

Implémentation et Application du Filtre de Kalman pour le Lissage de Signaux Bruités : Cas du Suivi de sujet dans un Flux Vidéo

Simon Eveillé

Université de Technologie de Belfort-Montbéliard, Belfort, France

Mots clés : Filtre de Kalman, Lissage de signal, Suivi d'objet, Flux vidéo, Filtrage, Modèle de reconnaissance YOLO, Mesure de position, Suivi de trajectoire

Résumé

Le suivi d'objets dans un flux vidéo est une problématique clé en vision par ordinateur, avec des applications dans la surveillance, la robotique et la navigation autonome. Ce travail propose une méthodologie combinant le modèle de détection d'objets YOLO (You Only Look Once) et le filtre de Kalman pour suivre la trajectoire d'un piéton dans un flux vidéo bruité. YOLO est utilisé pour détecter et mesurer les positions des objets dans les images, tandis que le filtre de Kalman lisse les mesures bruitées et prédit la position future du piéton.

Les résultats montrent que cette approche réduit efficacement les oscillations et le bruit dans les trajectoires mesurées, tout en maintenant une précision élevée dans le suivi. Les graphiques comparant les positions mesurées par YOLO et celles prédites par le filtre de Kalman mettent en évidence la robustesse et la stabilité du système, en particulier face aux variations dynamiques et au bruit dans les observations.

I. Introduction

Le suivi d'objets dans un flux vidéo est un défi majeur en vision par ordinateur, avec des applications clés dans des domaines variés tels que la surveillance, la robotique, la navigation autonome ou encore la réalité

augmentée. Ce problème consiste à localiser et à suivre un ou plusieurs sujets au fil du temps dans une séquence d'images. Cependant, la précision du suivi est souvent mise à mal par le bruit inhérent aux systèmes de détection, les occlusions, les changements de perspective ou encore les variations d'éclairage. Pour surmonter ces limitations, il est essentiel de disposer de techniques capables de traiter efficacement des mesures bruitées et d'estimer avec fiabilité la position d'un objet au cours du temps.

Dans ce contexte, le filtre de Kalman s'impose comme une solution élégante et performante. Cet outil de filtrage probabiliste est couramment utilisé pour traiter des séries temporelles bruitées et prédire l'état d'un système dynamique. Grâce à son modèle mathématique basé sur des équations d'état et des observations, il permet d'estimer de manière optimale la position et la vitesse d'un objet, même en présence de données incertaines. Par ailleurs, l'essor des réseaux de neurones convolutifs a considérablement amélioré les capacités de détection d'objets en temps réel. Parmi ces réseaux, YOLO (You Only Look Once) s'est imposé comme une référence, grâce à sa rapidité et à sa précision, pour localiser des objets dans des flux vidéo.

Dans ce travail, nous proposons d'exploiter les forces complémentaires de ces deux outils : YOLO pour détecter et mesurer la position d'un objet dans chaque image de la

vidéo, et le filtre de Kalman pour lisser les mesures bruitées et estimer une trajectoire cohérente au fil du temps. Cette approche permet de répondre efficacement aux problématiques liées à la précision et à la robustesse du suivi d'objets dans des environnements dynamiques.

II. Revue de la littérature

L'article [1] sur le filtre de Kalman du livre de Dan Simon présente l'histoire du filtre ainsi que les personnes impliquées dans sa conception ainsi que les domaines pour lesquelles des applications lui ont été trouvées.

Ainsi, le filtre de Kalman a été développé par Rudolph Kalman, bien que Peter Swerling ait développé un algorithme très similaire en 1958. Le filtre porte le nom de Kalman parce qu'il a publié ses résultats dans une revue plus prestigieuse et que son travail était plus général et plus complet. Le filtre est parfois appelé filtre de Kalman-Bucy en raison des premiers travaux de Richard Bucy sur le sujet, menés conjointement avec Kalman. Les racines de l'algorithme remontent à la méthode des moindres carrés de Karl Gauss, âgé de 18 ans, en 1795. Comme beaucoup de nouvelles technologies, le filtre de Kalman a été développé pour résoudre un problème spécifique, en l'occurrence la navigation des engins spatiaux dans le cadre du programme spatial Apollo. Depuis, le filtre de Kalman a trouvé des applications dans des centaines de domaines divers, y compris toutes les formes de navigation (aérospatiale, terrestre et marine), l'instrumentation des centrales nucléaires, la modélisation démographique, la fabrication, la détection de la radioactivité souterraine, la logique floue et l'entraînement des réseaux neuronaux.

