisée pour les résultats de la seconde méthode : par ajustement à un gabarit.

IV/ Différence du Premier Ordre par référence à une Image de Fond:

IV. 1°) Principe

Le principe consiste à détecter le mouvement pixel à pixel par différence entre l'image courante I(n) et l'image de fond B(n), puis de remettre à jour l'image de fond B(n+1) en fonction de I(n), B(n) et du mouvement détecté.

La mise à jour du fond, c'est-à-dire l'obtention de l'image $B(n+1)$ est une combinaison linéaire des images $I(n)$ et $B(n)$ affectées des coefficients α et $(1 - \alpha)$. α est appelé terme d'oubli. Dans les publications, on peut lire qu'il est généralement pris égal à 0.01, mais il sera intéressant de comparer les résultats obtenus avec différentes valeurs : les valeurs extrêmes : 1.0 et 0.0 (que représentent alors les algorithmes implantés), 0.1, 0.5, 0.9 etc ...

Pour la séquence proposée, quelle est la valeur du paramètre d'oubli donnant, à votre avis les meilleurs résultats.

IV. 2°) Détection de Mouvement

Le mouvement n'est pas détecté au pixel P de coordonnées (x,y) de l'image courante $I(n)$ si et seulement si :

$$\begin{aligned} &| I(n)_{_R}(x,y) - B(n)_{_R}(x,y) | \leq \text{Seuil} \text{ ET} \\ &| I(n)_{_V}(x,y) - B(n)_{_V}(x,y) | \leq \text{Seuil} \text{ ET} \\ &| I(n)_{_B}(x,y) - B(n)_{_B}(x,y) | \leq \text{Seuil} \end{aligned}$$

Où $I(n)_{_R}(x,y)$, $I(n)_{_V}(x,y)$, $I(n)_{_B}(x,y)$, $B(n)_{_R}(x,y)$, $B(n)_{_V}(x,y)$ et $B(n)_{_B}(x,y)$ sont respectivement les composantes rouge, verte et bleue du pixel de coordonnées (x,y) dans l'image courante $I(n)$ et l'image de Fond $B(n)$ et Seuil une valeur de seuil fixée a priori.

IV. 3°) Mise à Jour de l'image du Fond

L'image du fond est remise à jour au pixel courant (x,y) si celui-ci est immobile. L'algorithme est présenté ci-dessous.

```
* Si P(x,y) est immobile : remise à jour
    B(n+1)_{\_R}(x,y) = \alpha I(n)_{\_R}(x,y) + (1 - \alpha) B(n)_{\_R}(x,y)
    B(n+1)_{\_V}(x,y) = \alpha I(n)_{\_V}(x,y) + (1 - \alpha) B(n)_{\_V}(x,y)
    B(n+1)_{\_B}(x,y) = \alpha I(n)_{\_B}(x,y) + (1 - \alpha) B(n)_{\_B}(x,y)
* Sinon : pas de remise à jour
    B(n+1)_{\_R}(x,y) = B(n)_{\_R}(x,y)
    B(n+1)_{\_V}(x,y) = B(n)_{\_V}(x,y)
    B(n+1)_{\_B}(x,y) = B(n)_{\_B}(x,y)
```

IV. 4°) Initialisation de l'image du Fond

L'image du fond initiale $B(1)$ est initialisée par la première image de la séquence $I(0)$. La détection de mouvement commence bien évidemment à l'image $I(1)$.

IV. 4°) Implantation

Plusieurs implantations correspondant au cahier des charges ci-dessus sont possibles. Vous devez détailler et justifier dans le compte rendu l'implantation que vous avez réalisée.

V / Ajustement par rapport à un gabarit (facultatif):**V. 1°) Principe**

Le principe consiste à construire un gabarit à partir des n premières images de la séquence, qui ne contiennent pas la cible en mouvement. Ces images permettent d'apprendre les effets combinés « du bruit » de la caméra, des variations d'illumination dues aux nuages par exemple dans la séquence fournie, et au mouvement des branches arbres, du au vent, dans le fond de la scène.

Un gabarit simple peut être réalisé à partir de deux images « minimale » et « maximale » dans laquelle l'image courante doit se trouver en absence de mouvement.

V. 2°) Réalisation du Gabarit

Les images de référence minimale et maximale qui forment le gabarit, s'obtiennent de la même manière, pixel à pixel, selon les 3 composantes R,V et B. Ainsi pour l'image minimale I_{min} :

$$\begin{aligned} I_{min_R}(x,y) &= \min \{ I(k)_{_R}(x,y) / k \in [0,n] \} \\ I_{min_V}(x,y) &= \min \{ I(k)_{_V}(x,y) / k \in [0,n] \} \\ I_{min_B}(x,y) &= \min \{ I(k)_{_B}(x,y) / k \in [0,n] \} \end{aligned}$$

De même pour l'image maximale I_{max} :

$$\begin{aligned} I_{max_R}(x,y) &= \max \{ I(k)_{_R}(x,y) / k \in [0,n] \} \\ I_{max_V}(x,y) &= \max \{ I(k)_{_V}(x,y) / k \in [0,n] \} \\ I_{max_B}(x,y) &= \max \{ I(k)_{_B}(x,y) / k \in [0,n] \} \end{aligned}$$

V. 3°) Détection du Mouvement

Le mouvement n'est pas détecté au pixel de coordonnées (x,y) de l'image courante $I(n)$ si et seulement si les valeurs de l'intensité $I(n)_{_R}(x,y)$, $I(n)_{_V}(x,y)$, et $I(n)_{_B}(x,y)$ restent à l'intérieur du gabarit, soit :

$$\begin{aligned} I_{min_R}(x,y) &\leq I(n)_{_R}(x,y) \leq I_{max_R}(x,y) \quad \text{ET} \\ I_{min_V}(x,y) &\leq I(n)_{_V}(x,y) \leq I_{max_V}(x,y) \quad \text{ET} \\ I_{min_B}(x,y) &\leq I(n)_{_B}(x,y) \leq I_{max_B}(x,y) \end{aligned}$$

VI / De manière pratique :

Nous utiliserons l'environnement EdEnviTI, comme pour les précédentes séries de TP. En revanche, comme les images sont .ppm, il faudra effectuer les changements suivants :

- $prof = 3$ au lieu de 1, pour une image couleur,
- la lecture de l'entête s'effectue à l'aide de la fonction `LectureEntetePPM` (au lieu de `PGM`),
- les macros d'accès aux valeurs R,V et B du pixel du point « point » de l'image « image » sont respectivement `PIXEL_R(image, point, PIXEL_V(image, point), PIXEL_B(image, point)` au lieu de `PIXEL(image, point)` pour une image en niveaux de gris.