

# Crowd context recognition with wearable sensors

A dissertation submitted to

ETH ZURICH

for the degree of  
Doctor of Sciences

presented by

MARTIN WIRZ

MSc ETH

Date of birth January 13th, 1984  
Citizen of Neuendorf, Switzerland

accepted on the recommendation of

Prof. Dr. Gerhard Tröster, examiner  
Prof. Dr. Dirk Helbing, co-examiner

2013

# Abstract

Crowd behaviors are the result of self-organizing phenomena governed by the local interactions among the involved individuals. Due to the limited perception range of the human sensory system and the complexity such patterns embrace, the involved individuals often fail to detect the existence of a crowd behavior, do not recognize their involvement and cannot perceive the extent and characteristics of the pattern. Awareness of occurring crowd behaviors, however, is critical in many every-day situations.

This thesis aims at developing and studying wearable computing-based approaches to infer context information of human crowds. Considered technology consists of sensor modalities and communication capabilities commonly available in state-of-the-art smartphones as such are widely deployed in our society.

There is a need for real-time awareness of the context of a crowd for successful crowd management during mass gatherings. This enables a situational assessment which helps to detect critical crowd constellations at an early stage. This thesis investigates the addressed research questions in the perspective of crowd context recognition for providing support during mass gatherings. Towards the end of the thesis, the scope is also broadened up and the potential of crowd context recognition is evaluated in two other use cases.

This thesis stimulated the development of a data collection system able to record behavior information with mobile devices from hundreds of users simultaneously. The system processes the collected information centrally to infer crowd context information. The system can be used for participatory sensing campaigns where users voluntarily participate and share their context information. The system was successfully deployed during a city-wide mass gathering to collect location trajectories from over 800 attendees. The collected data was used for online identification and visualization of relevant crowd context information. The usefulness of the inferred information and the expressiveness of the generated visualizations were confirmed by police forces accessing the system during the event.

A participatory sensing strategy for crowd context recognition implies a major challenge: Only a subset of all attendees of a mass gathering is going to share sensor information. This affects the explanatory power of the inferred information. This thesis addresses this challenge on two fronts:

- Ensuring a large user base is important to obtain significant crowd context information. Thus, the user acceptance to share privacy sensitive context information was investigated and incentivization strategies to foster participation were derived. 98 visitors were questioned during a mass gathering to learn about their privacy concerns for sharing context information and to understand motivational factors for information sharing. This has led to design considerations for our participatory sensing system for crowd monitoring during mass gatherings.

- A methodology which allows to infer a crowd density estimation by tracking only a subset of the crowd members has been developed. The method identifies a mapping between the distribution of pedestrians which are sharing their location and the crowd density. A correlation coefficient between the crowd density estimates and ground truth information of up to  $r = 0.83$  and a crowd density estimation error of  $\sigma = 0.34 \text{ m}^{-2}$  was obtained, depending on the application. These crowd density estimates can be classified into the four service classes *free flow*, *restricted flow*, *dense flow* and *jammed flow* with an accuracy of 0.84 (precision = 0.67, recall = 0.49).

This thesis further introduced and evaluated methods to detect the *flocking* behavior of pedestrians to detect groups of people walking together. A method was developed which relies on pedestrians' location trajectories obtained from GPS sensors. An evaluation with data collected during an experiment involving 10 subjects showed that the correct number of flocks can be detected with an accuracy of 0.79 and with an accuracy of 0.91, subjects were correctly assigned to the correct flock. As location information is not always present, this thesis also investigated an approach to infer if people are walking together by correlation signal features derived from wearable acceleration sensors. An evaluation with a data set including recordings of 10 subjects confirmed that pedestrians walking in groups possess a higher behavior similarity than pedestrians not walking in the same group and that this information can be captured with body-worn acceleration sensors.

Besides crowd control during mass gatherings, there is a plethora of other socio-technical system where the recognition of crowd context information can be useful. This thesis further demonstrates the usefulness of crowd context recognition on two other use cases:

- A system has been deployed and evaluated that can help paraglider pilots to extend their flights by detecting locations with good thermal characteristics. These locations are inferred by sensing the emergent collective behaviors among pilots. The technical feasibility has been shown with test flights and the usefulness has been confirmed through expert interviews.
- A system has been developed to increase a crowd's mood in a dance club by allowing attendees of a party to collaboratively influence the music selection process. An unobtrusive, implicit feedback mechanism by sensing the dance engagement using wearable acceleration sensors has been introduced. The work verified that music can be selected such that the generated song sequence matches the taste of the crowd and demonstrated the feasibility of inferring an individual's dance engagement using smartphones' integrated accelerometers by achieving a detection accuracy of 0.89.

