

Merkmalsselektion von psychoakustischen Größen bei der datenbasierten Modellierung der subjektiven Geräuschqualität

Lukas Outzen¹, Tobias P. Ring², Markus Scholz³, Sabine C. Langer⁴

¹ *Institut für Akustik, TU Braunschweig, Email: l.outzen@tu-braunschweig.de*

² *Institut für Akustik, TU Braunschweig, Deutschland, Email: t.ring@tu-braunschweig.de*

³ *IAV GmbH, Gifhorn, Email: markus.scholz@iav.de*

⁴ *Institut für Akustik, TU Braunschweig, Email: s.langer@tu-braunschweig.de*

Einleitung

Bei der Geräuschwahrnehmung von Fahrzeugen spielt der empfundene akustische Komfort eine bedeutende Rolle. Da diese subjektive Empfindung jedoch nicht direkt messbar ist, wird die Geräuschqualität in der Praxis durch aufwendige Probandentests erhoben. In diesem Beitrag wird eine alternative Methode vorgestellt, bei der die aus den Probandentests resultierenden subjektiven Geräuschbewertungen durch ein künstliches neuronales Netz (KNN) objektiviert werden. Dabei wird durch das Modell der Zusammenhang zwischen den aus den Geräuschen berechneten akustischen Parametern und den von den Probanden empfundenen Subjektivurteilen der Geräuschqualität hergestellt.

Die Betrachtungen erfolgen am Beispiel der Geräusche von Fahrzeugklimageräten, deren breitbandiges Rauschen abhängig von den typenspezifisch zusätzlich auftretenden Geräuscheigenschaften (z.B. Tonhaltigkeiten durch rotierende Komponenten) unterschiedlich angenehm empfunden wird. Die Geräuscheigenschaften der Klimageräte werden von konventionellen akustischen und psychoakustischen Parametern erfasst, während die subjektiven Bewertungen der Klimagerätgeräusche von mehreren Probanden in einem Hörversuch erhoben werden.

Der Fokus des vorliegenden Beitrags liegt auf dem Prozess der Merkmalsselektion. Diese dient dazu, eine Kombination von Eingangsgrößen (akustischen Parametern) für das KNN zu finden, die zu einer möglichst hohen Vorhersagegenauigkeit führt. Je nach Anwendungsfall tragen nur bestimmte Geräuschcharakteristika zur subjektiven Empfindung des Geräuschs bei. Da irrelevante Eingangswerte die Güte des Modells reduzieren, ist eine geeignete Merkmalsselektion essenziell für eine gute Annäherung der empfundenen Geräuschqualität [1]. Deshalb werden in diesem Beitrag verschiedene Merkmalsselektionsstrategien verglichen und ihre Effektivität auf Basis der resultierenden Merkmalskombinationen interpretiert.


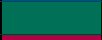












Datenbasis

Für die zugrundeliegenden Untersuchungen stehen 110 Aufnahmen der Geräusche von unterschiedlichen Klimageräten in einem vergleichbaren Betriebszustand zur Verfügung. Aus diesen wurden einerseits akustische Parameter extrahiert, die als Modelleingangswerte dienen. Die Geräuschbewertungen, die die Zielwerte für das Modell darstellen, werden in einem Hörversuch erhoben.

(Psycho-)akustische Parameter

Um die Geräuschcharakteristika der Klimagerätgeräusche zu erfassen, werden mithilfe der Analyse-Software *ArtemiS* der Firma *HEAD acoustics* verschiedene akustische Parameter berechnet. Diese sind in Tabelle 1 mit der jeweiligen Anzahl ihrer Werte dargestellt. Außerdem zeigt die Tabelle die Farben, die im weiteren Verlauf dieses Beitrags für die jeweiligen Parameter verwendet werden.

