

# Multisensor-Datenfusion zur Echtzeitlage-visualisierung und Kurzfristprognose bei Großevents

Armin KÖFLER<sup>1</sup>, Thomas SCHNABEL<sup>1</sup>, Roland PERKO<sup>1</sup>,  
Alexander ALMER<sup>1</sup> und Helmut SCHROM-FEIERTAG<sup>2</sup>

<sup>1</sup>JOANNEUM RESEARCH Forschungsges.mbH, Graz · armin.koefler@joanneum.at

<sup>2</sup>AIT Austrian Institute of Technology GmbH, Wien

## Zusammenfassung

Bei Massenveranstaltungen kommt es aus sicherheitstechnischer Sicht immer wieder zu kritischen Situationen und Gefährdung von Personen aufgrund fehlender genauer Informationen bezüglich der Verteilung von Menschenmassen. Eine permanente flächendeckende Überwachung mit Sensoren ist allerdings aus mehreren Gründen problematisch. In dieser Arbeit wird ein hybrider Ansatz gezeigt, in welchem unterschiedliche Sensoren verwendet und durch die intelligente Fusion dieser Daten die Grundlage für Bewegungsstromanalysen und Kurzfristprognosen bildet. Weiterhin wird erläutert, wie diese Daten in der Echtzeitvisualisierung verwendet wurden, welche realen Tests durchgeführt wurden sowie welche Aspekte beim Einsatz beachtet werden müssen.

## Einleitung

Bei Massenveranstaltungen kommt es immer wieder zur Gefährdung von Personen, in Fällen wie am Bergisel am 4.12.1999 oder bei der Love-Parade am 24.7.2010 sogar zu Todesfällen und Schäden an öffentlichen Infrastrukturen. Trotz umfangreicher Planung von Sicherheitskonzepten und der behördlichen Genehmigung anhand genauer Personenzahlgrenzen und dynamischer Entwicklungen treten immer wieder tragische Zwischenfälle auf. Der Grund ist oftmals das Fehlen von genauen Informationen über die aktuelle Personenzahl, Dichte und die räumliche Verteilung am Veranstaltungsgelände. Eine ständige flächendeckende Überwachung mit Sensoren ist meist in Hinblick auf Kosten und Aufwand nicht durchführbar.

Das KIRAS-Projekt „EN MASSE“ fokussiert daher darauf, eine hybride Methode zu entwickeln, die unterschiedliche Sensoren aus unterschiedlichen Einsatzgebieten für eine zuverlässigere Erfassung von Besucherströmen zusammenführt, diese auszuwerten, Prognosen der kommenden Entwicklungen zu erstellen und beides für Sicherheitskräfte geografisch basiert zu visualisieren.

## Überblick

Aufgrund der Problemstellungen oftmals zu spät erkannter Entwicklungen kritischer Situationen durch die Verteilung von Besuchern von Großevents wurde im Rahmen von KIRAS das Projekt EN MASSE durchgeführt. Dieses spezialisierte sich auf die Schwerpunkte

multisensorielle Erfassung von Personenbewegungen, die modellbasierte Fusion der einzelnen Datenquellen und deren Analyse sowie die Kurzfristprognose und echtzeitfähige, geobasierte Lagebilddarstellung. Die Erfahrungen von JOANNEUM RESEARCH (JR) und AIT, welche bereits in der Vergangenheit Forschung in diesen Bereichen durchgeführt hatten, wurden durch die aktive Teilnahme von Eventveranstaltern sowie zuständigen Sicherheitsexperten unterstützt.

## Sensordatenaufnahme und Analyse

Speziell für Großevents relevant war die Nutzung unterschiedlicher Sensoren, welche entsprechend ihrer Eigenschaften an neuralgischen Punkten aufgestellt wurden und in optische und drahtlose Sensoren eingeteilt werden.

