
Introduction

Grâce aux progrès réalisés dans les domaines de l'électronique et de la micro-électronique, l'acquisition de séquences vidéo est devenue une tâche d'une incroyable trivialité. Ainsi, en vision par ordinateur, les algorithmes travaillant sur les séquences vidéo ont connu un essor remarquable ces dernières années [SZE 10]. Si on feuillette un livre consacré à la vision par ordinateur écrit il y a 30 ans [BAL 82], on remarque que la notion de mouvement n'y est abordée presque qu'en termes d'estimation : il ne s'agissait alors que de détecter les mouvements et non de les analyser. En particulier, l'analyse du flot optique [BAR 94], très populaire à l'époque, ne permettait que de caractériser les changements temporels au sein de la séquence. Petit à petit, avec l'amélioration rapide de la qualité des capteurs et donc de la résolution des images qu'ils délivrent, mais également de la puissance de calcul et de la mémoire des ordinateurs, il est devenu possible, voire essentiel, en plus de détecter les mouvements, de les analyser : d'où viennent-ils ? Quels comportements traduisent-ils ? Ainsi, de nouveaux algorithmes ont fait leur apparition [SHI 94], dont l'objectif était de détecter et de suivre des entités dans une séquence vidéo. Ceux-ci sont regroupés sous l'appellation d'algorithmes de suivi (*tracking*). De nos jours, le suivi mono ou multi-objets dans les séquences vidéo est une des thématiques majeures de la vision par ordinateur. Il a en effet de nombreuses applications pratiques, notamment en interaction homme-machine, en réalité augmentée, en contrôle de trafic, en surveillance, en imagerie médicale ou biomédicale ou encore dans les jeux interactifs. La diversité des problématiques à résoudre ainsi que les challenges informatiques posés par le suivi d'objets dans les séquences vidéo motivent des recherches toujours plus nombreuses chaque année. Ainsi, pour ne parler que de 2012, dans les trois conférences majeures en vision par ordinateur (CVPR, ECCV et BMVC), trois ateliers et deux tutoriaux lui ont été consacrés, et on peut bien sûr citer l'atelier PETS (*International Workshop on Performance Evaluation of Tracking and Surveillance*), qui propose tous les 2 ans un concours autour du suivi. Par ailleurs, de plus en plus de bases de données sont disponibles pour permettre aux chercheurs de confronter leurs

résultats [SIG 10a, WU 13]¹. L'activité de recherche intense autour du suivi d'objets dans les séquences vidéo s'explique par le fait que celui-ci soulève de nombreux challenges. En effet, il nécessite d'extraire efficacement des images l'information concernant le ou les objets à suivre, puis de la modéliser afin d'en obtenir une représentation à la fois précise et compacte, et répondre à un compromis entre qualité et efficacité du suivi. Il faut ensuite pouvoir effectuer des liens temporels entre les instances des objets à chaque instant, tout en sachant gérer les éventuelles apparitions et disparitions d'objets dans la scène. Enfin, il faut parfois extraire des métadonnées pour répondre aux besoins d'une application spécifique (analyse d'un comportement, détection d'un événement, etc.). Ajoutons à ces difficultés celles induites par les changements de pose de l'objet (apparence, déformation), les variations d'illumination de la scène, le bruit présent dans les images, les occultations entre objets, etc. Le suivi d'objets se révèle donc un exercice très complexe, d'autant plus que, dans les applications pratiques, le niveau d'exigence en termes de qualité du suivi et de rapidité de calcul ne cesse d'augmenter.

Ces dernières années, les méthodes séquentielles de Monte-Carlo [DOU 01, GOR 93, ISA 98a], plus connues sous le nom de filtres particulaires, se sont imposées comme les algorithmes par excellence pour le suivi visuel. Elles ont pour objectif d'estimer la densité de filtrage liant les états des objets à suivre aux observations passées et présentes en l'approchant par un échantillon pondéré. Outre leur simplicité de mise en œuvre, ces approches sont capables de maintenir au cours du temps des hypothèses multiples, ce qui les rend robustes face aux challenges du suivi visuel. De plus, de par leur nature probabiliste, leur formalisme, très générique, permet d'envisager des modélisations complexes pour les objets et les observations dont on dispose, dont les densités peuvent être non paramétriques et/ou multimodales. Leur utilisation demande toutefois de rester vigilant à ne pas sortir de leur cadre mathématique, certes simple mais rigoureux. De plus, il faut prendre garde à ce que les coûts algorithmiques soient raisonnables (en incorporant, par exemple, des hypothèses d'indépendance quand elles se justifient). C'est donc très naturellement dans ce contexte méthodologique que nous nous sommes positionnés, notamment en observant que certains des avantages principaux offerts par le filtre particulaire ne pouvaient pas, à l'heure actuelle, être exploités autrement qu'en faisant un certain nombre d'hypothèses souvent simplificatrices. En particulier, si le fait de maintenir de multiples hypothèses au cours du temps est un réel avantage du filtrage particulaire, leur nombre minimal pour assurer le maintien d'une bonne estimation de la densité de filtrage doit être d'autant plus élevé que la modélisation choisie pour les données induit une représentation dans des espaces de grande dimension. Ceci pose de réels problèmes dès lors que l'on souhaite affiner une représentation en intégrant, par exemple, toute la richesse descriptive fournie par les capteurs dont on dispose.

