

Komplexe Modellierung: Bioakustik – Automatisches Erkennen von Vogelstimmen

In diesem Beitrag wird die sehr komplexe Fragestellung des automatischen Erkennens von Vogelstimmen anhand von Audioaufnahmen behandelt. Da es sich hierbei um einen Gegenstand der aktuellen Forschung handelt, der in verschiedenen Disziplinen thematisiert wird, darunter die digitale Signalverarbeitung (Informatik bzw. Elektrotechnik), neuronale Netze (Informatik bzw. Mathematik) oder Fourier-/Wavelet-Analyse (Mathematik), ist es im Rahmen des zur Verfügung stehenden Platzes nicht möglich Details zu diskutieren. Es geht vielmehr um die Darstellung verschiedener Möglichkeiten, die sich für eine Umsetzung in der schulischen oder universitären Lehre ergeben. Diese können durch Beschränkung auf Teilfragestellungen und Verwendung von Black-Box-Modellen auf unterschiedlichem inhaltlichen Niveau und mit variierendem Zeitbedarf realisiert werden.

1. Einleitung

Singvögel am Gesang zu erkennen ist eine Fähigkeit, die einige Übung erfordert. Mit etwas Training können viele Menschen nach kurzer Zeit zumindest sehr markante Arten wie Tauben, Amseln und Krähen auseinanderhalten – Experten können sogar Unterarten beispielsweise aus der Familie der Meisen oder Funken anhand ihres Gesangs unterscheiden. Will man die Identifizierung von Vögeln über ihren Gesang einem Computer übertragen, so legt man ihm Audiodateien vor und muss diejenigen Eigenschaften ermitteln, die möglichst deutlich – und robust – die Unterschiede herausstellen. Dabei ist eine geradezu unendliche Vielfalt an Möglichkeiten vorhanden, unter denen eine gute Auswahl zu treffen ist. Es fängt an beim Format: Arbeiten wir mit einem klassischen Zeit-Amplituden-Diagramm (s. Abb. 1, links) oder untersuchen wir das zugehörige sogenannte Spektrogramm, welches ein Zeit-Frequenz-Diagramm ist (s. Abb. 1, rechts)? Und wie ist der (mathematische) Zusammenhang zwischen beiden Darstellungen? Können wir die Eigenschaften, an denen Menschen Vogelstimmen erkennen, auf diese Darstellungen übertragen oder gibt es evtl. weitere Differenzierungsmöglichkeiten, die vielleicht nur in der technischen Darstellung erkennbar und für Menschen ohne Hilfsmittel gar nicht auszuwerten sind? Bei der Bearbeitung dieser komplexen Fragestellung steht man zunächst vor der Aufgabe, für die weitere Analyse geeignete Audiodateien zu erzeugen. Möchte man etwa mit dem Smartphone im Wald Vogelstimmen aufnehmen, die später von einer Software identifiziert werden, so gibt es eine Reihe von Hindernissen zu überwinden: Ein Ziel ist es, Audiosignale zu erhalten, die möglichst frei von Störgeräuschen sind.

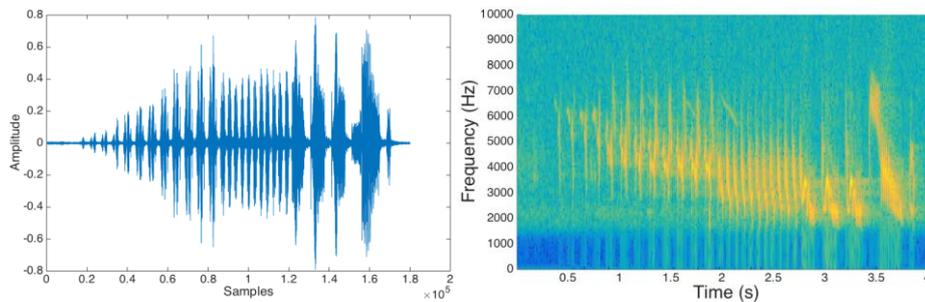


Abb. 1: Zeit-Amplituden-Diagramm (links) bzw. Zeit-Frequenz-Diagramm (rechts) des Gesangs eines Buchfinks (Quelle: Visualisierung mit MATLAB®)

