

Ein multimediales Verfahren zur subjektiven Evaluierung der Sprachqualität von Kfz-Freisprecheinrichtungen

Wilfried Gädicke und Markus Lieb

Volkswagen AG, 38442 Wolfsburg, Deutschland, Email: [\[vorname\].\[nachname\]@volkswagen.de](mailto:{vorname}.{nachname}@volkswagen.de) (ä =ae)

Einleitung

Eine gute Sprachqualität von Freisprecheinrichtungen im Fahrzeug ist nicht nur eine Frage des Komforts, sondern sie trägt auch maßgeblich zur Erhöhung der Sicherheit der telefonierenden Fahrzeuginsassen bei. Schließlich bedeutet eine gute Sprachverständlichkeit für den Fahrer, dass er sich weniger auf das Verständnis des Telefongesprächs konzentrieren muss und somit der größte Teil seiner Aufmerksamkeit dem Verkehr gelten kann.

Auf dem Markt gibt es eine Vielzahl von Freisprechsystemen, und es entsteht daraus die Notwendigkeit, die Sprachqualität möglichst objektiv zu bewerten und zu vergleichen, um eine zunehmende Sprachqualität in den Fahrzeugen zu gewährleisten. Zur objektiven Bewertung von Sprachqualität sind bereits Verfahren bekannt. Alle diese Verfahren sind jedoch Modelle der Wahrnehmung der Sprachqualität durch Probanden. Aus diesem Grund bilden die Resultate von subjektiven Sprachqualitätsbewertungen stets die Referenz, von der auch die Genauigkeit der objektiven Verfahren abhängt. Die einfachste Möglichkeit zur Verbesserung der Genauigkeit subjektiver Probandentests, nämlich eine Erhöhung der Probandenzahl, hat den Nachteil, dass dadurch ein erheblicher zusätzlicher Zeit- und Kostenaufwand entsteht. Im Bereich der Bewertung von hochqualitativen digitalen Audiocodecs wurden bereits Verfahren entwickelt, die eine wesentlich differenziertere subjektive Bewertung der Audioqualität erlauben. Bei der Sprachqualitätsbewertung lag der Schwerpunkt dagegen auf der Entwicklung von Bewertungsskalen, jedoch wurde dem Bewertungsverfahren dabei weniger Aufmerksamkeit gewidmet.

Aus oben genannten Gründen lohnt es sich, die Aufmerksamkeit auf die subjektiven Bewertungsverfahren für Audioqualität zu lenken.

Zuerst wird in dieser Veröffentlichung das MUSHRA-Verfahren zur Bewertung von Audioqualität vorgestellt. Danach werden die Vor- und Nachteile von den in der Sprachqualitätsbewertung üblichen absoluten und komparativen Skalen erläutert. Abschließend wird ein neues multimediales Verfahren zur differenzierten Bewertung der Sprachqualität von Klangereignissen vorgeschlagen, das die Vorteile von absoluten und komparativen Skalen vereint.

MUSHRA

Die MUSHRA „Multi Stimulus test With Hidden Reference and Anchors“-Methode [1] wurde für Audiosignale entwickelt. Sie eignet sich für Signale mittlerer Qualität sowie für stark gestörte Signale.

Beim MUSHRA-Verfahren werden alle Bewertungen bezüglich einer Referenz durchgeführt. Um zu große Distanzen

bei der Bewertung der Perzeption von stark gestörten Signalen zu vermeiden, wird jedoch empfohlen, kein ungestörtes hochqualitatives Audiosignal als Referenz zu verwenden. Zuerst werden die Stimuli mehrerer Testsysteme mit der Referenz verglichen. Dabei kann sich unter den Teststimuli ebenfalls eine verborgene Referenz befinden. Des Weiteren kann der Versuchsleiter die Referenz jederzeit gegen einen anderen Teststimulus austauschen. Außerdem können auch die Stimuli der Testsysteme untereinander verglichen werden, was zu einem hohen Auflösungsgrad der Testergebnisse führt. Bei diesem Verfahren kommt eine Skala (z. B. Wertebereich 0-100) mit folgender Skalenteilung zum Einsatz: Excellent, Good, Fair, Poor, Bad.