Tabelle 1: Verwendete akustische Parameter

Akustischer Parameter	Werte	Farbe
Schalldruckpegel	1	
Terzspektrum	34	
Spezifische Lautheit	24	
Lautheit	1	
Schärfe	1	
Spezifische Tonhaltigkeit	53	
Tonhaltigkeit	1	
Spezifische Schwankungsstärke	21	
Schwankungsstärke	1	
Spezifische Rauigkeit	24	
Rauigkeit	1	
Spezifische Impulshaltigkeit	24	
Impulshaltigkeit	1	
Alle Parameter	187	

Der Schalldruckpegel und das Terzspektrum können hierbei als konventionelle akustische Parameter bezeichnet werden, während die Übrigen entwickelt wurden, um die menschliche Wahrnehmung von Hörereignissen zu beschreiben und quantifizieren, und somit in den Bereich der Psychoakustik fallen. Einige dieser psychoakustischen Parameter können zum einen als Einzahlwert berechnet, aber auch frequenzgruppenspezifisch ermittelt und somit in mehreren Werten ausgedrückt werden.

Hörversuch

Die zur Verfügung stehenden Geräusche wurden von vier Probanden subjektiv bewertet. Dabei wurde eine zehnstufige kardinale Notenskala verwendet, die in der Automobilindustrie weit verbreitet ist [2]. Anhand dieser Skala werden die Klimagerätgeräusche von den Probanden zwischen eins (nicht akzeptabel) und 10 (nicht feststellbar) eingeordnet. Anschließend werden die Bewertungen für jedes Geräusch gemittelt.

Datenbasierte Modellierung

Wie einleitend erwähnt, wird zur Modellierung des Zusammenhangs zwischen den berechenbaren akustischen Parametern und dem Geräuschqualitätsempfinden der Probanden ein künstliches neuronales Netz (KNN) verwendet. KNN sind eine Methode des maschinellen Lernens, dessen zugrundeliegende Idee es ist, dass die Informationsverarbeitung nicht durch explizite Regeln festgelegt ist, sondern basierend auf vorhandenen Daten Muster erkannt und so Zusammenhänge „gelernt“ werden. Im Gegensatz zu herkömmlichen Methoden wie der linearen Regression können durch ein KNN auch nichtlineare Zusammenhänge hergestellt werden [3].

Für das Modelltraining werden dem KNN als Eingangswerte die in Tabelle 1 dargestellten akustischen Parameter übergeben. Als Zielwerte des überwachten Lernalgorithmus werden dem KNN die gemittelten subjektiven Bewertungen der Probanden übergeben. Während des Trainingsprozesses werden die Modellparameter iterativ so angepasst, dass die vom KNN vorhergesagten Bewertungen möglichst geringe Abweichungen von den Zielwerten haben.

Für die Untersuchungen dieses Beitrags wird das KNN als Regressionsmodell angewandt, wodurch eine Vorhersage der subjektiven Bewertungen auf dem kontinuierlichen Wertespektrum zwischen 1 und 10 möglich wird. Die Hyperparameter des KNN werden auf Werte festgelegt, die mit der gegebenen Datenbasis gemäß dahingehender Voruntersuchungen gute Ergebnisse erzielen. Als Optimierungsmethode wird das L-BFGS-Verfahren verwendet. Das KNN hat eine versteckte Schicht mit 100 Neuronen, deren Aktivierung durch die ReLU-Funktion erfolgt. Der Regularisierungsparameter wird auf einen Wert von 0,1 festgelegt.

Von den 110 Geräuschen werden 10 für den finalen Test des Modells zufällig ausgewählt. Damit stehen 100 Geräusche für Training und Validierung des Modells zur Verfügung. Um die Reproduzierbarkeit der Ergebnisse zu erhöhen, werden für jedes Modell 25 Durchläufe einer 10-fachen Kreuzvalidierung durchgeführt und die Ergebnisse gemittelt.

Ergebnisse

Zunächst wird die Vorhersagefähigkeit des Modells unter Nutzung einzelner akustischer Parameter untersucht. Gemäß Tabelle 1 wird dem Modell hierbei teilweise ein einzelner Wert übergeben, teilweise aber auch ein Vektor mit einer Größe von 21 bis 53 Werten.

Die Abbildung 1 visualisiert die Ergebnisse dieser Untersuchung. Dabei wird die Vorhersagefähigkeit mit dem Bestimmtheitsmaß R^2 quantifiziert. Dieses Bewertungskriterium kann als der Anteil der Zielwerte interpretiert werden, der durch das Modell erklärt werden kann. Ein R^2 -Wert von 1 bedeutet etwa, dass die subjektiven Bewertungen vollständig durch die Modelleingangswerte erklärbar sind. Bei einem R^2 -Wert von 0 wiederum wären die Vorhersagen so ungenau als würde ein naives Modell verwendet werden, das immer den durchschnittlichen Zielwert voraussagt.