Bei den optischen handelte es sich um Videokameras (AXIS P3364), mit welchen Bereiche automatisch auf den Befüllungsgrad hin analysiert wurden. Über die abgedeckte Fläche und das Tracking der an- und wegströmenden Personen konnte auch die Personenanzahl geschätzt werden. Zusätzlich dazu wurde die in Vorgängerprojekten von JR entwickelten Algorithmen zur Erkennung von Einzelpersonen (PERKO et al. 2013) weiterentwickelt, deren Ergebnisse in weiterer Folge in das Datafusionsframework und später zur Personenstromanalyse sowie Dichtebestimmung der Menschenmassen einfließen. Zum anderen wurden mit laserbasierten Zählsensoren (LASE) an definierten Eingängen Personen sowie deren Bewegungsrichtung gezählt.

Bei den drahtlosen Sensoren handelte es sich solche, die Mobiltelefone erfassen, deren Zuordnung zu den vorhandenen Mobilfunkzellen gegeben war und somit auf die Position des Gerätes selbst zurückgeschlossen werden konnte. Als zweite drahtlose Variante wurde Bluetooth verwendet, um eingeschaltete Geräte innerhalb des einstellbaren Empfangsradius zu registrieren und an anderer Stelle wieder zu erkennen. Bis auf die Ortung mittels Mobilfunk boten alle eingesetzten Sensoren mit der erreichbaren Genauigkeit die Möglichkeit daraus das Bewegungsverhalten ableiten zu können.

## Multisensor-Datenfusion

Für eine gemeinsame Auswertung, als Input für die Simulation und Kurzfristprognose, sowie die schließlich benötigte Lagebildgenerierung wurde ein flexibles, mehrere Ebenen umfassendes Data-Fusion-Framework, auf Basis des .NET-Frameworks entwickelt. Es erlaubt die Nutzung unterschiedlichster Sensortypen mit unterschiedlichen Datenformaten, Genauigkeiten und Zuverlässigkeit als Eingangsdaten, welche modellbasiert für die angestrebten Ergebnisse herangezogen werden.

Angelehnt an (CARVALHO et al. 2003) wurde das Framework erstellt, das in sieben Schichten mit verschiedenen Aufgaben eine Datenfusion ermöglicht. Abstrahiert lassen sich die Schichten wie folgt beschreiben:

1. **Sensorschicht.** Sie repräsentiert die Inputschnittstelle. Für jeden unterstützten Sensor wird im Framework eine den physischen Sensor beschreibende Sensorkomponente definiert welche Daten für folgende Schichten erfasst.

2. Vorverarbeitungsschicht. Sofern nötig werden von der Sensorschicht erhaltene Daten behandelt, sodass die in nachfolgenden Schichten verarbeitet werden können. Beispielhafte Aufgaben sind Fehlerdetektion oder Filterung.
3. Low-Level-Datenfusion. Werden aus der Vorverarbeitungsschicht erhaltene Daten vor der Datenanalyse (Schicht 4) fusioniert, so passiert das in der Low-Level-Datenfusion. Eine beispielhafte Aufgabe dieser Schicht ist die Aggregation von Daten von Zählsensoren über ein Zeitintervall.
4. Datenanalyse. Daten aus den vorhergehenden Schichten werden algorithmisch ausgewertet und höherwertige Information transformiert.
5. High-Level-Datenfusion. Aus der Datenanalyse stammende Information werden fusioniert.
6. Interpretation. Fusionierte Information wird nach dem Anwendungsfall entsprechenden Kriterien interpretiert. Weist fusionierte Information beispielsweise auf einen hohen Zustrom von Menschen auf ein Areal hin, so wird die Information auf eine Warnstufe reduziert.
7. Entscheidungsschicht. Erhaltene Interpretationen dienen als Input für die Generierung von Entscheidungen. Erkennt das System beispielsweise, dass sich zwei große Menschengruppen aus zwei Richtungen einem klein dimensionierten Areal nähern, so kann eine dringliche Warnung an einen Entscheidungsträger geschickt werden.