Un autre excellent article rédigé par Greg Welch et Gary Bishop, intitulé *An Introduction to the Kalman Filter*, constitue une référence incontournable pour comprendre les principes fondamentaux de cet algorithme.

Cet article didactique a pour objectif d'offrir une introduction claire et progressive au filtre de Kalman, en expliquant non seulement ses bases théoriques mais aussi ses applications pratiques.

L'un des points majeurs de l'article est la présentation structurée du schéma général du filtre, qui se divise en deux étapes principales : la période de **prédiction** et la période de **correction**. Ces deux phases fonctionnent de manière cyclique et itérative, permettant au filtre de s'adapter en permanence à des systèmes dynamiques :

1. **La période de prédiction** : Cette phase utilise un modèle dynamique pour anticiper l'état futur du système à partir de l'état précédent et des équations du mouvement. Elle génère une **estimation a priori** de la position et de la vitesse de l'objet (ou de toute autre grandeur suivie), basée uniquement sur les connaissances internes du modèle. Cette prédiction est accompagnée d'une estimation de l'incertitude, représentée par une matrice de covariance qui décrit la précision attendue de cette estimation.
2. **La période de correction** : Dans cette phase, les mesures bruitées obtenues par un capteur (comme celles produites par YOLO dans notre cas) sont comparées à l'estimation a priori. Le filtre ajuste alors la prédiction précédente en combinant ces nouvelles observations avec les estimations du modèle, en pondérant les deux sources d'information selon leur degré de fiabilité. Ce processus génère une **estimation a posteriori**, qui devient la meilleure approximation possible de l'état réel du système. Cette mise à jour concerne à la fois la valeur estimée de l'état et l'incertitude associée.

L'article met particulièrement en lumière la distinction entre ces deux périodes, en

soulignant que le filtre de Kalman ne se contente pas de "lisser" les données mais effectue un processus mathématique rigoureux basé sur la théorie de l'estimation bayésienne. La prédiction exploite les dynamiques connues du système (via des

équations différentielles ou des modèles simplifiés), tandis que la correction affine cette prédiction à partir des observations réelles, corrigeant ainsi les biais potentiels et réduisant les erreurs accumulées.

