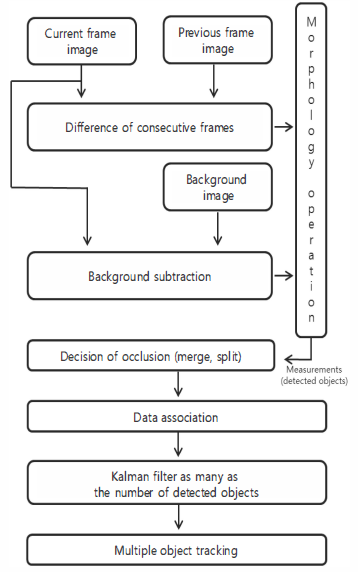
Note reference

1. Latha Anuj, M T Gopala Krishna, “Multiple camera based multiple object tracking under occlusion: A survey”, 2017 International Conference on Innovative Mechanisms for Industry Applications (ICIMIA). IEEE Conferences, pp. 432-437, 2017.
2. Haanju Yoo, et al, “Online Scheme for Multiple Camera Multiple Target Tracking Based on Multiple Hypothesis Tracking”, IEEE Transactions on Circuits and Systems for Video Technology, IEEE, Vol 27, Number 3, pp. 454-469, 2017.
3. Soonmin Hwang, et al, “Fast multiple objects detection and tracking fusing color camera and 3D LIDAR for intelligent vehicles”, Ubiquitous Robots and Ambient Intelligence (URAI), IEEE Conferences, pp. 234-239, 2016.
4. Jiayi Ma, et al, “Robust Feature Matching for Remote Sensing Image Registration via Locally Linear Transforming”, IEEE Transactions on Geoscience and Remote Sensing, IEEE Journals and Magazines, Vol 53, Number 13, pp. 6469-6481, 2015.
5. Yonglong Tian, Ping Luo, Xiaogang Wang, Xiaoou Tang, “Pedestrian detection aided by deep learning semantic tasks”, Computer Vision and Pattern Recognition (CVPR), IEEE Conferences, pp. 5079-5087, 2015.
6. Junwei Han, Dingwen Zhang, Gong Cheng, Lei Guo, Jinchang Ren, “Object Detection in Optical Remote Sensing Images Based on Weakly Supervised Learning and High-Level Feature Learning”, IEEE Transactions on Geoscience and Remote Sensing, IEEE Journals and Magazines, Vol 53, Number 6, pp. 3325-3337, 2015.
7. Shipra Ojha, Sachin Sakhare, “Image processing techniques for object tracking in video surveillance – A survey”, International Conference on Pervasive Computing (ICPC), IEEE Conferences, pp. 1-6, 2015.
8. Gedas Bertasius, Jianbo Shi, Lorenzo Torresani, “DeepEdge: A multi-sccale bifurcated deep network for top-down contour detection”, Computer Vision and Pattern Recognition (CVPR), IEEE Conferences, pp. 4380-4389, 2015.
9. Wei Shen, Xinggang Wang, Yan Wang, Xiang Bai, Zhijiang Zhang, “DeepContour: A deep convolutional feature learned by positive-sharing loss for contour detection”, Computer Vision and Pattern Recognition (CVPR), IEEE Conferences, pp. 3982-3991, 2015.
10. Yan Yang, Xiaodong Wang, Jiande Wu, Haitang Chen, Zhaoyuan Han, “An improved mean shift object tracking algorithm based on ORB feature matching”, The 27th Chinese Control and Decision Conference (CCDC), IEEE Conferences, pp. 4996-4999, 2015.
11. Wanli Ouyang, Xiao Chu, Xiaogang Wang, “Multi-source Deep Learning for Human Pose Estimation”, Computer Vision and Pattern Recognition (CVPR), IEEE Conferences, pp. 2337-2344, 2014.
12. Wei Li, Rui Zhao, Tong Xiao, Xiaogang Wang, “DeepReID: Deep Filter Pairing Neural Network for Person Re-identification”, Computer Vision and Pattern Recognition (CVPR), IEEE Conferences, pp. 152-259, 2014.
13. Jong-Min Jeong, Tae-Sung Yoon, Jin-Bae Park, “Kalman filter based multiple objects detection-tracking algorithm robust to occlusion”, SICE Annual Conference (SICE), IEEE Conferences, pp. 941-946, 2014.
14. Dumitru Erhan, Christian Szegedy, Alexander Toshev, Dragomir Anguelov, “Scalable Object Detection Using Deep Neural Networks”, Computer Vision and Pattern Recognition (CVPR), IEEE Conferences, pp. 2155-2162, 2014.
15. Alexander Toshev, Christian Szegedy, “DeepPose: Human Pose Estimation via Deep Neural Networks”, Computer Vision and Pattern Recognition (CVPR), IEEE Conferences, pp. 1653-1660, 2014.
16. Kaihua Zhang, Lei Zhang, Ming-Hsuan Yang, “Real-time Object Tracking Via Online Discriminative Feature Selection”, IEEE Transactions on Image Processing, IEEE Journals and Magazines, Vol 22, Number 12, pp. 4664-4677, 2013.
17. Rosten E, Drummond T, “Fusing points and lines for high performance tracking”, Tenth IEEE International Conference on Computer Vision(ICCV’05), IEEE Conferences, Vol 2, pp.1508-1515, 2005.
18. Michael Calonder, Vincent Lepetit, Christoph Strecha, Pascal Fua, “Brief: binary robust independent elementary features”, European Conference on Computer Vision (ECCV), Springer, pp. 778-792, 2010.

