

EVIVA – fluggestütztes Beobachtungs- und Analysesystem für Eventmonitoring mit videobasierter Verhaltensanalyse

Alexander ALMER, Roland PERKO, Thomas SCHNABEL,
Lucas PALETTA und Ludwig KASTNER

JOANNEUM RESEARCH, Graz · alexander.almer@joanneum.at

Einleitung

Bei Ereignissen mit großen Menschenansammlungen ermöglicht die umfassende Beobachtung des Bewegungsverhaltens der Personen eine frühzeitige Erkennung kritischer Situationen (überhöhte Menschendichten, Staubbildung, beginnendes Panikverhalten, etc.). Im Projekt EVIVA¹ wurden Methoden für eine luftbasierte Beobachtung auf Basis von Videodaten entwickelt, welche eine computergestützte Analyse potenziell kritischen Bewegungsverhaltens ermöglicht. Die Beobachtung des Verhaltens kann einerseits zur Prävention von Krisensituationen genutzt werden, um Sicherheitskräfte rechtzeitig ordnend in die Gruppendynamik eingreifen zu lassen (z. B. zur Verhinderung von Staubbildung bei Open-Air-Veranstaltungen). Andererseits kann in bereits bestehenden Krisensituationen (z. B. die Evakuierung eines Stadtgebietes) durch frühzeitige Identifikation von Gefahrenzonen und rechtzeitiges Einschreiten von Sicherheitskräften Panikverhalten verhindert werden. Mit den in EVIVA entwickelten Komponenten können große Areale überwacht und bei Detektion einer potenziell gefährlichen Situation spezifische Aufmerksamkeitsalarme am Management- und Führungssystem angezeigt werden. Dieses integrierte Management- und Führungssystem stellt eine wesentliche Erweiterung zum klassischen Einsatzmanagement dar und bietet eine wesentliche Unterstützung bei der Beurteilung von Situationen sowie der gezielten Steuerung von Sicherheitskräften. Wesentliches Ziel war, es den Einsatzzentralen optimale Werkzeuge für die effiziente Unterstützung von Entscheidungsprozessen zur Verfügung zu stellen und damit ein wirkungsvolleres Sicherheitsmanagement bei Großveranstaltungen zu ermöglichen.

Ausgangssituation

Dramatische Beispiele wie Hillsborough² oder auch die Bergisel-Katastrophe³ sowie die Loveparade⁴ Katastrophe in Duisburg zeigen, dass die Gewährleistung der Sicherheit ganz elementar mit der Steuerung der Dynamik der Bewegung von Menschenmassen zusammenhängt. Kritische Situationen werden dabei nur beherrschbar, wenn sie frühzeitig erkannt werden und damit rechtzeitig interveniert werden kann.

¹ <http://www.kiras.at/gefoerderte-projekte/detail/projekt/eviva/>

² Tragödie von Hillsborough (15.04.1989): Fußballstadion – 96 Tote und 730 Verletzte.

³ Bergisel-Katastrophe (1999): Veranstaltung Air&Style – 5 Tote und 5 Pflegefälle.

⁴ Loveparade in Duisburg (24.07.2010): 21 Besucher starben und 541 wurden verletzt.

Für die Überwachung von Personengruppen werden terrestrische Videosysteme bereits umfangreich eingesetzt. Der kombinierte Einsatz von terrestrischen und fluggestützten (Flugzeug, Hubschrauber, UAS) Videosystemen bietet die Möglichkeit einer wesentlichen Effizienzsteigerung im Hinblick auf die Früherkennung von kritischen Situationen, auf rasche Entscheidungen für koordinierte Interventionen sowie der Steuerung dieser Maßnahmen und Beurteilung des Erfolgs. Zentrale Voraussetzung für ein gemeinsames, zielgerichtetes Handeln in Krisensituationen ist ein möglichst umfassendes und objektives Lagebewusstsein und Lageverständnis (gemeinsames Lagebild). Dies setzt eine nutzer- und situationsgerechte Bereitstellung von führungs- und entscheidungsrelevanten Informationen an alle beteiligten Akteure voraus. Wesentliche Erfahrungen konnten bei der Bundespolizeidirektion durch Analysen der Abläufe im Rahmen des Wiener Donauinselfestes gemacht werden. Im Rahmen des Projektes PUKIN⁵ unterstützte Diamond Aircraft Industries (DAI) mit einer DA 42 MPP mit eingerüstetem HDTV-Sensor das Sicherheitsmanagement. Insgesamt waren 2.9 Millionen Besucher in 3 Tagen zu verzeichnen. Die Notwendigkeit einer integrierten und automatisierten Verfahrensweise zur Unterstützung der Entscheidungsträger und Einsatzkräfte wurde durch die vorhandene Informationsflut und die zeitlich kritischen Abläufe bei Entscheidungsprozessen eindeutig dokumentiert.