1. Une liste est disponible au lien http://clickdamage.com/sourcecode/cv_datasets.php.

L'objectif de cet ouvrage est de présenter diverses contributions concernant la gestion des grands espaces de représentation d'état et d'observation, ce qui nous paraît être un des verrous majeurs du filtrage particulaire. Nous avons distingué trois axes principaux qui ont guidé les travaux de recherche et qui feront l'objet des chapitres 2 à 4.

Le premier axe concerne le choix de la modélisation des données afin d'en alléger la représentation ainsi que d'accélérer leur extraction. Les travaux de cet axe sont essentiels pour simplifier les calculs relatifs à l'estimation par filtrage particulaire. En effet, pour être résolu de manière robuste et précise, les problèmes de suivis actuels nécessitent l'exploitation d'une multiplicité d'informations/observations disponibles, dont la qualité ne cesse d'augmenter. Cela requiert donc des descriptions toujours plus fines de la scène dynamique que l'on traite, qui ont ainsi tendance à alourdir considérablement les calculs. S'il existe aujourd'hui des techniques fiables et efficaces d'extraction de données qui ont permis de mieux exploiter l'information image, elles ne sont pas forcément adaptées au cas du filtrage particulaire et à sa représentation multi-hypothèses car elles induisent une répétition des calculs qui peut s'avérer désastreuse pour l'efficacité du filtre. Ainsi, les histogrammes sont un mode de représentation très courant en suivi par filtrage particulaire, mais leur extraction peut rapidement devenir un goulet d'étranglement pour les temps de réponse du filtre et il convient donc de trouver des méthodes adaptées pour les extraire. La dimension des espaces sur lesquels on travaille ayant une influence significative sur les temps de réponse, être capable de combiner un ensemble de caractéristiques, d'observations, d'informations, et de fournir un modèle décrit dans un espace de dimension réduite devient donc également une tâche critique à laquelle se sont attelés de nombreux chercheurs. Comme nous le verrons dans la suite, des combinaisons efficaces peuvent être réalisées soit pendant le processus de correction, ce qui peut être vu comme une fusion *a posteriori*, soit être intégrées directement au processus de suivi (propagation et correction).

Plusieurs solutions théoriques et/ou algorithmiques ont été proposées qui ont permis d'obtenir, soit une diminution significative des temps de calcul des fonctions de vraisemblance, notamment par des extractions optimisées d'histogrammes, soit des modélisations originales de problèmes classiques en suivi, tels que la déformation de l'objet au cours du temps, ses multiples représentations possibles (apparences, modalités et fragments), ou encore la détection d'objets (nouveaux ou pas) entre deux images.

Le deuxième axe concerne l'exploration de l'espace d'état. En effet, le filtre particulaire possède une phase de propagation des hypothèses (particules) ayant pour but de les diffuser vers les zones à forte vraisemblance, dans lesquelles un fort poids leur sera affecté lors d'une étape de correction. Lorsque l'espace d'état est de grande taille, détecter ces zones à forte vraisemblance s'avère difficile car une exploration exhaustive de l'espace nécessite qu'on multiplie significativement

le nombre d'hypothèses, ce qui s'avère impossible dans des temps de calcul raisonnables. Ce problème peut être résolu de deux manières différentes. La première constitue notre deuxième axe de recherche et consiste à « choisir » les zones de l'espace à explorer, ce que nous appelons « focaliser ». Ainsi, on peut maintenir un nombre d'hypothèses raisonnable tout en explorant astucieusement les zones dans lesquelles elles seront propagées, zones supposées à forte vraisemblance. Cela peut se faire, soit par le biais de détections avant la propagation (*detect-before-track*), soit par la définition de fonctions de proposition dédiées. C'est la deuxième solution que nous avons choisi de développer dans cet ouvrage et qui consiste à décomposer l'espace d'état en sous-espaces de petites tailles sur lesquels les calculs sont réalisables avec « peu » d'hypothèses.