13) Tác giả Jong-Min Jeong trong bài nghiên cứu của mình [13] đã đưa ra giải pháp giải quyết vấn đề phủ lấp giữa các đối tượng di chuyển trong camera. Rõ rang trong thực tế, khi các đối tượng di chuyển qua lại trong camera ngẫu nhiên không theo một hướng nhất định thì việc hai đối tượng che phủ lẫn nhau trong camera là rất thường xuyên. Khi các đối tượng chồng lấp lên nhau như vậy, ta không thể sử dụng các phương pháp phát hiện đối tượng như background subtraction và motion information, supervisor learning để xác định đối tượng bị che phủ đằng sau được mà chỉ có thể sử dụng các phương pháp ước lượng, phỏng đoán. Do đó, trong nghiên cứu [13] tác giả Jong-Min Jeong và các đồng nghiệp đã sử dụng kalman filter và đề xuất phương pháp của mình nhằm giải quyết bài toán che phủ giữa các đối tượng. Đầu tiên, họ sử dụng background subtraction và motion information để phát hiện nhiều đối tượng di chuyển trong camera. Sau đó, họ xác định được số lượng các đối tượng di chuyển trong frame. Bước thứ hai, họ sử dụng Kalman Filter cho mỗi đối tượng ghi nhận được. Tuy nhiên, việc sử dụng một Kalman Filter cho một đối tượng ghi nhận được sẽ dẫn đến tình trạng ở frame ảnh tiếp theo họ không thể biết chính xác được đối tượng nào sẽ tương ứng với bộ Kalman Filter nào trước đó. Chính vì thế, họ đề xuất giải thuật xác định đối tượng ghi nhận và bộ Kalman Filter đúng của nó sử dụng hàm chi phí bao gồm các đặc trưng cũng như là xác định được hai đối tượng che phủ hợp nhất lại với nhau hay tách rời nhau. Hình 1 thể hiện các bước trong phương pháp đề xuất của tác giả và các đồng nghiệp.



Hình 1: Sơ đồ khối phương pháp đề xuất [13]

Hai bước quan trọng mà tác giả và các đồng nghiệp đề xuất để giải quyết được bài toán truy vết các đối tượng bị che phủ lẫn nhau là bước xác định các đối tượng che phủ đang hợp nhất lại với nhau hay đang tách ra và bước gán đối tượng phát hiện được trong frame ảnh tiếp theo đúng với bộ Kalman Filter của nó trong frame ảnh trước. Để phát hiện được các đối tượng đang hợp nhất hay tách rời nhau trong vùng che phủ, họ sử dụng tỉ lệ giữ chiều cao và chiều rộng của đối tượng phát hiện được so sánh với ngưỡng đề xuất. Cụ thể:

Với m: số lượng các đối tượng phát hiện được trong frame ảnh thứ k.

k: frame ảnh thứ k.

và là ngưỡng trên và ngưỡng dưới của tỉ lệ giữa chiều cao và chiều rộng của đối tượng phát hiện được trong frame ảnh.