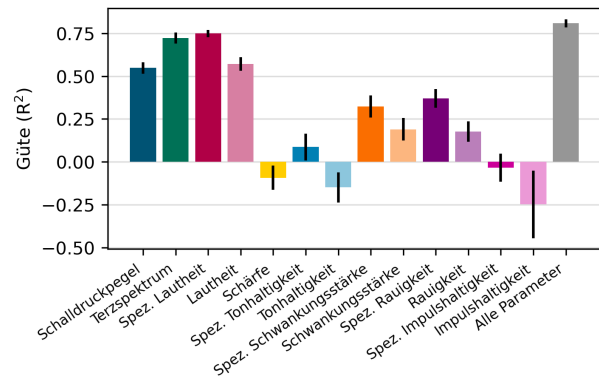


Abbildung 1: Erzielte Güte bei Verwendung einzelner und aller akustischen Parameter

Die mit den einzelnen akustischen Parametern erzielte Güte spiegelt auch ihren Informationsgehalt bezüglich der empfundenen Geräuschqualität wider. Während einige Parameter (wie der Pegel, das Terzspektrum und die (spezifische) Lautheit) einzeln bereits einen relevanten Zusammenhang zu den Geräuschbewertungen aufzuweisen scheinen, resultieren andere (wie die Schärfe, die (spezifische) Tonhaltigkeit und die (spezifische) Impulshaltigkeit) in einer niedrigen Vorhersagegüte.

Zusätzlich zu der erreichten Güte mit den einzelnen akustischen Parametern ist in Abbildung 1 als grauer Balken die Güte des Modells dargestellt, dem alle zur Verfügung stehenden Parameter übergeben werden. Die Güte erreicht hier mit $R^2 > 0,8$ den höchsten Wert. Dies entspricht einer mittleren Abweichung der vorhergesagten Werte von den tatsächlichen Zielgrößen von etwa 0,37 auf der Bewertungsskala von 1 bis 10. Das bedeutet, dass zusätzlich zur spezifischen Lautheit, die einzeln die höchste Relevanz aufweist, die anderen akustischen Parameter zusätzliche Informationen enthalten. Allerdings ist es trotzdem wahrscheinlich, dass die unterschiedlichen Parameter auch Redundanzen aufweisen. Redundante und irrelevante Informationen erhöhen nicht nur die Rechenleistung, sondern können zudem eine Verringerung der Vorhersagegüte und der Generalisierungsfähigkeit bewirken [1].

Da im vorliegenden Fall die Nutzung aller akustischen Parameter vorteilhaft ist, überwiegt hier der Gütegewinn durch die zusätzlichen Parameter gegenüber dem Güteverlust durch die nicht aussagekräftigen Informationen. Trotzdem liegt die Vermutung nahe, dass eine noch höhere Vorhersagegüte durch eine Kombination der geeigneten und einen Ausschluss der redundanten und irrelevanten akustischen Parameter möglich wäre.

Merkmalsselektion

Das Ziel der Merkmalsselektion ist es, die Menge der Eingangswerte hinsichtlich der mit ihnen erreichbaren Vorhersagegüte zu optimieren. Dabei gilt es, die Parameterkombination zu finden, die einen möglichst hohen Informationsgehalt enthält. Die restlichen Parameter, die keinen zusätzlichen Informationsgewinn bedeuten und so die Effizienz und Güte des Modells verschlechtern, sollen hierbei identifiziert und ausgeschlossen werden.

Im Folgenden werden drei Methoden zur Merkmalsselektion diskutiert und angewandt. Bei der ersten findet die Merkmalsselektion durch ein Filterkriterium unabhängig vom maschinellen Lernalgorithmus statt. Als Kriterium dient hierbei die lineare Korrelation der Eingangswerte mit den Zielwerten. Bei den beiden anderen Methoden, der Vorwärtsselektion und der Rückwärtselimination, wird im Gegensatz dazu mit dem Lernalgorithmus über mehrere Merkmalskombinationen iteriert.