Diese können durch die Aufnahmetechnik bedingt sein und sich einfach durch einen signifikanten Rauschpegel äußern, es können allerdings auch andere Störungen wie Blätterrauschen oder Plätschern eines Bachlaufs auftreten, im städtischen Umfeld ist die Vielfalt unnatürlicher Geräusche entsprechend größer. Und letztlich stellt es auch ein Problem dar, wenn ein ganzer Chor unterschiedlicher Vogelgesänge aufgenommen wird, aus dem einzelne Stimmen identifiziert werden sollen. In dieser Arbeit möchten wir uns auf die vereinfachte Situation konzentrieren, dass eine technische saubere Aufnahme eines einzelnen Vogelgesangs ohne Störgeräusche vorliegt und analysiert werden soll. Interessante Fragestellungen ergeben sich im Zusammenhang mit der Verwendung von Parabolspiegeln zur Verbesserung der Aufnahmequalität (Bock, Bracke & Gabler, 2015).

2. Modellierung und mathematischer Hintergrund

Wie bereits angedeutet gibt es eine Vielfalt mathematischer Zugänge zur Identifizierung von Vogelstimmen, wobei wir uns hier auf die Arbeit mit der klassischen Fourier-Analyse beschränken wollen. In einem Proseminar *Mathematische Modellierung* an der TU Kaiserslautern wurde die Fragestellung im Wintersemester 2015/16 von einer Gruppe Studierender im zweiten Studienjahr bearbeitet. Die Lösung basiert auf (diskreten) Fourier-Transformationen von Audiodateien fester Länge, die allerdings noch mit manueller Unterstützung aus einer allgemeinen Aufnahme ausgeschnitten werden müssen. Die so erhaltenen Vektoren fester Länge repräsentieren in einem gewissen Sinn das typische Frequenzspektrum einer Vogelart und werden durch die *Pearson-Korrelation* miteinander verglichen. Dadurch können zu identifizierende Signale mit einer vorhandenen Datenbank abgeglichen und ein Vogel identifiziert werden. Notwendige Voraussetzung ist dabei natürlich, dass ein oder mehrere Exemplare der zu identifizierenden Art bereits erfasst wurden. Durch die Beschränkung auf das zeitunabhängige Frequenzspektrum ist die in diesem Projekt gewählte Lösung allerdings eingeschränkt in der Hin-

sicht, dass zusätzliche Informationen, welche in der zeitabhängigen Unterteilung eines Gesangs etwa in Strophen, Phrasen, Silben oder Elemente stecken, nicht genutzt werden können.

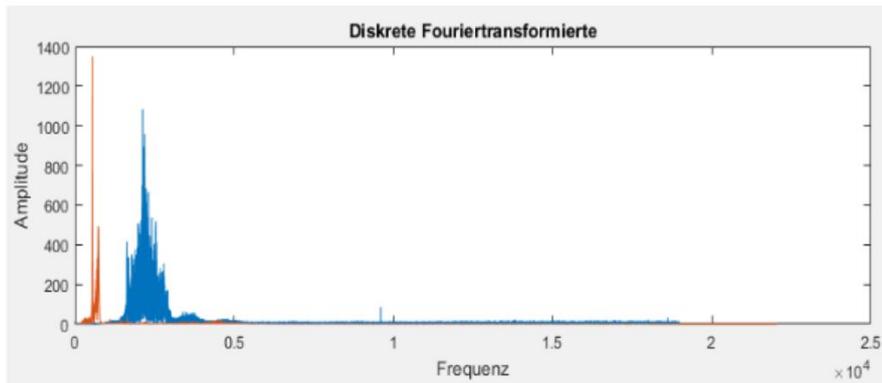


Abb. 2: Diskrete Fourier-Transformierte einer Amsel (blau) und eines Kuckucks (rot), welche anschließend mit Hilfe der Pearson-Korrelation miteinander verglichen werden.