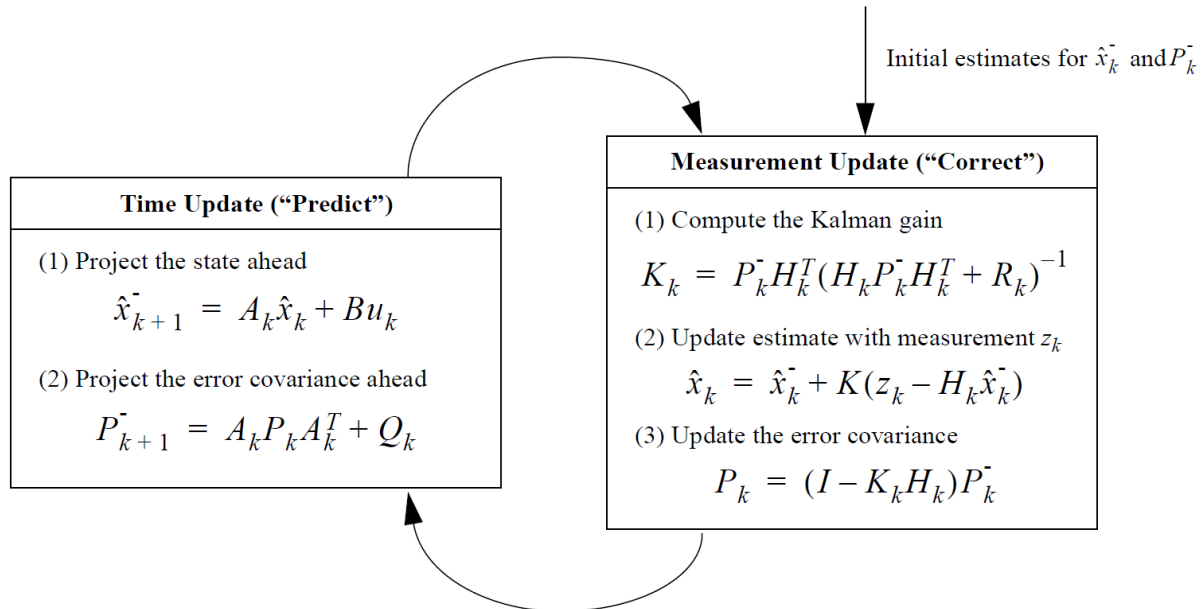


Figure 1 - Schéma général du filtre de Kalman présenté dans l'article de Welch G. et Bishop G.

III. Méthodologie

L'objectif de cette étude est de développer et d'évaluer un pipeline combinant un modèle de détection d'objets (YOLO) et un filtre de Kalman pour le suivi précis d'un objet, ici un piéton, dans un flux vidéo bruité. La méthodologie repose sur plusieurs étapes clés, détaillées ci-dessous :

A. Détection initiale du sujet avec YOLO

Le modèle YOLO (You Only Look Once) est utilisé pour détecter les objets dans chaque image du flux vidéo. YOLO est un réseau de neurones convolutifs rapide et précis qui, pour chaque image, identifie les objets et retourne leurs coordonnées sous forme de boîtes englobantes (bounding boxes). Pour chaque boîte correspondant au piéton étudié

(classe "person"), la position centrale de la boîte englobante est calculée et utilisée comme mesure brute de la position du piéton. Il s'agit d'un signal exposé à divers bruits : variations de la taille de la boîte englobante, mouvements intempestifs de la caméra (bien que relativement stable dans notre cas) et bruit intrinsèque du modèle de réseau de neurones YOLO.

B. Paramétrage du Filtre de Kalman pour le suivi de sujet

Dans notre approche, nous implémentons notre propre filtre de Kalman en Python afin de lisser le bruit sur la mesure de la position du piéton mais aussi pour anticiper l'état suivant du système. Autrement dit, le filtre nous permet de prédire la position la plus probable pour notre sujet lors de la prochaine prise de mesure.

Notre algorithme est structuré en 2 étapes, exactement comme il est présenté dans l'article [2] de Welch G. et Bishop G.

Prédiction

La phase de prédiction s'appuie sur un modèle de mouvement constant où les positions et vitesses futures sont estimées en utilisant des équations matricielles. Les matrices au cours de cette étape sont définies comme suit :

- **Matrice d'état (X) :** Elle caractérise l'état du système. Dans notre cas elle est composée des positions selon x et y, ainsi que des composantes selon x et y de la vitesse du sujet.

$$X = \begin{pmatrix} x(t) \\ y(t) \\ \dot{x}(t) \\ \dot{y}(t) \end{pmatrix}$$

- **Matrice d'évolution (A) :** Modélise le déplacement du piéton à chaque intervalle de temps, en prenant en compte la position et la vitesse. Elle caractérise l'évolution du système entre deux étapes qui se suivent.

$$A = \begin{pmatrix} 1 & 0 & dt & 0 \\ 0 & 1 & 0 & dt \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Dans notre cas, nous avons choisi un modèle en vitesse constante, étant donné que la vitesse du sujet n'a que très peu varié au cours de l'enregistrement.

La valeur de dt correspond au délai qui sépare 2 mesures, soit un 30^{ème} de seconde étant donné que la vidéo a été enregistrée en 30 fps (frame per second).

- **Matrice de commande (B) :** Nulle dans notre cas, car aucune force ou contrôle externe n'influence le sujet.

$$B = 0 \times Id(4)$$

À chaque étape de prédiction, la position prédite est stockée avant de la comparer avec la mesure réelle.

Correction

Lorsqu'une nouvelle mesure de position est disponible, la phase de correction est déclenchée. Le filtre ajuste son estimation en combinant les informations provenant de la mesure (position observée) et de la prédiction (modèle dynamique). Ce processus utilise un gain de Kalman calculé dynamiquement, qui pondère la contribution relative des deux sources d'information en fonction de leur fiabilité.