Über das Verhalten von Menschenmassen in Gefahrensituationen existieren nur wenige Studien und empirische Daten, in denen dynamische Phänomene von Menschenmassen identifiziert und relevante Kriterien abgeleitet wurden (HELBING et al. 2007). Eine wesentliche Zielsetzung ist somit die Identifikation der kritischen Parameter und Werte, die Klassifikation von Situationen sowie die Untersuchung der Zuverlässigkeit der Kriterien anhand weiterer empirisch erfasster Daten. Um kritische Situationen rasch beurteilen bzw. entsprechende Gegenmaßnahmen sofort einleiten zu können, sind aus der Sicht der Endnutzer drei Grundanforderungen wesentlich:

- frühzeitiges Erkennen kritischer Situationen, möglichst bevor diese ihr volles Gefahrenpotenzial entwickeln;
- jederzeitige und vollständige Klarheit über alle verfügbaren Möglichkeiten der Intervention. Dies inkludiert Informationen wie: Aufenthaltsort aller einsetzbaren Sicherheitskräfte, verfügbare, gesperrte und geöffnete Verkehrswege, usw.;
- Monitoring der Wirksamkeit der eingeleiteten Maßnahmen für eine entsprechende Steuerung; Aufbau eines „near real-time“ Informationsflusses unter Einbindung aller relevanten/verantwortlichen Personen bzw. existierender Informationen.

Systemkomponenten

Ziel im Projekt EVIVA war die Entwicklung folgender Komponenten auf Basis eines leistungsfähigen multisensoralen Videosystems (vgl. Abb. 1):

- Informationsanalyse (echtzeitnahe Analysen) mit dem Ziel eines effizienten „Decision Support“.

⁵ PUKIN – Periodische Überwachung kritischer Infrastrukturen; national gefördertes KIRAS Forschungsprojekt – Bundesministerium für Verkehr, Innovation und Technologie (BMVIT); siehe www.kiras.at.

- Verhaltensanalyse und Simulationen zur Abschätzung des Gefahrenpotenzials von Menschengruppen.
- Intelligente Einsatzzentrale (“Location Based Information Management”).
- Unterstützung von Feldoperationen durch den Informationsaustausch mit mobilen Einheiten.

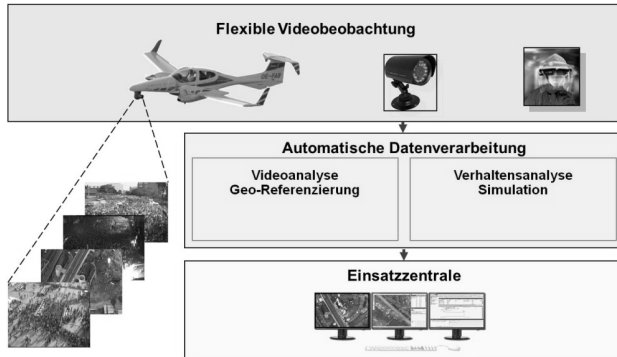


Abb. 1: Systemkomponenten

Systemdemonstration

Die Systemkomponenten wurden im Rahmen von zwei Veranstaltungen getestet. Die für die Entwicklung und Evaluierung der Systemkomponenten benötigten Videodaten aus Flugaufnahmen standen aus organisatorischen und rechtlichen Gründen nur stark beschränkt zur Verfügung. Weitere Videoaufnahmen wurden im Rahmen des Lakeside Festival 2011 von einem Turm durchgeführt, um eine entsprechende Aufnahmekonfiguration zu simulieren. Im Juni 2012 erfolgte im Rahmen des Donauinsselfestes eine umfangreiche Befliegung durch den Projektpartner DAI⁶ mit einem FLIR Star Safire HD-Videosystem⁷. Die Datenanalyse bis zur Ableitung sicherheitsrelevanter Parameter erfordert die Abbildung 2 dargestellten Prozessierungsschritte.