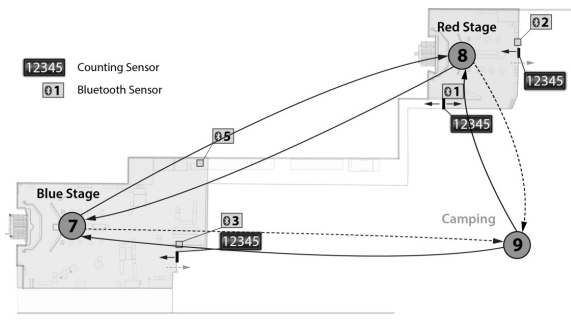
Eine konkrete Anwendung definiert die Ausprägung des Frameworks durch eingesetzte Sensoren, den gewünschten Ergebnissen und anderen Eigenschaften. Die Ausprägung besteht aus einer hierarchischen Kombination der Schichten, wobei einzelne Schichten fehlen (Daten/Informationen überspringen die Schicht) oder mehrfach parallel vorkommen können (unterschiedliche Sensoren werden gleichzeitig verwendet). Ein grafisches Planungswerkzeug dient der Erzeugung konkreter Ausprägungen.

Für die Nutzung im Rahmen von EN MASSE wurden dabei zuerst die einzelnen Sensortypen mit ihren Parametern festgesetzt und darauf basierend das Analysemodell erstellt. Hierbei flossen vor allem die Gewichtungswerte der Sensoren ein, um eine normalisierte Datensituation zu schaffen. Mit Bluetooth Scannern erfasste Mobilgeräte deckten beispielsweise nur 2 % der Besuchergesamtanzahl ab. Die Ergebnisse der Videoanalyse boten für die verwendeten extrahierten Einzelbilder eine Erkennungsrate von ~95 %, welche durch manuell durchgeführte Referenzzählungen für unterschiedliche Testaufnahmen ermittelt wurde (PERKO et al. 2014).

## Situationsvisualisierung

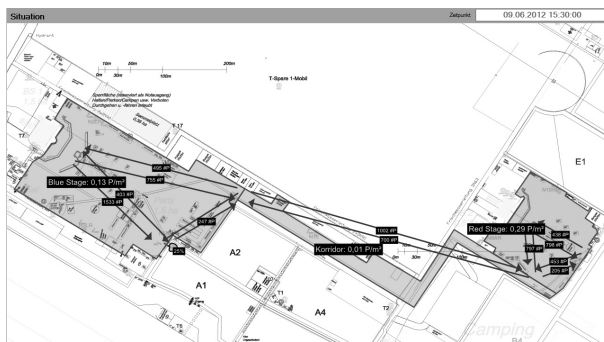
Zur Verwendung der Analysedaten durch die Bedarfsträger wurde eine Applikation zur Visualisierung eines Lagebildes erstellt. Diese nutzt zum einen die kontinuierlich durch die Analyseworkflows erhaltenen Daten und zum anderen die Ergebnisse der Kurzfristprognose, welche die prognostizierte Situation in 15 bzw. 30 Minuten bereitstellte. Der Nutzer hat dabei auch die Möglichkeit sich Situationen zu unterschiedlichen Zeitpunkten anzeigen zu lassen und somit auch einen Vergleich zwischen den prognostizierten Daten und der realen durch die Sensoren gemessene Datensituation herzustellen. Die Darstellung umfasst dabei einerseits Bereiche, welche durch ihre geografische Ausdehnungen sowie Ein- und Ausgänge definiert wurden, andererseits die Übergänge zwischen den einzelnen Sensoren. Somit konnten auch Bewegungsgeschwindigkeiten ermittelt werden. Eine farb-kodierte Darstellung von Personendichte bzw. Geschwindigkeit erlauben rasches erfassen

der Lage. Detaillierte Informationen können bei Bedarf abgerufen werden. Eine schematische Darstellung eines Geländes und der Sensorpositionen ist unter Abbildung 1 ersichtlich.