Voici les matrices les matrices à paramétrer pour cette étape :

- **Matrice d'observation (H) :** Lie les variables d'état (position et vitesse) aux mesures de position fournies par YOLO. La matrice H permet le passage du domaine des prédictions (vecteur contenant position et vitesse) au domaine des mesures (vecteur contenant uniquement les informations de position).

$$H = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

- **Covariance des mesures (R) :** Représente l'incertitude des mesures bruitées.

$$R = \begin{pmatrix} r_x & 0 \\ 0 & r_y \end{pmatrix}$$

Avec r_x l'incertitude sur la mesure selon l'axe x et r_y l'incertitude sur la mesure selon l'axe y.

- **Matrice de bruit de processus (Q) :** Capture l'incertitude du modèle.

$$Q = q \times Id(4)$$

C. Initialisation du filtre

L'initialisation du filtre de Kalman nécessite une estimation initiale de la position, de la vitesse et de la covariance d'erreur initiale. Les deux premières mesures de YOLO sont utilisées pour calculer la vitesse initiale du piéton.

D. Intégration dans le flux vidéo

Le filtre de Kalman est intégré dans une boucle de traitement en temps réel. À chaque image capturée par la caméra, les étapes suivantes sont exécutées :

1. YOLO détecte les objets et retourne les coordonnées des piétons.
2. Si le filtre est initialisé, une prédiction de la prochaine position est effectuée.
3. Si une nouvelle mesure est disponible, la phase de correction est appliquée.
4. La position corrigée est enregistrée et visualisée.

Enfin, les différentes position mesurées et lissées sont enregistrées afin d'afficher la trajectoire du sujet au cours de l'enregistrement vidéo.

IV. Présentation des résultats

A. Trajectoire globale

Le premier graphique [annexe N°1] représente la trajectoire du piéton dans le champ de vision de la caméra. Les courbes rouges indiquent les positions mesurées par YOLO, tandis que les courbes bleues représentent les positions lissées et prédites par le filtre de Kalman. Ce graphique montre clairement que :

- Les positions mesurées (rouge) présentent des variations rapides et bruitées.

- Les positions prédites par le filtre de Kalman (bleu) suivent une trajectoire plus stable et cohérente, réduisant l'impact du bruit.

Cette stabilisation est essentielle pour un suivi robuste, en particulier dans des environnements où les mesures peuvent être affectées par des erreurs, des occlusions ou des variations rapides.

B. Zoom sur les variations verticales

Le second graphique offre un zoom sur les variations verticales de la trajectoire, mettant davantage en lumière les performances du filtre de Kalman. Nous pouvons faire les observations suivantes :

- Les mesures YOLO montrent des oscillations significatives dues à des perturbations dans le flux vidéo ou des incertitudes dans la détection.
- Le filtre de Kalman, en intégrant les dynamiques du mouvement, produit une estimation plus lisse et plus proche de la trajectoire réelle.

Ce comportement illustre la capacité du filtre à compenser les limitations des mesures brutes en exploitant les informations temporelles et dynamiques.

V. Conclusion

Les résultats obtenus démontrent l'efficacité de l'approche proposée, combinant le modèle de détection YOLO et le filtre de Kalman pour le suivi d'un piéton dans un flux vidéo. L'intégration de ces deux outils exploite les forces complémentaires de chaque méthode : YOLO offre une détection rapide et précise des objets dans chaque image, tandis que le filtre de Kalman permet de traiter les mesures bruitées pour produire une estimation lissée et cohérente de la

trajectoire. Cette combinaison répond efficacement aux défis du suivi d'objets, notamment en présence de bruit, de variations dynamiques et d'incertitudes dans les observations.

En particulier, les graphiques montrent que le filtre de Kalman améliore significativement la robustesse des trajectoires reconstruites en réduisant les oscillations dues aux erreurs de détection, tout en maintenant une précision élevée dans le suivi. Cet aspect est essentiel dans des environnements réels, où des conditions telles que les occlusions, les variations d'éclairage ou les mouvements imprévus peuvent perturber les systèmes de suivi traditionnels.