Das MUSHRA-Verfahren kann ohne große Anpassungen auch für die Sprachqualitätsbewertung von Freisprecheinrichtungen verwendet werden. Als Referenzsignal bietet sich hierfür ein auf Telefonbandbreite tiefpassgefiltertes Sprachsignal an. Der Vorteil liegt darin, dass hierbei die Bewertung auf einer absoluten Skala durchgeführt werden kann und gleichzeitig die Systeme untereinander verglichen werden können. Nachteilig für die Sprachqualitätsbewertung ist an diesem Verfahren, dass aufgrund der Verwendung einer Referenz keine Kompatibilität zu den weit verbreiteten MOS-Bewertungen gegeben ist. Der Aufwand zum Vergleich aller Stimulikombinationen steigt mit der Anzahl der zu bewertenden Systeme erheblich.

Zur Behebung dieser Nachteile ist bei der Volkswagen AG ein neues Verfahren entstanden, das ohne eine Referenz auskommt. Es bietet dennoch eine differenziertere Bewertung, als die in der Sprachqualitätsanalyse üblichen Bewertungen mit einfachen Skalen, ohne dabei jedoch notwendigerweise die Kompatibilität zu den weit verbreiteten Skalen zu verlieren.

Ein multimediales Verfahren zur Sprachqualitätsbewertung

Die am häufigsten verwendete Skala ist die MOS-Skala, deren deutsche Repräsentation oft aus den folgenden fünf diskreten Kategorien besteht: ausgezeichnet, gut, ordentlich, dürftig, schlecht. Ein zu bewertendes Klangereignis muss vom Probanden einer Kategorie zugeordnet werden. Der MOS entsteht dann durch Bildung eines einfachen Mittelwertes über alle Probandenurteile der einzelnen Klangereignisse. Sollen dagegen Klangereignisse verglichen und Abstände zwischen den Klangereignissen bestimmt werden, kommen vergleichende Skalen mit folgender Skalenteilung zum Einsatz: viel besser (+3), besser (+2), wenig besser (+1), etwa gleich (0), weniger schlecht (-1), schlechter (-2), viel schlechter (-3).

Ein Vorteil des vergleichenden Verfahrens liegt darin, dass man bei der Bewertung eine höhere Genauigkeit erzielen

kann und die Klangereignisse dabei in eine eindeutige Reihenfolge gebracht werden. Da jedoch ein absoluter Bezugspunkt fehlt, ist eine Vergleichsmöglichkeit der Ergebnisse unterschiedlicher Versuchsdurchführung selten gegeben. Auch die Verwendung einer Referenz führt oft nicht zu einer Vergleichbarkeit der Ergebnisse, da stets eine Kontextabhängigkeit vorhanden ist (eine besonders gute Referenz kann alle weiteren Urteile negativ beeinflussen).

Die Idee des neuen multimedialen Verfahrens ist einfach: Anstelle einer Referenz werden hierbei alle Klangereignisse zueinander in Beziehung gesetzt. Somit können alle Klangereignisse außer dem jeweiligen zu bewertenden Klangereignis als Referenz betrachtet werden. In einem Optimierungsprozess werden dann vom Probanden nacheinander die Bewertungen für die einzelnen Klangereignisse und der perzeptuelle Abstand zu allen anderen Klangereignissen optimiert. Grundlage für das Verfahren bildet eine absolute Skala. Hierfür kann nur eine kontinuierliche Skala in Frage kommen. Es ist aber durchaus möglich, eine Skalenteilung entsprechend den Kategorien der MOS-Skala vorzunehmen. Es können auch beliebige andere kontinuierliche Skalen eingesetzt werden. Um dieses Verfahren zu realisieren wurde eine einfache grafische Benutzeroberfläche für Windows-PCs erstellt. Auf dieser grafischen Oberfläche wird für jedes Klangereignis ein eingängiges graphisches Symbol an einer zufälligen Position eingeblendet. Jedes dieser Symbole kann vorzugsweise mit einer Maus frei auf der Oberfläche bewegt werden. Beim Betätigen des Symbols erklingt das zugehörige Klangereignis. Auf der Ebene, in der die Symbole bewegt werden können, wird die gewünschte Skala eingeblendet. Der Proband hat nun die Aufgabe, die Symbole auf der absoluten Skala anzuordnen, so dass sie seinem Qualitätsempfinden entsprechen. Da alle Symbole gleichzeitig eingeblendet sind, kann der Proband auch jederzeit Vergleiche mit zuvor bewerteten Klangbeispielen vornehmen. Durch die Verknüpfung mit einem Symbol wird eine komparative Methode mit einer absoluten Methode kombiniert.