Merkmalsselektion durch Korrelationsanalyse

Die Filterung der Eingangswerte nach hohen Korrelationskoeffizienten basiert auf der Idee, dass die linear stark korrelierenden Werte auch für das nichtlineare KNN die größte Aussagekraft besitzen. Um dies zu untersuchen, wird der Korrelationskoeffizient der 187 einzelnen Eingangswerte mit den Zielwerten bestimmt und eine variable Anzahl n_k an höchstkorrelierenden Werten mit dem Modell getestet. Die Abbildung 2 zeigt die Ergebnisse dieser Untersuchung.

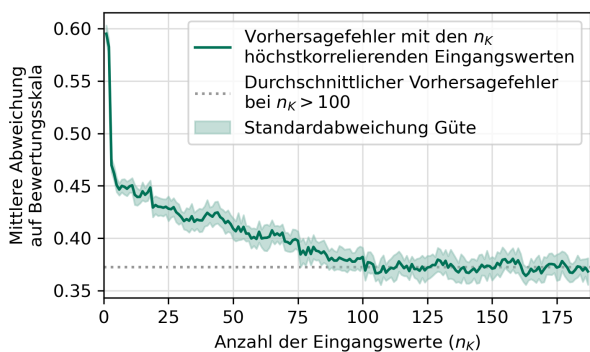


Abbildung 2: Ergebnisse der Merkmalsselektion durch die Auswahl höchstkorrelierender Werte

Während bei einer steigenden Anzahl an linear korrelierenden Eingangswerten bis hin zu $n_k \approx 100$ der Vorhersagefehler abnimmt, bleibt die Güte danach ungefähr konstant bei einer Abweichung auf der Bewertungsskala von durchschnittlich 0,37. Somit bringt diese Methode zwar durch eine Verringerung der Menge an Eingangswerten einen Effizienzgewinn, jedoch keinen qualitativen Mehrwert für das Modell. Nachteil dieser Methode und damit eine mögliche Erklärung dieses Ergebnisses ist, dass durch die Auswahl der am stärksten linear korrelierenden Einzelwerte eine große Menge an redundanten Informationen aufgenommen wird, während Merkmale, die nichtlinear mit dem Zielwert verbunden sind und deren Aussagekraft vor allem durch Interdependenzen mit anderen Merkmalen entsteht, nicht berücksichtigt werden. Dieses Problem soll durch Methoden beseitigt werden, bei denen eine handhabbare Menge an Merkmalskombinationen strategisch ausgewählt und mit dem Lernalgorithmus ausprobiert wird.

Vorwärtsselektion

Das Vorgehen bei der Vorwärtsselektion ist folgendermaßen strukturiert:

1. Initiale Menge der Eingangsmerkmale ist leer.
2. Alle verfügbaren Merkmale werden einzeln temporär zu den Eingangsmerkmalen hinzugefügt und die re-

sultierende Güte für jede Kombination bestimmt.

3. Das Merkmal, dessen Hinzunahme die beste Güte zufolge hatte, wird fest zu den Eingangsmerkmalen hinzugefügt.
4. Wiederholung ab Schritt 2 mit den übrigen Merkmalen. Ggf. Abbruch, wenn erreichte Güte bei einzelner Hinzunahme jedes weiteren Merkmals schlechter ist als in der vorherigen Iteration.

Somit werden im Laufe der Vorwärtsselektion iterativ diejenigen Merkmale hinzugefügt, die den größten Mehrwert hinsichtlich einer höheren Güte aufweisen.

Da die Modellparameter eines KNN zufällig initiiert werden, können die Ergebnisse solcher Methoden bei mehreren Anwendungen unterschiedlich ausfallen. Da eine einzelne Vorwärtsselektion somit nicht aussagekräftig wäre, zeigt die Tabelle 2 zehn solcher iterativen Auswahlprozesse. Bei Berücksichtigung der farblichen Codierung ist hier die Reihenfolge der Hinzunahme von Parametern in die Menge der Eingangsmerkmale dargestellt. Bei der Wiederholung 1 etwa wird als erster Parameter die spezifische Lautheit übernommen, bevor nach Hinzufügen der Lautheit als Parameter 6 keine Verbesserung der Güte und somit keine weitere Hinzunahme von Merkmalen erfolgt.

