

Abstract

Drone-based crowd monitoring is a key technology for applications in surveillance, public safety, and event management, primarily due to its dynamic, aerial perspective that surpasses the limitations of traditional ground-based systems. Recently, a new trend has emerged in tiny object localization and tracking, characterized by the use of point-oriented object sensing, which enables accurate monitoring of densely packed individuals in low-altitude aerial imagery. In this dissertation, advancements in this area are presented, including novel approaches for point-oriented object localization and a new solution for point-oriented object tracking. For localization and counting tasks, a series of enhancement mechanisms is introduced. These include the integration of motion-based features, the use of task-oriented synthetic data, and addressing the influence of varying image input resolutions in neural networks. A direct incorporation of drone altitude into the neural network architecture is also investigated, a new module that processes all pixels of high-resolution images without downscaling is proposed, and a novel loss function tailored to point-oriented localization is introduced. For object tracking and trajectory counting, an algorithm is proposed that enhances trajectory continuity and unique counting reliability in drone-based crowd monitoring, enabling the accurate tracking of individuals across video sequences. The approach extends the Simple Online and Real-time Tracking (SORT) framework by replacing the bounding-box assignment with a point-distance metric. It is further enhanced with three cost-effective techniques: camera motion compensation, altitude-aware assignment, and classification-based trajectory validation. Additionally, Deep Discriminative Correlation Filters (DDCF) are integrated, which reuse spatial feature maps from localization algorithms to improve computational efficiency and handle missed detections. To support this research, two new datasets, UP-COUNT and UP-COUNT-TRACK, are introduced, addressing challenges in modern drone imagery, including simultaneous camera and object motion, as well as changing flight altitudes. All proposed methods are quantitatively evaluated on both the publicly available DroneCrowd dataset and new datasets, demonstrating significant improvements in localization and tracking performance and achieving state-of-the-art results in drone-based people and trajectory counting. This dissertation makes substantial contributions to computer vision in aerial robotics, offering practical tools for rapid crowd size and movement estimation. These tools have been demonstrated to be applicable in real-world scenarios.

Streszczenie

Monitorowanie tłumu z wykorzystaniem dronów stanowi kluczową technologię w obszarach nadzoru, bezpieczeństwa publicznego oraz zarządzania wydarzeniami, głównie dzięki dynamicznej, lotniczej perspektywie, która jest pozbawiona ograniczeń tradycyjnych systemów naziemnych. W ostatnim czasie pojawił się nowy trend w lokalizacji i śledzeniu bardzo małych obiektów, oparty na punktowej detekcji obiektów, umożliwiając precyzyjne monitorowanie gęsto zgromadzonych osób w obrazach zarejestrowanych na niskim pułapu lotu. W niniejszej rozprawie przedstawiono postępy w tym obszarze, w tym nowe podejścia do punktowej lokalizacji obiektów oraz nowatorskie rozwiązanie dla punktowego śledzenia obiektów. Na potrzeby zadań lokalizacji i zliczania wprowadzono szereg usprawniających mechanizmów, obejmujących: integrację cech opartych na ruchu, wykorzystanie syntetycznych danych oraz zbadanie wpływu zmiany rozdzielczości obrazów wejściowych w sieciach neuronowych. Dodatkowo zbadano wpływ uwzględnienia wysokości lotu drona bezpośrednio w architekturze sieci neuronowej, zaproponowano nowy moduł sieci neuronowej przetwarzający wszystkie piksele obrazów wysokiej rozdzielczości bez potrzeby ich skalowania oraz opracowano nową funkcję kosztu dostosowaną do punktowej lokalizacji. W zakresie śledzenia obiektów i zliczania unikalnych trajektorii opracowano algorytm poprawiający ciągłość trajektorii oraz wiarygodność zliczania w kontekście monitorowania tłumu z dronów, umożliwiając precyzyjne śledzenie pojedynczych osób w sekwencjach wideo. Podejście to rozszerza algorytm SORT (Simple Online and Real-time Tracking), zastępując dopasowanie oparte na ramkach ograniczających metryką odległości punktowej. Ulepszenia obejmują także trzy efektywne z punktu widzenia kosztu obliczeniowego techniki: kompensację ruchu kamery, dopasowanie obiektów uwzględniające wysokość lotu oraz walidację trajektorii opartą na klasyfikacji. Ponadto, zintegrowano Deep Discriminative Correlation Filters (DDCF), które wykorzystują przestrzenne mapy cech bezpośrednio z algorytmu lokalizacji w celu zwiększenia efektywności obliczeniowej i lepszego radzenia sobie z brakującymi detekcjami. W ramach wsparcia badań wprowadzono dwa nowe zbiory danych: UP-COUNT oraz UP-COUNT-TRACK, odpowiadające na wyzwania nowoczesnego obrazowania z dronów, takie jak jednoczesny ruch kamery i obiektów oraz zmienna wysokość lotu. Wszystkie zaproponowane metody zostały poddane ilościowej ocenie zarówno na publicznie dostępnym zbiorze danych DroneCrowd, jak i na nowo wprowadzonych, wykazując istotną poprawę dokładności lokalizacji i śledzenia, osiągając przy tym najlepsze dotychczasowe wyniki w zadaniach zliczania osób i trajektorii z perspektywy powietrznej. Praca ta wnosi istotny wkład w rozwój widzenia komputerowego w robotyce lotniczej oraz dostarcza praktycznych narzędzi do efektywnej oceny liczebności i ruchu tłumu, co zostało potwierdzone w rzeczywistych scenariuszach.