

Anwendung von Spracherkennern für die Klassifikation von Schnarchlauten

Rüdiger Hoffmann¹, Thomas Richter²

¹Institut für Akustik und Sprachkommunikation, TU Dresden, Ruediger.Hoffmann@ias.et.tu-dresden.de

²voiceINTERconnect GmbH Dresden, richter@voiceinterconnect.de

Einleitung

In der akustischen Praxis treten immer wieder Klassifikationsprobleme mit nichtsprachlichen Signalen auf, die erfolgreich durch die Anwendung von Spracherkennern gelöst werden können (vgl. auch [1]). Im vorliegenden Fall bestand die Aufgabe, Schnarchlaute zu klassifizieren. Die Eignung von zwei Worterkennern (einem DTW- und einem HMM-Erkennen) wurde untersucht, um in einem gegebenen Testkorpus sowohl zwischen Schnarch- und sonstigen Lauten als auch zwischen pathologischem und „harmlosem“ Schnarchen zu unterscheiden. Letzteres ist von Bedeutung für die Erkennung lebensbedrohlicher Zustände (Schlafapnoe). Die hier vorgestellten Ergebnisse stammen hauptsächlich aus der Diplomarbeit [2].

Anliegen

Viele Menschen (ca. 10 bis 30 % der Erwachsenen) schnarchen im Schlaf. Daraus ergeben sich zwei Aufgabenstellungen: Erstens ist Schnarchen, auch wenn es (noch) keine gesundheitliche Gefahr darstellt, für andere Personen lästig. Daraus ergeben sich Konzepte für Geräte, die Schnarchen erkennen und den Schnarcher auf geeignete Weise zu einer Positionsänderung o. ä. veranlassen und die eher dem Konsumgüterbereich zuzuordnen sind (z. B. [3]). Man findet daher auch einschlägige Patente [4].

Zweitens kann Schnarchen mit einer lebensbedrohlichen Erkrankung verbunden sein (Schlaf-Apnoe, Atemstillstand von ca. 1 s bis zu 1 min), die zum Sauerstoffmangel im Herzmuskel und im Gehirn führt. In diesem Fall ist es wichtig, Atemaussetzer zu erkennen, was u. a. auch über das Schnarchen möglich ist. Daher findet man auch in der medizinischen Literatur vereinzelte Untersuchungen des Schnarchsignals [5 - 9].

Beide Anwendungen zielen auf ein Klassifikationsproblem: Im ersten Fall muss man Schnarchen von Nichtschnarchen unterscheiden können, da bisherige, relativ simple Konstruktionen mit Schwellwertschaltern etc. auch bei externen Geräuschen ansprechen können. Zweitens ist eine Unterscheidung von ungefährlichem und medizinisch gefährlichem Schnarchen erforderlich, um die Apnoe-Behandlung unterstützen zu können.

Die Aufgabe ist komplex und kann daher mit verschiedenen Ansätzen behandelt werden. Aufgrund des technologischen Reifegrades der Spracherkennung liegt es nahe, neben anderen Verfahren (Beispiel wäre die Verwendung des Crest- oder Scheitelfaktors als Zeitbereichsmethode [7, 8]) auch den Einsatz von Spracherkennern zur Klassifikation von Schnarchlauten zu untersuchen. Der vorliegende Beitrag konzentriert sich auf diesen Aspekt.

Sprach- und Schnarchsignale haben aus der Sicht der Objekterkennung die Gemeinsamkeit, dass sie zwar jeder Mensch mühelos identifizieren kann, dass aber die technische Verarbeitung schwierig ist, weil es sich um sehr variable zeitliche Strukturen handelt, deren Erkennung zudem durch Störgeräusche stark beeinträchtigt wird.

Das Schnarchsignal

Das Schnarchen entsteht wohl durch eine Erhöhung des Widerstands der Atemwege im Schlaf, häufig auch durch anatomisch bedingte Engen im Pharynx [9]. Im Extremfall wird die Atmung völlig blockiert (obstruktive Schlafapnoe). Nach dem Ort der Entstehung unterscheidet man palatales und non-palatales Schnarchen. Aus Abbildung 1 ist erkennbar, dass ein palatales Schnarchgeräusch durch einen annähernd gleichbleibenden, quasiperiodischen zeitlichen Verlauf ausgezeichnet ist.