Tabelle 2: Vergleich mehrerer Wiederholungen der Vorwärtsselektion (Legende s. Tabelle 1)

Wiederholung	1	2	3	4	5	6	7	8	9	10
Parameter 1	Red	Red	Red	Red	Red	Red	Red	Red	Red	Red
Parameter 2	Blue	Pink	Blue	Pink	Blue	Pink	Blue	Pink	Blue	Pink
Parameter 3	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
Parameter 4	Light Blue	Light Blue	Yellow	Light Blue	Yellow	Light Blue	Yellow	Light Blue	Yellow	Light Blue
Parameter 5	Purple	Purple	Purple	Purple	Purple	Purple	Purple	Purple	Purple	Purple
Parameter 6	Pink	Pink	Green	Pink	Green	Pink	Green	Pink	Green	Pink
Parameter 7	White	White	White	White	White	White	White	White	White	White
Parameter 8	White	Yellow	White	Purple	White	Pink	White	Pink	White	White
Parameter 9	White	Pink	White	Blue	White	Purple	White	White	White	White
Parameter 10	White	Light Blue	White	White	Blue	White	White	White	White	White

Bei allen Wiederholungen wird einheitlich die spezifische Lautheit als erstes Merkmal aufgenommen. Dies entspricht den Beobachtungen aus Abbildung 1, bei denen diese als einzelner akustischer Parameter das beste Ergebnis erzielt. Eine weitere Gemeinsamkeit ist die Aufnahme der akustischen Parameter Schalldruckpegel, Terzspektrum und Lautheit. Somit scheinen die Informationen der stark mit der Lautheit korrelierenden akustischen Parameter nicht redundant – sie ergänzen sich vielmehr hinsichtlich einer besseren Güte. Weiterhin scheinen die Tonhaltigkeit und die spezifische Rauigkeit über den Lautheitseindruck hinausgehende Informationen zur Modellierung der subjektiven Bewertungen zu enthalten.

Die Abweichung der vorhergesagten von den tatsächlichen Bewertungen beträgt bei der Vorwärtsselektion durchschnittlich 0,359 und die Anzahl der Eingangswerte durchschnittlich 125. Somit kann durch die Vorwärtsselektion eine Verbesserung der Vorhersagegüte bei einer Verringerung der Rechenleistung erreicht werden.

Rückwärtselimination

Im Fall bestimmter Interdependenzen zwischen Merkmalen ist es möglich, dass die Methodik der Vorwärtsselektion gewisse Zusammenhänge nicht berücksichtigen kann. Wenn beispielsweise zwei akustische Parameter allein stehend keine Aussagekraft haben, jedoch in Kombination relevante Informationen enthalten, kann die Vorwärtsselektion diese nicht berücksichtigen. Nur die Interdependenzen mit den bereits in den Eingangsmerkmalen enthaltenen akustischen Parametern können berücksichtigt werden und nicht noch zu entdeckende potenzielle Interdependenzen der noch nicht aufgenommenen Parameter.

Die Methode der Rückwärtselimination versucht dieses Problem zu beheben, indem hier mit allen Merkmalen gestartet wird und iterativ diejenigen Merkmale, die am wenigsten zur Güte des Modells beitragen, entfernt werden. Die Vorgehensweise hierzu funktioniert wie folgt:

1. Initiale Menge der Eingangsmerkmale besteht aus allen Merkmalen.
2. Alle verfügbaren Merkmale werden einzeln von der Menge der Eingangsmerkmale abgezogen, und die resultierende Güte ohne das jeweilige Merkmal für jede Kombination bestimmt.
3. Das Merkmal, dessen Entfernung die beste Güte zufolge hatte, wird aus der Menge der verfügbaren Merkmale entfernt und so in keiner der nachfolgenden Iterationen mehr untersucht.
4. Wiederholung ab Schritt 2 mit den übrigen Merkmalen. Ggf. Abbruch, wenn erreichte Güte bei der Entfernung jedes weiteren Merkmals schlechter ist als in der vorherigen Iteration.

Somit können von dieser Methode alle Interdependenzen zwischen den Parametern berücksichtigt werden.