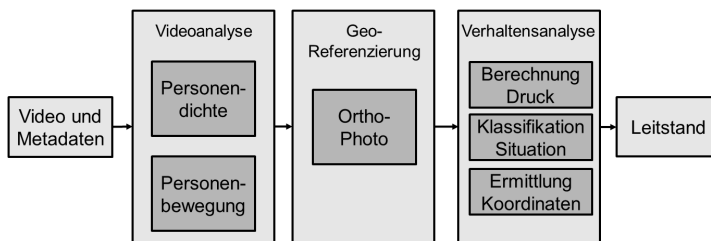


Abb. 2: Prozessierungsschritte des vorgeschlagenen Systems

Abbildung 3 zeigt den Workflow für die Ableitung einer Personendichte aus den Videodaten. Mithilfe von Methoden aus der Bildanalyse werden aus den Bildern Merkmale ex-

⁶ <http://www.diamond-air.at>

⁷ <http://gs.flir.com/surveillance-products/star-safire-hd/hd>

trahiert, z. B. eine Personendektection (FELZENSZWALB et al. 2010). Lernverfahren ermöglichen dann eine iterative Regression dieser Merkmale auf die Personendichte, wobei zum Lernen der benötigten Gewichte manuell annotierte Bilder vorliegen müssen (vgl. (LEMPITSKY et al. 2010)). In der Lernphase werden somit allen Merkmalen Gewichte zugeordnet, welche eine Personendichtenbestimmung auf Testbildern ermöglichen.

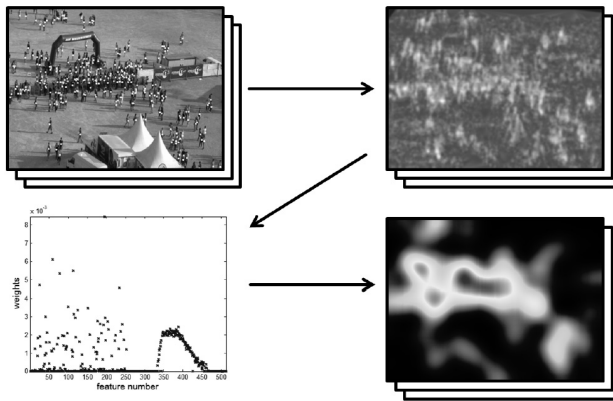


Abb. 3: Workflow für die Bestimmung der Personendichte

Die Bewegungen der Personen werden über den optischen Fluss unter Verwendung von totaler Variation bestimmt (ZACH et al. 2007). Für eine weiterführende Modellierung ist eine Georeferenzierung der abgeleiteten Parameter erforderlich. Abbildung 4 zeigt ein Videobild überlagert mit einer Schätzung für die Personendichte und Bewegungsrichtung sowie dieselbe Information georeferenziert überlagert in einer Google-Earth-Darstellung.



Abb. 4: Darstellung der Personendichte und Abschätzung der Bewegungsrichtung: (links) Ergebnisse in Bildgeometrie und (rechts) georeferenziert in Google Earth überlagert

Abbildung 5 zeigt die errechnete Anzahl der Personen für 7000 Videoframes. Die beiden Kurven (blau und grün) repräsentieren die berechneten Ergebnisse unter Verwendung von unterschiedlichen Prozessierungsparametern, während die Kreise in rot die manuell bestimmten Personenanzahlen darstellen. Die mittlere erzielte Genauigkeit liegt in den beiden Szenarien bei 4 % bzw. 9 %.