**Abb. 1:**  
Schematische Darstellung von Gelände und Sensoren

In Abbildung 2 wird die Applikation dargestellt, die in der Lage ist, basierend auf den erfassten Sensordaten fusionierte Auswertungen zu erstellen und diese sowohl in der aktuellen Situation als auch in einer Prognose für wahlweise 15 oder 30 Minuten darzustellen. Als Datengrundlage für den Screenshot dient die erfasste Situation beim NOVAROCK Festival 2012.



**Abb. 2:**  
Lagebild NOVAROCK Festival 2012

Auf dem Lagebild ersichtlich sind die drei Regionen „Blue Stage“ (links), „Korridor“ (Mitte) und „Red Stage“ (rechts), für die sowohl farblich als auch textuell (grüne Schrift) der aktuelle Füllstand basierend auf in den letzten 15 Minuten aggregierten Sensordaten und dem Zustand der vorhergehenden Epoche darstellt. Die Textinformation den gibt den Füllstand als durchschnittliche Anzahl von Personen pro m<sup>2</sup> an. Die farbliche Darstellung der Regionsfläche korrespondierend anhand von Schwellwerten den Kategorien „Niedriger Füllstand“ (grün), „Mittlerer Füllstand“ (gelb), „Hoher Füllstand“ (rötlich) und „Kritischer Füllstand“ (rot). Die Intensität der Färbung passt sich graduell, sodass ein Operator intuitiv Änderungen innerhalb einer Kategorie wahrnehmen kann.

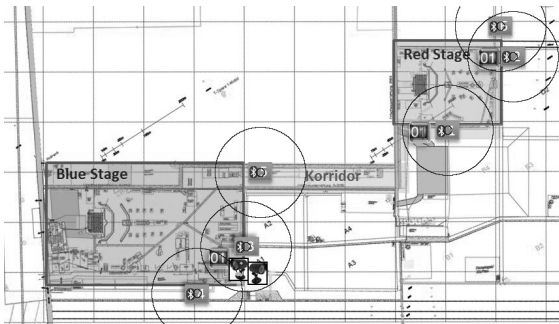
Gelb beschriftete rote Pfeile beschreiben die Anzahl der Personen, die sich in diesem Zeitraum von einem Knoten zu einem anderen bewegt haben. Vor dem Eingang der „Blue Stage“, welcher rechts unten an der „Blue Stange“ zu finden ist, wird mittels Videoauswertung die prozentuelle Ausfüllung des Zugangs mit Personen ersichtlich. Auch diese Füllstandsanzeige wird basierend auf Schwellwerten farblich veränderlich angezeigt.

## Kurzfristprognose

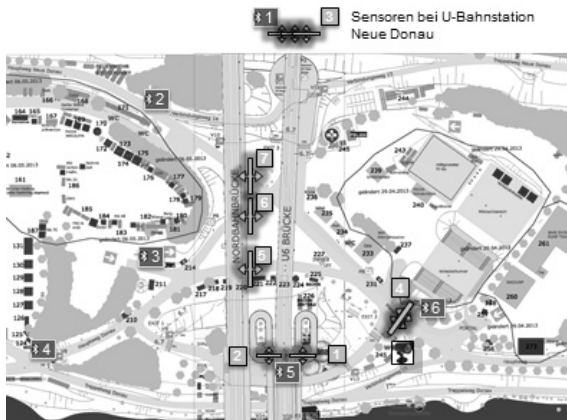
Basierend auf den fusionierten Daten über den Istzustand wurden einerseits die Füllstände definierter Bereiche sowie die Bewegung von Personengruppen zwischen definierten Knoten für auf 15 Minuten aggregierte Epochen bestimmt. Mittels linearer Interpolation wurden für jeweils nachfolgenden beiden Epochen Prognosen erstellt und im Lagebild dargestellt.

## Reale Einsätze

Im Rahmen von EN MASSE konnten zwei Feldtests durchgeführt werden. Sowohl am NOVAROCK 2012 sowie auch am Donauinsselfest 2013 wurden die unterschiedlichen Sensoren installiert und Daten aufgenommen, welche in Folge analysiert und verifiziert werden konnten. Bei beiden Veranstaltungen wurden Zählsensoren (LASE), Bluetooth-Scanner und Kameras als Datenquellen eingesetzt.