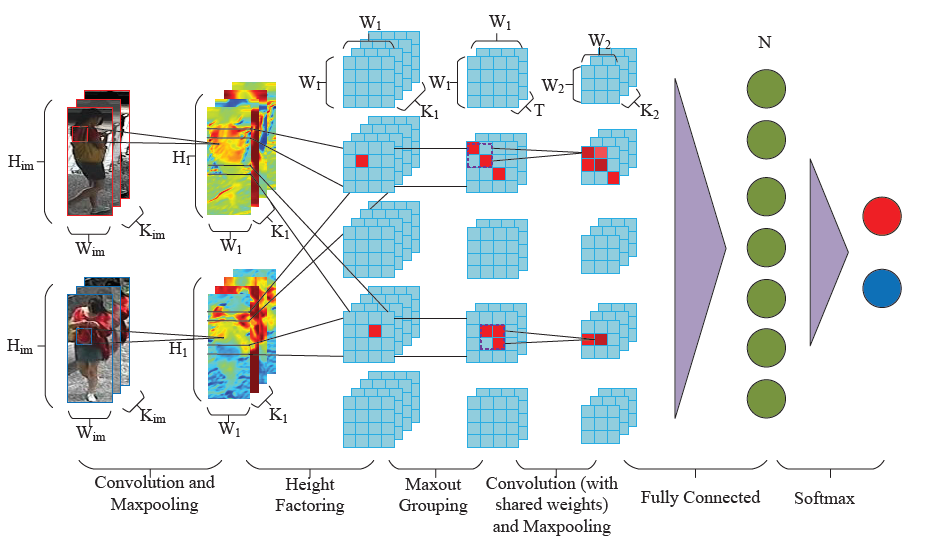
Bước quan trọng thứ hai là bước làm thế nào để xác định đúng đối tượng ghi nhận được với bộ Kalman Filter tương ứng của nó. Để thực hiện được điều này, họ sử dụng hai yếu tố là yếu tố về khoảng cách giữa giá trị dự đoán với giá trị của đối tượng ghi nhận được và yếu tố về diện tích của đối tượng giữa các frame ảnh với nhau.

14) Scalable Object Detection Using Deep Neural Networks

Tác giả Dumitru Erhan và các đồng nghiệp [14] đã đề xuất phương pháp phát hiện nhiều đối tượng trong một frame ảnh được gọi là “DeepMultiBox”. Đối với bài toán phát hiện đối tượng trong frame ảnh, ta có rất nhiều cách như là background subtraction và motion detection, supervisor learning… Phương pháp đề xuất này sử dụng DNNs là một trong các giải thuật thuộc lớp supervior learning. Mục tiêu của nghiên cứu là xây dựng được phương pháp dự đoán được tập các vùng chứa đối tượng gọi là bounding box, bounding box là một hình chữ nhật bao quanh đối tượng trong không gian 2D. Dữ liệu đầu ra của phương pháp phát hiện đối tượng được đề xuất bởi Dumitru Erhan và các đồng nghiệp ở đây bao gồm một tập các bounding box với các điểm tọa độ thể hiện vị trí của bounding box trong frame ảnh và giá trị cho biết độ tin cậy (tính chính xác) của việc xác định nhãn của đối tượng tương ứng với bounding box đó. Đóng góp chính của nghiên cứu này là xây dựng được một mạng noron học sâu để phát hiện được đối tượng và thu được dữ liệu đầu ra như mô tả trên.

12) DeepReID: Deep Filter Pairing Neural Network for Person Re-identification

Tác giả Wei Li và các đồng nghiệp của mình trong nghiên cứu [12] đã đề xuất xây dung một mạng nơ ron học sâu nhằm xác định định danh của người xuất hiện trong đoạn camera. Ở nghiên cứu [12] tác giả đề xuất mạng nơ ron bộ loc bắt cặp (FPNN – Filter Pairing neural Network) để xác định định danh của đối tượng di chuyển thường được sử dụng trong các hệ thống camera phân tán không có vùng trùng lắp hay bất cứ thông tin liên hệ trực tiếp với nhau dựa trên các đặc trưng về màu sắc của đối tượng. Điểm đóng góp của nghiên cứu này đó là thay vì sử dụng các đặc trưng thủ công bằng thao tác trực tiếp thì tác giả đề xuất phương pháp thu và học các đặc trưng từ dữ liệu có được một cách tự động. mạng nơ ron bộ lặc bắt cặp được mô tả như hình 1:



Hình 1: Filter pairing neural network[12]

10) An improved mean shift object tracking algorithm based on ORB feature matching

Tác giả Yan Yang và các đồng nghiệp trong bài nghiên cứu [10] đã đề xuất sử dụng đặc trưng ORB (Oriented FAST and Rotated BRIEF) để cải thiện hiệu suất của phương pháp truy vết đối tượng sử dụng Mean Shift. Giải thuật Mean Shift thông thường sử dụng đặc trưng về màu sắc của đối tượng để truy vết. Các đặc trưng màu ở đây được thu nhận từ không gian màu RGB và chuyển sang không gian màu HSV nhằm giảm bớt sự tác động từ các yếu tố ngoại như ánh sáng…Nhưng với nghiên cứu [10], tác giả sử dụng đặc trưng ORB là một sự cải tiến dựa trên phát hiện đặc trưng FAST[17] và mô tả đặc trưng BRIEF[18]. So với SIFT và SURF thì ORB cải tiến hơn về tốc độ tính toán cũng như đảm bảo tính bất biến của đặc trưng trong các trường hợp các đối tượng bị thay đổi vì xoay, thu phóng hay sự chiếu sáng từ bên ngoài.