Wie bei der Vorwärtsselektion ist in Tabelle 3 die Merkmalsauswahl von zehn Wiederholungen der Rückwärtselimination dargestellt. Auch hier sind die Merkmale nach ihrer Relevanz sortiert. Bei der Wiederholung 1 indiziert etwa der Parameter 1, dass das Terzspektrum als letztes aus der Menge der Eingangsmerkmale entfernt worden wäre. Der letzte Parameter einer Wiederholung (in diesem Fall Parameter 6, die spezifische Lautheit) ist hingegen derjenige, dessen Entfernung als erstes eine Verschlechterung der Güte zufolge hatte.

Tabelle 3: Vergleich mehrerer Wiederholungen der Rückwärtselimination (Legende s. Tabelle 1)

Wiederholung	1	2	3	4	5	6	7	8	9	10
Parameter 1										
Parameter 2										
Parameter 3										
Parameter 4										
Parameter 5										
Parameter 6										
Parameter 7										
Parameter 8										

Im Gegensatz zur Vorwärtsselektion ist in diesem Fall übereinstimmend das Terzspektrum das wichtigste Merkmal. Somit scheint dieses in Kombination mit anderen Merkmalen die relevantesten Informationen zu enthalten, obwohl es gemäß Abbildung 1 allein stehend nicht die größte Aussagekraft hat. Ähnlich wie bei der Vorwärtsselektion verbleiben weiterhin die Lautheit, der Schalldruckpegel und die spezifische Rauigkeit in allen Wiederholungen in der Menge der Eingangsmerkmale.

Im Falle der Rückwärtselimination beträgt die Abweichung der vorhergesagten von den tatsächlichen Bewertungen durchschnittlich 0,357 und die Anzahl der Eingangswerte 115. Somit wird im Vergleich zur Vorwärtsselektion hier durch eine effektivere Nutzung der Interdependenzen eine noch bessere Güte bei einer weiteren Verringerung der Eingangswerte erreicht.

Zusammenfassung

In diesem Beitrag wird der Ansatz vorgestellt, subjektive Empfindungen von Störgeräuschen durch ein datenbasiertes Modell zu objektivieren. Dabei werden als Modell eingangswerte berechenbare Parameter aus der Akustik und der Psychoakustik verwendet und als Zielwert eine Geräuschbewertung auf einer Notenskala prognostiziert.

Insbesondere wird der Aspekt der Merkmalsselektion betrachtet und drei Methoden hierzu verglichen. Dabei kann festgestellt werden, dass die einfache Auswahl der am stärksten mit dem Zielwert korrelierenden Eingangswerte die Vorhersagegüte nicht verbessern kann. Im Gegensatz dazu resultieren die Methoden der Vorwärtsselektion und der Rückwärtselimination, bei denen durch unterschiedliche Strategien Merkmalskombinationen ausprobiert werden, in einer Verbesserung des Modells. Durch die Rückwärtselimination können zudem Interdependenzen zwischen den Merkmalen effektiv berücksichtigt werden.

Die Ergebnisse zeigen, dass mit der vorgestellten Methode die subjektive Empfindung der Geräuschqualität modelliert werden kann. Für den Praxiseinsatz sollte eine umfassendere Datenbasis, insbesondere in Bezug auf die Repräsentativität der Hörversuche, aufgebaut werden.

Acknowledgement

Dieser Beitrag basiert auf der Masterarbeit des Erstautors mit dem Titel „Objektivierung von subjektiven Akustikbewertungen mit Hilfe von künstlichen neuronalen Netzen“, die im Zeitraum von November 2019 bis Juni 2020 am Institut für Akustik in Zusammenarbeit mit der IAV GmbH in Gifhorn angefertigt wurde.

Literatur

- [1] Saunders, C.: Subspace, Latent Structure and Feature Selection. Springer, Berlin, 2006
- [2] Zeller, P.: Handbuch Fahrzeugakustik. Springer Fachmedien, Wiesbaden, 2018
- [3] Dörn, S.: Programmieren für Ingenieure und Naturwissenschaftler. Springer Vieweg, Berlin, 2018
- [4] Tadist, K.: Feature selection methods and genomic big data: a systematic review. Journal of Big Data 6.1 (2019)