L'approche proposée est par ailleurs flexible et peut être généralisée à d'autres scénarios de suivi d'objets dans des applications variées, telles que la surveillance vidéo, la navigation autonome, ou encore la robotique mobile. La modularité du filtre de Kalman, avec ses capacités d'adaptation à différents modèles dynamiques, ouvre également la voie à son utilisation pour le suivi d'objets ayant des mouvements plus complexes, tels que des véhicules ou des drones.

Cependant, cette étude présente quelques limitations. Tout d'abord, la précision du suivi dépend fortement des performances du modèle YOLO pour détecter l'objet cible dans chaque image. Si l'objet est perdu pendant plusieurs frames, le filtre de Kalman seul ne peut compenser cette absence, car il repose sur des mesures régulières pour ajuster ses prédictions. De plus, dans les cas de mouvements abrupts ou très rapides, les hypothèses de modèle linéaire du filtre peuvent s'avérer insuffisantes, nécessitant des adaptations telles que l'utilisation d'un filtre de Kalman étendu (EKF) ou d'un filtre de particules.

Pour les travaux futurs, plusieurs pistes peuvent être explorées pour améliorer et élargir cette méthodologie. D'une part, l'intégration de réseaux de neurones récurrents (RNN) ou de modèles basés sur des transformations temporelles pourrait renforcer la capacité à prédire des trajectoires complexes. D'autre part, l'utilisation de modèles YOLO plus avancés ou spécialisés pourrait accroître la précision de la détection dans des environnements difficiles. Enfin, il serait intéressant d'étudier des systèmes multi-objets combinant des approches similaires pour suivre simultanément plusieurs cibles tout en minimisant les ambiguïtés liées à l'identité des objets.

En résumé, cette étude démontre que la combinaison de YOLO et du filtre de Kalman constitue une approche prometteuse pour le suivi d'objets dans des flux vidéo bruités. En optimisant les étapes de détection et de suivi, cette méthode offre une solution robuste, précise et adaptable, contribuant ainsi au développement de systèmes de vision par ordinateur performants et fiables. Cette recherche pose les bases pour des applications avancées en vision par ordinateur, tout en ouvrant la voie à des améliorations futures dans ce domaine en constante évolution.