The conclusion of this thesis is that crowd context information can be inferred by means of wearable technology. Following a participatory sensing approach is promising by addressing the challenges of data sparsity. This issue can be mitigated by developing reliable calibration and interpolation models and by cleverly designing incentives. Hereby, providing relevant information as a reward for sharing con-

text information can stimulate participation. Such information is of great value to users if it is derived from the behavior (or knowledge) of the collective and cannot easily be assessed by an individual independently and thus helps to extend their sensory range.

# Zusammenfassung

Massenverhalten in Menschenmengen sind das Resultat selbst-organisierender Phänomene. Solche Phänomene können durch lokale Interaktionen zwischen den beteiligten Personen entstehen. Aufgrund des begrenzten Wahrnehmungsbereiches des menschlichen sensorischen Systems und der Komplexität solcher Verhaltensmuster, können die involvierten Personen oft nicht erkennen, dass ein solches Massenverhalten auftritt. Zudem können die involvierten Personen ihre Beteiligung nicht feststellen und sind nicht in der Lage, das ganze Ausmass wahrzunehmen. Zu wissen, dass ein Massenverhalten existiert, ist jedoch in zahlreichen Situationen von grosser Wichtigkeit.

Ziel dieser Arbeit ist die Entwicklung und Evaluation von Methoden zur Erkennung solcher Massenverhalten mittels tragbarer Computertechnologie. Dazu sollen Sensoren und Kommunikationsmöglichkeiten verwendet werden, die bereits heute in modernen Mobiltelefonen integriert sind. Dies, weil solche Geräte in unserer Gesellschaft bereits eine weite Verbreitung gefunden haben.

Es ist besonders während Massenveranstaltungen wichtig, den Kontext einer Menschenmenge erfassen zu können. Dies hilft, Menschenströme beobachten zu können und kann bei einer Lagebeurteilung helfen. Eine solche Lagebeurteilung unterstützt die Früherkennung kritische Massenkonstellationen.

Diese Arbeit entwickelt Methoden zur Echtzeiterkennung des Kontextes einer Menschenmenge während Massenveranstaltungen. Zusätzlich wird in dieser Arbeit auch die Anwendbarkeit von Kontexterkenennung in Menschenmengen an zwei weiteren Anwendungsfällen gezeigt.

Ein Teil dieser Arbeit war die Entwicklung eines Datenerhebungssystems, welches Sensorwerte von Mobiltelefonen hunderter Benutzer gleichzeitig erfassen kann. Zudem erlaubt das System die gesammelten Informationen zentral zu verarbeiten, um den Kontext einer Menschenmenge zu erfassen. Das System kann für partizipative Datenerhebungen verwendet werden, bei welchen die Benutzer aus eigenem Willen und Interesse ihre Kontextdaten bereitstellen.

Das System wurde erfolgreich während einer Massenveranstaltung eingesetzt. Dabei wurden Positionsdaten von über 800 Besuchern gesammelt. Die gesammelten Daten wurden für die Echtzeit-Identifikation und Visualisierung von relevanten Kontext-Informationen der Menschenmenge verwendet. Während der Veranstaltung hatte die Polizei Zugriff auf diese Visualisierungen. Nach dem Anlass konnten die Nützlichkeit der abgeleiteten Informationen sowie die Aussagekraft der generierten Visualisierungen von der Polizei bestätigt werden.

Ein partizipatives Datenerhebungssystem mit dem Ziel Massenverhalten zu erkennen birgt eine grosse Herausforderung: Nur ein Teil aller Besucher wird sich bereit erklären, Sensordaten beizusteuern. Dies hat Auswirkungen auf die Aussagekraft der abgeleiteten Informationen. Diese Arbeit befasst sich mit dieser Herausforderung und behandelt zwei Punkte:

- Die Sicherstellung einer grossen Nutzerbasis ist wichtig um eine signifikante Menge an Datenwerten zu erhalten. In einer Studie wurden 98 Besucher einer Massenveranstaltung bezüglich ihren Bedenken hinsichtlich der Privatsphäre beim Teilen kontextsensitiver Daten befragt. Zusätzlich wurde nach Motivationsfaktoren und Anreizmechanismen für das Bereitstellen solcher kontextsensitiver Daten gefragt. Dies führte zu Entwurfsentscheidungen für unser partizipatives Datenerhebungssystem.
- Eine Methodik wurde entwickelt, um eine Dichteschätzung der Menschenmenge zu erhalten, wenn die Position lediglich eines Teils aller Besucher abrufbar ist. Das Verfahren erzeugt eine Abbildung zwischen der Verteilung jener Personen, welche ihre Position mitteilen und der Dichte der gesamten Menschenmenge. Eine Korrelation zwischen den geschätzten Dichten der Menschenmenge und wahren Messungen von  $r = 0.83$  wurde ermittelt. Einen Schätzungsfehler von  $\sigma = 0.34 \text{ m}^{-2}$  wurde erreicht. Diese Dichteschätzung der Menschenmenge kann mit einer Genauigkeit von 0.84 (Präzision = 0,67, Sensitivität = 0,49) in die vier Klassen *uneingeschränkter Menschenfluss*, *eingeschränkter Menschenfluss*, *dichter Menschenfluss* und *verstopfter Menschenfluss* eingeteilt werden.