Nous avons proposé plusieurs approches permettant de mieux focaliser dans l'espace d'état et, ainsi, accélérer le processus de suivi par filtrage particulaire. Deux types de contributions dédiées au suivi multi-objets ont permis de ne pas avoir à évaluer toutes les combinaisons possibles d'association entre mesures et objets. Le premier type vise, tout d'abord, à modéliser une fonction de proposition permettant de propager les particules uniquement dans les zones en mouvement qui ont été, au préalable, détectées, et qui sont vues comme des mesures, puis à classer ces particules en les associant aux objets de la scène. Le deuxième prend en compte la dynamique passée des objets et propose un modèle d'association de données très peu contraint, puisque dépendant de peu de paramètres. Ces modèles permettent de calculer simplement les probabilités d'association entre les mesures et les objets. Nous montrerons qu'il est également possible d'introduire des informations spatiales floues dans le filtre particulaire. Cela a permis la modélisation d'une nouvelle fonction de proposition prenant en compte, non seulement l'observation courante, mais également l'historique des relations spatiales floues ayant caractérisé les trajectoires passées des objets. Le résultat permet un suivi beaucoup plus souple et mieux adapté aux changements brusques de trajectoire ou de forme. L'intérêt de ce type de modélisation est montré à travers des applications variées, telles que le suivi d'un objet, en gérant des mouvements erratiques, le suivi multi-objets, en traitant le cas des occultations et enfin le suivi multi-formes, en traitant les objets déformables.

Comme le précédent, le troisième axe vise à rendre possible en pratique l'exploration des grands espaces d'états. Mais, ici, on ne cherche plus à travailler dans l'espace complet dont on a réduit l'hyper-volume d'exploration. On propose plutôt de le décomposer en sous-espaces de petites tailles dans lesquels les calculs peuvent être réalisés en des temps raisonnables car ceux-ci permettent d'estimer des distributions ayant beaucoup moins de paramètres que celles de l'espace complet. Ce dernier est ainsi défini comme un espace joint. Nous nous intéressons dans cet ouvrage aux méthodes de décomposition « non approchées », c'est-à-dire celles qui garantissent asymptotiquement que les particules échantillonnent correctement la distribution de filtrage sur l'espace complet. Ces méthodes ne font donc aucune hypothèse simplificatrice et exploitent uniquement des indépendances existant dans le

problème de suivi. Parmi ces techniques, l'échantillonnage partitionné (*partitioned sampling*) est aujourd'hui très utilisé, mais connaît quelques limites. En effet, il parcourt séquentiellement tous les sous-espaces afin de construire, progressivement, l'ensemble de particules sur l'espace joint (l'espace complet) et cela peut soulever des problèmes lorsque l'ordre dans lequel ces sous-espaces sont parcourus est totalement arbitraire : si la sous-hypothèse réalisée dans le premier sous-espace est mauvaise, elle contribuera à diminuer le score global de l'hypothèse, mêmes si les autres sous-hypothèses sont bonnes. Ainsi le suivi sera de mauvaise qualité.

Il existe des contributions qui ont permis de résoudre ce problème. Tout d'abord, la possibilité d'ajouter au processus d'estimation l'ordre dans lequel les objets doivent être traités. Cet ordre est estimé séquentiellement, en même temps que les états des objets, ce qui permet de considérer en dernier les objets les moins fiables pour le suivi. On peut également exploiter des indépendances conditionnelles intrinsèques au problème de suivi (sans faire d'hypothèse abusive). Cela amène naturellement à exploiter des réseaux bayésiens dynamiques plutôt que des chaînes de Markov afin de modéliser le processus de filtrage. L'exploitation des propriétés d'indépendance de ces réseaux nous a permis de développer une nouvelle méthode de permutation de certains sous-échantillons de sous-particules qui permet de mieux estimer les modes de la densité de filtrage tout en garantissant que la densité estimée reste inchangée. Cette méthode permet de réduire, non seulement les erreurs de suivi, mais également les temps de calcul. Cette idée de permutation est exploitée pour proposer une nouvelle méthode de rééchantillonnage qui permet, elle aussi, d'améliorer significativement le suivi.

Le plan de cet ouvrage est le suivant. Dans le chapitre 1, nous présentons les éléments théoriques essentiels nécessaires à la compréhension du filtrage particulaire. Nous expliquons ensuite comment cette méthodologie est utilisée dans le cadre du suivi visuel, en particulier quels en sont les points fondamentaux à considérer. Ceci nous permet enfin de décrire quelques limites et verrous actuels du suivi par filtrage particulaire et, ainsi, de justifier notre positionnement scientifique. Le chapitre 2 présente des contributions concernant la modélisation et l'extraction des données à traiter, ainsi que le choix de leur représentation pour les simplifier et ainsi accélérer les calculs. Dans le chapitre 3, nous décrivons quelques contributions permettant d'explorer l'espace d'état en se focalisant sur certaines zones spécifiques, jugées plus intéressantes que les autres. Le chapitre 4 montre, à travers quelques travaux, comment décomposer l'espace d'état en sous-espaces dans lesquels les calculs sont réalisables. Enfin, nous proposons une conclusion et une ouverture vers le futur du suivi, en particulier par filtrage particulaire dans le chapitre 5.