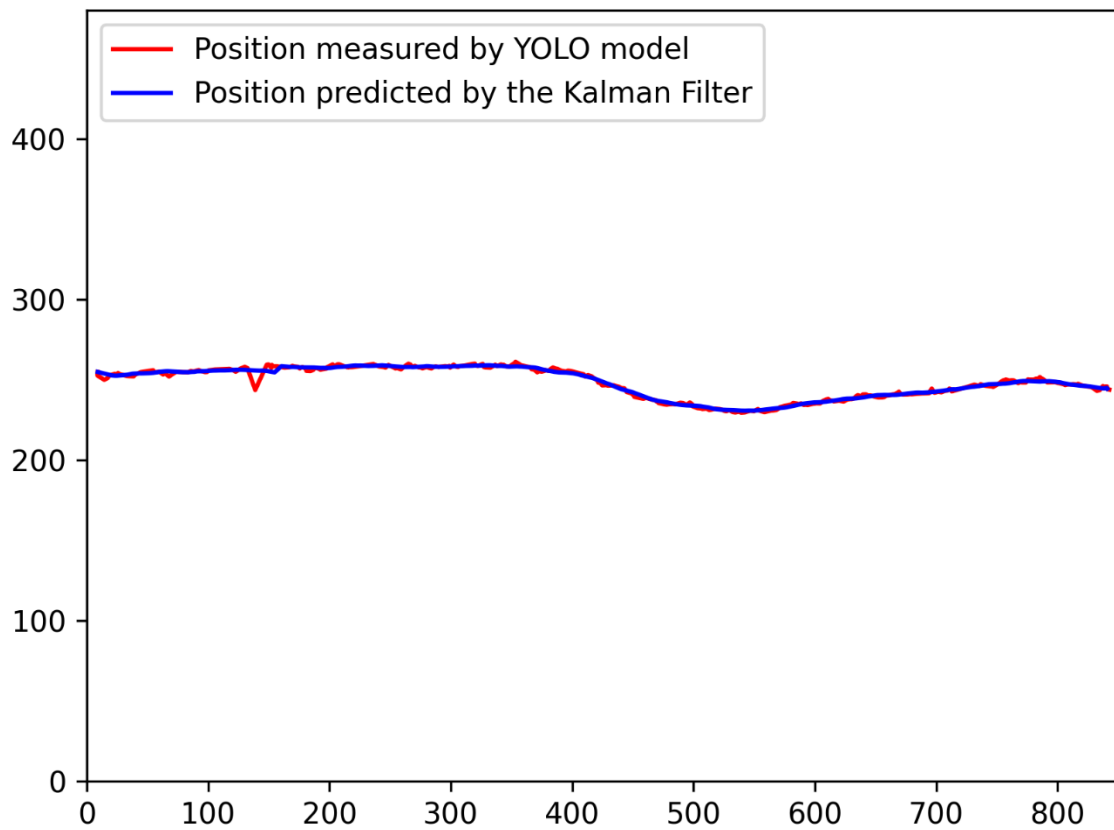
VI. Références

- [1] SIMON, Dan. Kalman filtering. Embedded systems programming, 2001, vol. 14, no 6, p. 72-79.
- [2] WELCH, G. An Introduction to the Kalman Filter. 1995.

VIII. Annexes

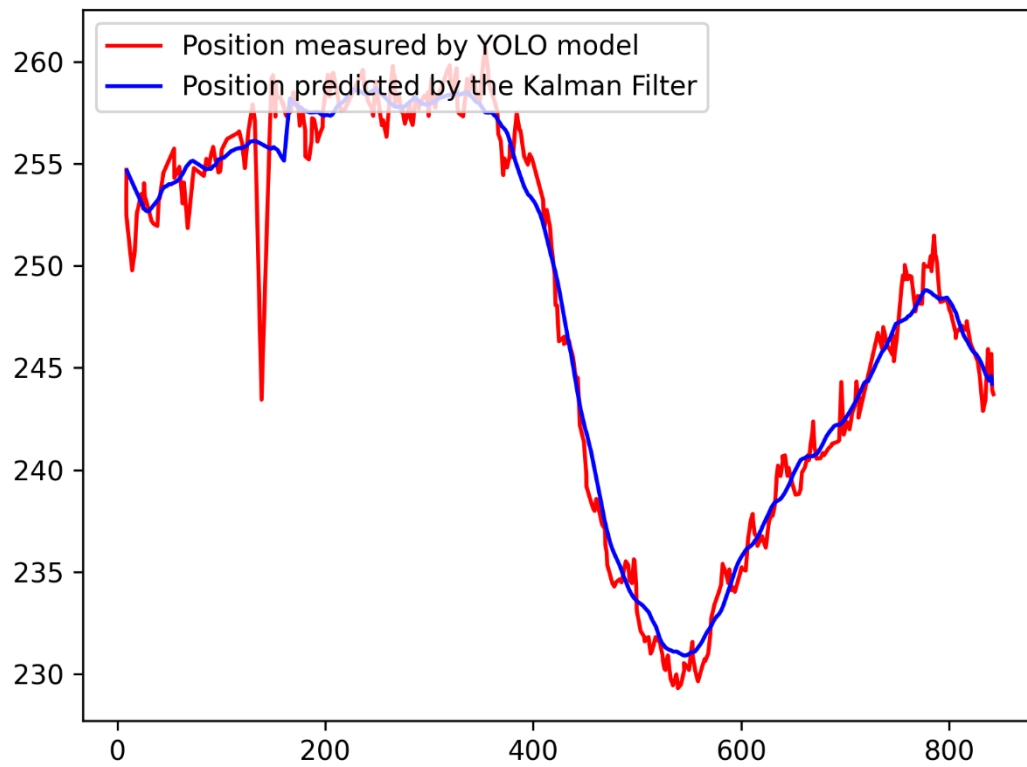
Annexe n°1 : Trajectoire du sujet dans le champs de vision de la caméra

Trajectory of the pedestrian in the camera view



Annexe n°2 : Trajectoire du piéton dans le champs de vision de la caméra avec étirement vertical pour mieux voir le lissage du bruit

Trajectory of the pedestrian in the camera view



Annexe n°3 : Exemple d'image (frame n°190) analysée par YOLO avec affichage des boîtes englobantes