Des Weiteren werden in dieser Arbeit Methoden zur Detektion des Herdenverhalten unter Fussgängern entwickelt. Diese Methoden helfen, Gruppen von Menschen, die sich zusammen bewegen, zu erkennen. Eine Methode wurde entwickelt, welche Gruppen anhand von Positionstrajektorien via GPS Sensoren ermittelt. Ein Experiment, durchgeführt mit 10 Probanden, lieferte Daten für die Evaluation. Mit einer Genauigkeit von 0.79 konnte die korrekte Anzahl der Gruppen ermittelt werden und mit einer Genauigkeit von 0.91 konnten die Probanden zur korrekten Gruppe zugewiesen werden. Da Positionsdaten nicht immer eruierbar sind, wurde in dieser Arbeit zusätzlich ein Ansatz entwickelt, welcher über die Korrelation von Signaleigenschaften tragbarer Beschleunigungssensoren ermittelt, ob Fussgänger zusammen laufen oder nicht. Eine Auswertung mit Daten von 10 Probanden wurde durchgeführt. Diese zeigt, dass Fussgänger welche sich in einer Gruppen fortbewegen, eine höhere Ähnlichkeit in ihrem Verhaltensmuster haben als solche, die sich nicht zusammen fortbewegen. Zudem wurde gezeigt, dass diese Eigenschaft mit am Körper getragenen Beschleunigungssensoren erfasst werden kann.

Neben dem Beobachten von Besucherströmen während Massenveranstaltungen gibt es eine Vielzahl weiterer Anwendungsbereichen, bei welchen die Kontexterkennung in Menschenmengen nützlich sein kann. Diese Arbeit präsentiert zwei weiteren Anwendungsfälle:

- Ein System wurde implementiert und evaluiert, mit welchem Gleitschirmpiloten ihre Flüge verlängern können. Das System kann Gebiete mit idealen thermischen Bedingungen erkennen. Diese Gebiete werden durch das Flugverhalten der anderer Gleitschirmpiloten ermittelt. Die technische Machbarkeit wurde mit Testflüge gezeigt und die Nützlichkeit wurde durch Interviews mit Experten bestätigt.

- Es wurde ein System entwickelt, um die Stimmung in einem Tanzclub zu erhöhen. Dabei können die Besucher die gespielte Musik selber mitbestimmen. Mittels tragbaren Beschleunigungssensoren wird die Tanzaktivität der Besucher ermittelt. Diese Aktivität wird als eine unaufdringliche, implizite Stimmabgabe zur aktuell gespielten Musik verstanden. Je grösser der Anteil an tanzenden Personen ist, desto grösser ist die Zustimmung zur aktuell gespielten Musik. Ist die Zustimmung gross genug, wird das System den Musikstil beim nächsten Lied beibehalten. Ansonsten wird die Musikrichtung gewechselt. Die Arbeit wies nach, dass damit Musik ausgewählt werden kann, sodass die erzeugte Musikabfolge den Vorlieben der Teilnehmern entspricht. Zudem demonstriert die Arbeit die Möglichkeit, die Tanzaktivität mittels Beschleunigungssensoren in Mobiltelefonen mit einer Genauigkeit von 0.89 bestimmen zu können.

Das Fazit dieser Arbeit ist, dass Eigenschaften von Massenverhalten mittels tragbaren Sensoren gemessen werden können. Hierbei wurde gezeigt, dass ein partizipatives Datenerhebungssystem ein vielversprechender Ansatz ist. Jedoch muss dem daraus resultierenden Mangel an Daten spezielle Beachtung geschenkt werden. Diesem Problem kann einerseits durch die Entwicklung zuverlässiger Kalibrierungs- und Interpolationsmodellen entgegen gewirkt werden. Andererseits kann eine geschickte Gestaltung von Anreizmechanismen die Teilnahme erhöhen. Dabei kann durch Bereitstellen relevanter Informationen als Belohnung für den Austausch von Kontextinformationen die Teilnahme stimuliert werden. Solche Informationen sind für die Benutzer von grossem Wert, wenn sie nicht ohne weiteres durch individuelle Beurteilung der Lage eruierbar sind, sich aber vom Verhalten (oder Wissen) des Kollektives ableiten lassen können. Dies ermöglicht die Erweiterung des sensorischen Bereichs des Individuums.