**Abb. 3:**  
Lageplan NOVAROCK Festival 2012 inkl. der Sensorverteilung



**Abb. 4:**  
Sensorverteilung Donauinsselfest 2013

## Schlussfolgerungen und Ausblick

In dieser Arbeit wurden die Inhalte des KIRAS-Projekts EN MASSE mit dem Fokus auf die Fusion von multisensoriellen Daten vorgestellt sowie die Erkenntnisse und weiteren

Entwicklungsanforderungen diskutiert. Die durchgeführten Tests und Demonstrationen im Rahmen von zwei Festivals zeigten dabei die unterschiedlichen Qualitäten aber auch die Zuverlässigkeit der eingesetzten Datenquellen. Dabei kristallisierte sich heraus, dass speziell die drahtlosen Zählmethoden noch deutlich verbessert bzw. erweitert werden müssen. Beispielsweise würde die Erweiterung um WLAN oder iBeacon basierte Positionierungsmethoden eine wesentlich detaillierte Aussagekraft liefern können. Weiters ist es auch notwendig, die Erfassung mit optischen Mitteln zu verbessern, was die vor allem die Geschwindigkeit der Zählung betrifft, um diese echtzeitnahe durchführen zu können. Die Wahl der Sensoren für eine Veranstaltung hängt auch von der Veranstaltungsart ab. Besucher einer Veranstaltung im urbanen Gebiet können mittels derer Mobiltelefone potenziell leichter erfasst werden als bei mehrtägigen Festivals am Land, da die Geräte regelmäßiger geladen werden können. Ähnliches gilt für die Verwendung von WLAN bzw. Bluetooth. Hier müssen Anreize (z. B. durch Apps) geschaffen werden, diese Technologien am mobilen Gerät aktiv zu haben. Auch müssen dabei noch mehrere Versuche für eine fundierte statistische Aussage in Bezug auf die Zuverlässigkeit der gemessenen Daten zu erhalten.

Weiters ergeben sich Notwendigkeiten in der Einbeziehung weiterer Informationsquellen, welche Änderungen grundlegender Zustände des beobachteten Areals signalisieren. Werden beispielsweise temporär Absperungen oder Zäune geöffnet, so betreten oder verlassen Personen ein Areal, ohne dass diese Änderungen sensorisch erfasst werden können.

Bei wiederkehrenden Veranstaltungen können die erfassten und ausgewerteten Daten der Planung für das nächste Mal dienlich sein, indem sie zum Beispiel für eine höher zu erwartende Besucheranzahl Simulationen für die Planung der Dimensionierung von Ein- und Ausgangsportalen oder Durchgängen verwendet werden können.

Datenschutz ergab sich als nicht zu vernachlässigendem Faktor. Die Gewährleistung der Anonymität von Veranstaltungsbesucher muss gegeben sein und überzeugend kommuniziert werden. Aufgrund schlechter Publicity konnten Versuche nicht wie geplant durchgeführt werden, da der Veranstalter einen Besucherschwund fürchtete.

## Literatur

- CARVALHO, H. S., HEINZELMAN, W. B., MURPHY, A. L. & COELHO, C. J. N. (2003), A general data fusion architecture. In: Proceedings of the Sixth International Conference of Information Fusion, 2003. Cairns, 1465-1472.
- PERKO, R., SCHNABEL, T., FRITZ, G., ALMER, A. & PALETTA, L. (2013), Airborne based high performance crowd monitoring for security applications. In: KÄMÄRÄINEN, J.-K. & KOSKELA, M. (Eds.), Scandinavian Conference on Image Analysis (SCIA), Espoo, Finland: Springer LNCS, 2013, 7944 (18), 664-674.
- PERKO, R., SCHNABEL, T., ALMER, A. & PALETTA, L. (2014), Towards View Invariant Person Counting and Crowd Density Estimation for Remote Vision-Based Services, 23rd International Electrotechnical and Computer Science Conference, Portorož, Slovenia. Submitted for review.