Unser Untersuchungsmaterial stammte aus Aufzeichnungen eines Schlaflabors von vier Testpersonen. Das Datenmaterial wurde in „Schnarchwörter“ zerlegt; dies sind Segmente, die aus jeweils einer Einatem- und einer Ausatemphase des Schnarchens bestehen. Ab-

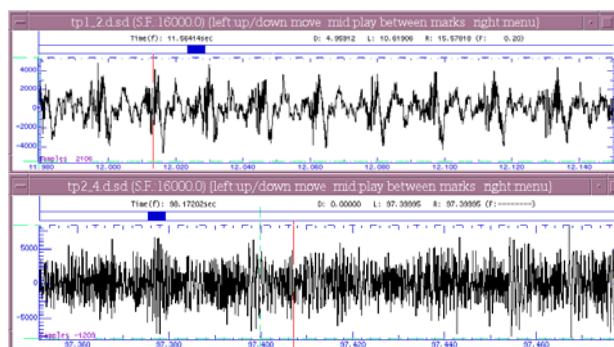


Abb. 1: Zeitfunktionen von Schnarchsignalen. Oben: palatales Schnarchen (170 ms), unten: non-palatales Schnarchen (130 ms).

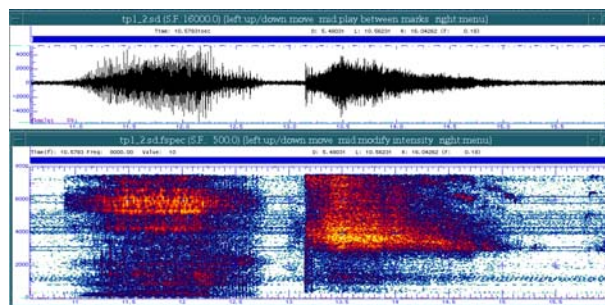


Abb. 2: Ein- und Ausatemphase eines Schnarchsignals (5,5 s) als Zeitfunktion und Spektrogramm (Frequenzbereich 8 kHz).

Kürzel	Beschreibung
S	silence, Ruhe, Pause
A	normales Atmen
N	non-palatales Schnarchen
P	palatales Schnarchen
B	burst, Plosionsgeräusch
E	Expiration, Ausatmung
M	Müll, Stör- / Umgebungsgeräusche

Tabelle 1: Verwendete Segmentklassen.

bildung 2 zeigt dafür ein Beispiel. Zur Etikettierung der Schnarchwörter wurden sieben Segmentklassen nach Tabelle 1 verwendet.

Verwendete Spracherkennung

Bekanntlich werden für Worterkennung je nach erforderlicher Performanz relativ einfache Vergleichsverfahren (DTW) oder aufwändige statistische Verfahren (HMM) verwendet. Für die hier vorgestellten Untersuchungen sollten beide Verfahren getestet werden. Auf die verwendeten Verarbeitungsschritte zur Merkmalgewinnung kann aus Platzgründen hier nicht eingegangen werden.

DTW-Erkennung

Mit der dynamischen Zeitanpassung (DTW = dynamic time warping) vergleicht man ein zu klassifizierendes Muster mit einem Muster, das zu einer bekannten Klasse gehört. Für die Versuche wurden Erkennungsmodule verwendet, die einer Eigenentwicklung entstammen [11].

HMM-Erkennung

Mit den in Tabelle 1 aufgezählten Klassen kann man zu einem „Schnarchwort“ ein Wortmodell analog zur Spracherkennung aufbauen, das in einem klassischen HMM-Erkennung eingesetzt wird. Konkret wurde der Erkennung verwendet, der als CD dem Buch [10] beigegeben ist.

Ergebnisse zur Unterscheidung von gefährlichem und ungefährlichem Schnarchen

Definition der Klassen

Aus den in Tabelle 1 definierten Segmentklassen lassen sich Sequenzen zusammenstellen, die in drei Wortklassen nach Tabelle 2 eingeteilt werden können.

WK	Klasse	Referenz-Sequenzen
2	Gefährlich	SNPSBES, SPSBES, SPSES, SPSAS, SNPS, SPS
1	Ungefährlich	SNSBES, SNSSES, SNSAS, SNS, SAS
0	Müll	SMS

Tabelle 2: Klasseneinteilung der Schnarchwortmodelle.

