

ABSTRACT

Automated drone surveillance can help to detect and defuse critical situations in large crowds at an early stage. Reliable analysis of movement patterns in drone videos is central to this, for example to prevent mass panics. Optical flow methods for determining motion vectors are an important tool, but their performance in drone scenarios with complex perspectives and dynamics has hardly been systematically investigated so far.

This thesis develops a prototypical and reproducible pipeline for estimating motion vectors in drone videos and compares three established methods: Lucas-Kanade (classical, sparse), Farnebäck (classical, dense), and RAFT (deep learning, dense). The basis consists of seven real drone videos from a festival with varying scene complexities, supplemented by manual ground-truth annotations. All methods are tested under identical conditions. The research question is: *Which of the selected classical and deep learning-based methods delivers the quantitatively and qualitatively best results for drone videos in terms of accuracy, robustness, and runtime when calculating motion vectors?*

The results show that classical methods (Lucas-Kanade and Farnebäck) provide more robust and accurate motion vectors in the examined drone videos than RAFT. Lucas-Kanade achieves the lowest mean endpoint error (approx. 1.29 px), Farnebäck is similarly precise and most reliably detects the diversity of directions in crowd dynamics. RAFT has higher errors and is less robust in complex drone scenarios. In terms of runtime, Lucas-Kanade is the only real-time capable method (approx. 7 ms/frame); Farnebäck and RAFT are considerably slower and unsuitable for time-critical applications.

The analysis provides clear recommendations: For real-time applications and limited resources, Lucas-Kanade is optimal; Farnebäck is suitable for detailed crowd analyses; RAFT can be an addition for offline evaluation and after targeted domain fine-tuning. This work establishes a solid basis for comparison, highlights methodological strengths and weaknesses, and provides practical recommendations for the selection of optical flow methods in drone-based movement scenarios. The outlook emphasizes the importance of larger ground-truth datasets, specific training of deep learning models, and optimizations for real-time and edge applications.

Thus, this master's thesis contributes to the methodological and practical advancement of movement analysis in drone videos and lays the foundation for future, practical applications in automated crowd surveillance.

ZUSAMMENFASSUNG

Automatisierte Drohnenüberwachung kann helfen, kritische Situationen in großen Menschenmengen frühzeitig zu erkennen und zu entschärfen. Die zuverlässige Analyse von Bewegungsmustern in Drohnenvideos ist dabei zentral, etwa zur Prävention von Massenpaniken. Optical-Flow-Methoden zur Bestimmung von Bewegungsvektoren sind ein wichtiges Werkzeug, doch deren Leistungsfähigkeit in Drohnenszenen mit komplexer Perspektive und Dynamik ist bislang kaum systematisch untersucht.

Diese Arbeit entwickelt eine prototypische und reproduzierbare Pipeline zur Schätzung von Bewegungsvektoren in Drohnenvideos und vergleicht drei etablierte Methoden: Lucas-Kanade (klassisch, sparse), Farnebäck (klassisch, dense) und RAFT (Deep Learning, dense). Basis sind sieben reale Drohnenvideos eines Festivals mit unterschiedlichen Szenenkomplexitäten, ergänzt durch manuelle Ground-Truth-Annotationen. Alle Verfahren werden unter identischen Bedingungen getestet. Die Forschungsfrage lautet hierbei: *Welche der ausgewählten klassischen und Deep-Learning-basierten Methoden liefert für Drohnenvideos in Bezug auf Genauigkeit, Robustheit und Rechenzeit die quantitativ und qualitativ hochwertigsten Ergebnisse bei der Berechnung von Bewegungsvektoren?*

Die Ergebnisse zeigen, dass klassische Verfahren (Lucas-Kanade und Farnebäck) in den untersuchten Drohnenvideos robustere und genauere Bewegungsvektoren liefern als RAFT. Lucas-Kanade erzielt den geringsten durchschnittlichen Endpunktfehler (ca. 1,29 px), Farnebäck ist ähnlich präzise und erkennt zudem die Richtungsvielfalt in der Crowd-Dynamik am zuverlässigsten. RAFT hat im Drohnenkontext höhere Fehler und ist weniger robust in komplexen Szene. Hinsichtlich Rechenzeit ist Lucas-Kanade als einziges Verfahren echtzeitfähig (ca. 7 ms/ Frame), Farnebäck und RAFT sind deutlich langsamer und für zeitkritische Anwendungen ungeeignet.

Die Analyse liefert klare Empfehlungen: Für Echtzeitanwendungen und begrenzte Ressourcen ist Lucas-Kanade optimal, Farnebäck eignet sich für detaillierte Crowd-Analysen, RAFT kann bei Offline-Auswertung und nach gezieltem Domänen-Fine-Tuning eine Ergänzung sein. Die Arbeit schafft eine fundierte Vergleichsbasis, zeigt methodische Stärken und Grenzen auf und liefert Handlungsempfehlungen für die Auswahl von Optical-Flow-Verfahren in drohnenbasierten Bewegungsszenarien. Der Ausblick betont die Bedeutung größerer Ground-Truth-Datensätze, spezifisches Training von Deep-Learning-Modellen sowie Optimierungen für Echtzeit- und Edge-Anwendungen.

Die Masterarbeit trägt somit zur methodischen und praktischen Weiterentwicklung der Bewegungsanalyse in Drohnenvideos bei und legt die Basis für zukünftige, praxisnahe Anwendungen in der automatisierten Crowd-Überwachung.