

Classification of audio sources using ad-hoc microphone arrays

(deutscher Titel: Klassifikation von Schallquellen unter Verwendung von ad-hoc Mikrofon-Arrays)

von Sebastian Gergen

Die automatische Klassifikation von akustischen Quellen ist ein wichtiger Bestandteil vieler Algorithmen der Audiosignalverarbeitung, zum Beispiel bei der Reduktion von Störanteilen in einem Signal oder der Analyse akustischer Szenen. Mögliche Anwendungen finden sich unter anderem in der Telekommunikation und in der automatischen Spracherkennung. Die Klassifikation ist insbesondere dann eine anspruchsvolle Aufgabe, wenn Signale, die mit einem Mikrofon empfangen werden, durch räumlich bedingten Nachhall beeinflusst werden und wenn mehrere Quellen in einem Raum simultan aktiv sind. Dies führt zu einer Fehlanpassung eines Klassifikationssystems, insbesondere wenn dieses mittels sauberer Daten trainiert wird. Zur präziseren Klassifikation der Quellensignale, kann die räumliche Verteilung von Mikrofonen und Quellen ausgenutzt werden.

In der vorliegenden Arbeit wird ein System vorgestellt, das die Signale eines ad-hoc Mikrofon-Arrays verarbeitet um simultan aktive Schallquellen in einer halligen Umgebung zu klassifizieren. Im ersten Verarbeitungsschritt werden dazu Merkmalsvektoren aus den Mikrofonsignalen extrahiert. Diese Vektoren beinhalten in komprimierter Form notwendige Informationen um Signale verschiedener Signalklassen unterscheiden zu können. Ungenauigkeiten des Startzeitpunktes der Signalaufnahme unterschiedlicher Mikrofone, die bei der Verwendung von ad-hoc Mikrofon-Arrays unvermeidbar sind, haben durch die Wahl eines diesbezüglich robusten Merkmalsvektors keinen Einfluss auf die Klassifikationsleistung. Innerhalb der Merkmalsberechnung werden zudem Maßnahmen ergriffen, um den Einfluss des Nachhalls auf den erzeugten Merkmalsvektor auszugleichen.

Im Anschluss an die Merkmalsextraktion werden die Mikrofone, die sich jeweils in der Nähe einer Quelle befinden und daher von dem Signal dieser Quelle dominiert werden, automatisch in Gruppen zusammengefasst. Dieser Clustering-Vorgang nutzt ausschließlich die Merkmalsvektoren und benötigt keine Informationen über die Positionen der Quellen und der Mikrofone im Raum. Schließlich wird für die Klassifikation der Quellen eine Kombination der Informationen aus der Merkmalsberechnung innerhalb und zwischen den Mikrofon-Gruppen durchgeführt. Auf diese Weise wird ein Merkmalsvektor für jedes Mikrofon-Cluster ermittelt, der zu einer präzisen Klassifikation der in dieser Gruppe dominanten Quelle genutzt werden kann.

Die Evaluierung des vorgestellten Systems zur Klassifikation akustischer Quellen mittels ad-hoc Mikrofon-Arrays basiert sowohl auf simulierten als auch auf aufgenommenen Mikrofonsignalen. Es wird gezeigt, dass der Einfluss des Nachhalls in der Merkmalsextraktion effektiv reduziert werden kann. Werden automatisch ermittelte Mikrofon-Cluster und entsprechende Strategien zur Kombination der Informationen zwischen den Clustern genutzt, lassen sich akustische Quellen trotz der Störeinflüsse mit hoher Genauigkeit klassifizieren.