Ergebnisse

Ohne hier auf Details eingehen zu können, werden in Tabelle 3 die Erkennungsraten der Klassen aus Tabelle 2 für die jeweils leistungsfähigste Erkennungskonfiguration zusammen mit jeweils einer etwas relativierenden Bemerkung angegeben. Für die Untersuchungen standen über 1000 „Schnarchwörter“ zur Verfügung, die natürlich jeweils in eine Trainings- und eine Testmenge eingeteilt wurden.

Erkennung	Erk.rate	Bemerkung
DTW	92,4 %	Auswertung wurde durch Einbeziehung des Crest-Faktors des Signals unterstützt
HMM	83,1 %	Für statistische Verfahren ist die Trainings-Datenbasis nicht völlig ausreichend.

Tabelle 3: Erkennungsergebnisse zu Tabelle 2.

Ergebnisse zur Unterscheidung des Schnarchens vom Nicht-Schnarchen

Wie einleitend bemerkt, gibt es Anwendungsfälle, bei denen lediglich die Unterscheidung des Schnarchens von anderen Geräuschen von Interesse ist. Um auch hierzu Ergebnisse zu erhalten, musste

natürlich die Menge der über 1000 „Schnarchwörter“ aus der Datenbasis in „Schnarcher“ und „Nichtsnarcher“ verändert werden. Auf dieser Basis ergaben sich die Erkennungsergebnisse von Tabelle 4.

Erkennung	Erk.rate
DTW	90,7 %
HMM	98,9 %

Tabelle 4: Erkennungsergebnisse Schnarchen/Nichtsnarchen.

Zusammenfassung

Die Untersuchung hat gezeigt, dass es möglich ist, mit Hilfe von Worterkennung, wie sie in der Spracherkennung eingesetzt werden, auch Schnarchsignale unter den Aspekten gefährlich / ungefährlich sowie Schnarchen / Nichtsnarchen zu klassifizieren. Die erreichten Erkennungsraten liegen (grob eingeschätzt) in einer Größenordnung, wie sie auch in der Spracherkennung geläufig ist.

Danksagung

Der Universität der Bundeswehr Hamburg, namentlich Herrn Dipl.-Ing. Th. Netzel, danken wir für die großzügige Überlassung des Untersuchungsmaterials aus einem von ihm bearbeiteten Projekt [12] und für nützliche Literaturhinweise.

Literatur

- [1] D. Hirschfeld, R. Hoffmann, P. Holstein, M. Koch: Sound- und Mustererkennung unter Nutzung von Spracherkennungstechnologien. Vorliegender DAGA-Tagungsband.
- [2] T. Richter: Erkennung von Biosignalen. Diplomarbeit, TU Dresden, Fak. Elektrotechnik und Informationstechnik, 2001.
- [3] Schnarch-Blocker. Conrad Electronic, <http://www.conrad.de>
- [4] A. Andjulovici: Einrichtung zur Selbstunterdrückung des eigenen Schnarchens beim Schlafen. Deutsches Patentamt, Offenlegungsschrift DE 3529936 A1, 21. 8. 1985 / 26. 2. 1987.
- [5] F. Dalmasso, R. Benedetto, G. Righini, R. Spagnolo: Digital processing of snoring sounds. Kopenhagen: European respiratory journal, 1990.
- [6] F. Dalmasso, R. Prota: Snoring: analysis, measurement, clinical implications and applications. Kopenhagen: European respiratory journal, 1996.
- [7] J. E. Osborne et al.: A new acoustic method of differentiating palatal from non-palatal snoring. Clinical Otolaryngology and Allied Sciences, 24 (1999) 2, 130 – 133.
- [8] J. E. Osborne et al.: Is your snoring palatal? Snore sound crest factor gives answer. Clinical Otolaryngology and Allied Sciences 24 (1999) 4, 370 – 377.
- [9] J. Schäfer: Spektralanalyse schlafabhängiger Atemgeräusche der oberen Luftwege. Stuttgart: Hippokrates-Verlag, 1993.
- [10] C. Becchetti, L. Prina Ricotti: Speech Recognition – Theory and C++ Implementation. Chichester etc.: Wiley & Sons, 1999.
- [11] G. Strecha: Multilinguale Etikettierung natürlicher Sprachsignale auf Basis synthetischer Referenzsignale. Diplomarbeit, TU Dresden, Fak. Elektrotechnik und Informationstechnik, 2000.
- [12] R. Lunderstädt, Th. Netzel: Verbesserung der nächtlich assistierten Beatmung bei Schlafapnoe. <http://www.unibw-hamburg.de/MWEB/ifa/fit/>