Im Juli 2015 bearbeitete eine Projektgruppe von Schülerinnen und Schülern der Klassenstufen 11 und 12 im Rahmen der mathematischen Modellierungswoche der TU Kaiserslautern (Informationen zur Modellierungswoche s. Bock & Bracke (2014)) dieselbe Fragestellung. Hierbei wurden neben dem klassischen Frequenzspektrum auch sogenannte Sonogramme verwendet, die aus einer gefensterten Fourier-Analyse resultieren und damit einen zeitabhängigen Verlauf der Frequenzen darstellen können, der große Parallelen zur Darstellung von Musiknoten aufweist und ein wichtiges Werkzeug von Ornithologen ist. Sehr spannende Möglichkeiten in der Modellierung bietet die Arbeit mit solchen Sonogrammen, wenn die den Ornithologen bekannten Unterscheidungsmerkmale in mathematisch sauber definierte und robust bestimmbare Parameter abgebildet werden sollen. Interessant ist dabei, dass auch ohne ein lückenloses theoretisches Verständnis der zugrundeliegenden Mathematik (indem die entsprechenden mathematischen Verfahren als Black-Box verwendet werden – was in diesem Fall absolut legitim ist) spannende Fragestellungen aufgeworfen und bearbeitet werden können, wobei präzise und rigoros argumentiert wird.

3. Umsetzungsmöglichkeiten in der Sekundarstufe

Erstaunlich scheint vermutlich die Aussage, dass die ursprüngliche Fragestellung bereits in der Sekundarstufe I zumindest in Teilaspekten bearbeiten werden kann. So wurde im Rahmen einer Masterarbeit eine Einführung in die Akustik über einen mathematischen Zugang im Physikunterricht der Klassenstufe 7 erfolgreich umgesetzt, der die Idee der Fourier-Analyse beinhaltet (Saygushev, 2016). Wie schon zuvor angesprochen, kann man das

Teilproblem der Verbesserung von Audioaufnahmen mit Hilfe von Parabolspiegeln sehr schön als mathematisch-physikalisches Projekt in der Mittelstufe umsetzen (Bock, Bracke & Gabler, 2014). Innerhalb einer dreijährigen Junior-Ingenieur-Akademie wurde sogar die ganze Fragestellung in der Mittelstufe bearbeitet (Bock & Bracke, 2013). Hier lag der Beginn in der experimentiellen und theoretischen Erarbeitung der mathematischen und physikalischen Grundlagen der Akustik. Anschließend wurden selbst erstellte Audioaufnahmen in einer Black-Box-Prozedur in Sonogramme transformiert, die in ihrer diskreten Form einfach Tabellen sind, welche sogar mit Excel bearbeitet werden können. Der Entwicklung und praktischen Erprobung eigener Algorithmen auf unterschiedlichem mathematischen Niveau sind dabei kaum Grenzen gesetzt.

4. Fazit

Die komplexe Aufgabe des Automatischen Erkennens von Vogelstimmen bietet Potential für eine schulische Umsetzung in Form unterschiedlicher Teilprojekte in den verschiedenen MINT-Disziplinen. Dabei ergeben sich insbesondere interessante Herausforderungen für mathematische Modellierungen, kombiniert mit der eigenständigen, kreativen Entwicklung von Algorithmen und der Umsetzung mit dem Computer.

Danksagung

Die vorliegende Arbeit wurde mit finanzieller Unterstützung durch das Projekt *SchuMaMoMINT* des Europäischen Sozialfonds (ESF) in Rheinland-Pfalz erstellt.

Literatur

- Bock, W. & Bracke, M. (2013). *Project Teaching and Mathematical Modeling in STEM Subjects: A Design Based Research Study*. Proc. of CERME 8.
- Bock, W. & Bracke, M. (2014). Angewandte Schulmathematik – Made in Kaiserslautern. In Neunzert, H. & Prätzel-Wolters, D. (Hrsg.). *Mathematik im Fraunhofer-Institut. Problemgetrieben – Modellbezogen – Lösungsorientiert*. Springer Berlin, Heidelberg.
- Bock, W., Bracke, M. & Gabler, K. (2015). Mathematische Modellierung und Konstruktion eines Parabolspiegels zur Verstärkung von Audiosignalen, *Der Mathematikunterricht 61 (2015) Heft 5*, Friedrich-Verlag.
- Saygushev, D. (2016). *Fourier-Analyse im Unterricht. Ein anwendungsbezogener Zugang mit Umsetzung in der Schule*. Masterarbeit in Mathematik (M.Ed.), Technische Universität Kaiserslautern.