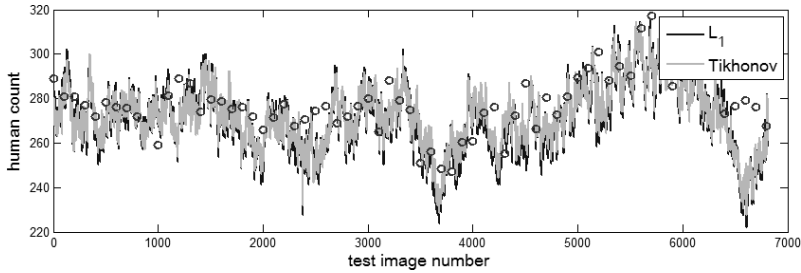


Abb. 5: Genauigkeit der Personenzählung für das Szenario Lakeside

Leitstand

Die Einsatzzentrale als zentraler Informationsknoten, ermöglicht einen nutzer- und situationsorientierten Zugang zu relevanter Information für eine optimale Unterstützung des Sicherheitsmanagements von Großveranstaltungen. Die Bereitstellung eines aktuellen und umfassenden Lagebildes ist dabei oberstes Ziel, wobei gleichzeitig größtmögliche Übersichtlichkeit als auch ein maximaler Detaillierungsgrad von Informationen angestrebt wird. Im Rahmen von EVIVA wurden neben den luftgestützten und terrestrischen Sensoren auch mobile Einheiten als Sensoren genutzt, um das Lagebild in der Einsatzzentrale mit aktuellen Informationen zu versorgen. Als Informationssysteme für mobile Einheiten wurden Smartphones, Tablet-PCs und speziell für den rauen Outdoor-Einsatz entwickelte Geräte verwendet. In der Einsatzzentrale wurden alle Informationen georeferenziert als Layer über einer Basiskarte dargestellt, wozu aktuelle Luftbilder, die automatisch errechneten gefährlichen Bereiche, eventspezifische Karten (die z. B. alle Bühnenaufbauten enthalten) genauso wie aktuelle Positionen aller mobilen Einheiten, gehörten. Die mobilen Einheiten konnten zusätzlich zur Sprachkommunikation auch Bild- und Textnachrichten an die Einsatzzentrale schicken und somit Vorfälle melden, welche grafisch in das Lagebild eingebunden wurden.



Abb. 6: Darstellung des Lagebildes in der Leitzentrale mit überlagerten Layern (Basiskarte, Luftbild, gefährliche Bereiche, Bühnenaufbauten und eventspezifischer Raster)

Problemstellungen und Ausblick

Ein wesentliches Problem im Projektablauf war die Verfügbarkeit des Aufnahmesensorsystems. Aus rechtlichen sowie organisatorischen Gründen war dieses im geplanten Umfang nicht gegeben, was es erforderlich machte, unterschiedliche flugzeuggetragene Videosysteme zu testen sowie mit terrestrischen Videosystemen entsprechende Aufnahmesituationen zu simulieren. Lizenzrechtliche Probleme seitens des Anbieters führten dazu, dass der Zugang zur Kamerasteuerung und zu den Daten der Geosensorik (GPS, IMU) nicht gegeben war. Dies führte im Projekt zu einer stark reduzierten Datensituation für Entwicklungen und Tests im Bereich der Geoprozessierung sowie Video- und Verhaltensanalyse und der Echtzeit-Datenprozessierung. Daher konnte keine gesicherte Aussage über die Leistungsfähigkeit des angestrebten Gesamtsystems auf Basis eines Videoaufnahmesystems gemacht werden.

Die im Rahmen des Projektes erarbeiteten Ergebnisse seitens der Bildanalyse haben gezeigt, dass eine prinzipielle Lösbarkeit der Problemstellung des Erkennens von sich bewegenden Menschengruppen bzw. die Personendichteschätzungen machbar ist, jedoch immer eine Schätzung bleibt. Weiters müssten die entwickelten Algorithmen mit weiteren Videosequenzen getestet und an unterschiedliche Aufnahmesituationen angepasst werden. Mit der Möglichkeit der gezielten Steuerung der Aufnahmesituation können zusätzlich die Rahmenbedingungen optimiert und die Ergebnisse der Bildanalyse deutlich verbessert werden. Durch die schrägen Bildaufnahmen ergeben sich auch deutliche Herausforderungen, da Personen teilweise verdeckt sind bzw. im unebenen Gelände (Brücken, Aufbauten etc.) die Georeferenzierung hochgenaue Oberflächenmodelle erfordert. Speziell auf diese weiterführenden Entwicklungsschwerpunkte konnte aufgrund der benötigten Ressourcen für das Austesten der unterschiedlichen Sensorsysteme nur wenig Augenmerk gelegt werden. Als finaler Schritt ist auch noch die Optimierung in Bezug auf die Analysezeit zu nennen, welche mit finalen Algorithmen konkret für eine echtzeitnahe Ausführung durchzuführen ist.

Literatur

- FELZENSZWALB, P. F., GIRSHICK, R. B., MCALLESTER, D., RAMANAN, D. (2010), Object detection with discriminatively trained part based models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32 (9), 1627-1645.
- HELBING, D., JOHANSSON, A. & AL-ABIDEEN., H. Z. (2007), The Dynamics of Crowd Disasters: An Empirical Study. *Phys. Rev. E* 75, 046109.
- LEMPITSKY, V. & ZISSERMAN, A. (2010), Learning to count objects in images. *Advances in Neural Information Processing Systems (NIPS)*, 23, 1324-1332.
- ZACH, C., POCK, T. & BISCHOF, H. (2007), A duality based approach for realtime TV-L1 optical flow. *Symposium on Pattern Recognition (DAGM)*, Heidelberg, Germany, 214-223.