Ein wesentlicher Vorteil dieser Methode liegt in der intuitiven Vorgehensweise des Probanden. So wird der Proband nur sinnvolle Vergleiche vornehmen und direkt benachbarte Klangereignisse miteinander vergleichen, um deren Abstand auf der absoluten Skala zu bestimmen. Die intuitive Vorgehensweise der Probanden bei diesem Verfahren ähnelt dem Ablauf von genetischen Algorithmen, die bei Optimierungsproblemen für ihre Effizienz bekannt sind. Der Proband korrigiert immer nur die Qualitätsbewertungen, mit denen er am unzufriedensten ist. Die Kosten im genetischen Algorithmus entsprechen hierbei dem Grad der Unzufriedenheit des Probanden mit seinem eigenen Bewertungsergebnis.

Was als weiterer Vorteil dieses Verfahrens betrachtet werden kann, ist die spielerische und sehr effiziente Durchführung, die ein schnelles Ermüden der Probanden verhindert. Der Proband kann selber bestimmen, wie schnell er bei der Beurteilung vorgeht.

Ein leicht ersichtlicher Vorteil des multimedialen Verfahrens ist, dass 2- (oder sogar 3-) dimensionale Skalen eingeblendet werden können, auf denen mehrere abhängige oder unab-

hängige Qualitätsdimensionen parallel bewertet werden können.

Um eine repräsentative Aussage der Ergebnisse zu erhalten, ist es natürlich sinnvoll, die Anordnung der Symbole von mehreren Probanden durchführen zu lassen und dann durch statistische Auswertung das Ergebnis zu optimieren.

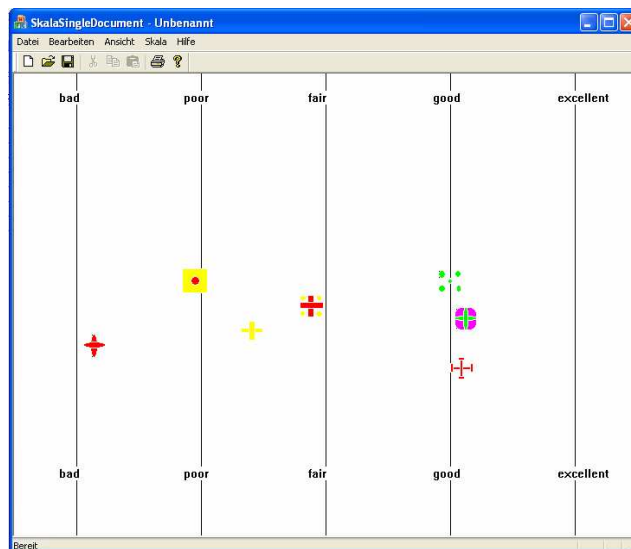


Abbildung 1: Grafische Benutzeroberfläche für das multimediale Verfahren zur Sprachqualitätsbewertung. Im Hintergrund wurde eine kontinuierliche MOS-Skala eingeblendet. Jedem der frei beweglichen farbigen Symbole wurde ein Klangereignis zugeordnet.

Zusammenfassung

In der Einleitung wird aufgezeigt, warum eine weitere Verfeinerung von subjektiven Sprachqualitätsmessverfahren sinnvoll erscheint. Ausgehend von der Betrachtung eines Verfahrens für die Qualitätsbewertung von hochqualitativen Audiocodern wird festgestellt, dass dieses Verfahren auch für den Telekommunikationsbereich sinnvoll eingesetzt werden kann. Danach stehen die in der Sprachqualitätsbeurteilung üblichen Verfahren im Mittelpunkt (Absolute Category Rating, Comparative Category Rating). Abschließend wird ein multimediales Verfahren vorgestellt, das die Vorteile der letztgenannten Methoden vereint und zusätzlich eine schnelle und intuitive Möglichkeit zur Sprachqualitätsbeurteilung bietet. Es ist vorgesehen, dieses *Tool* bei Volkswagen in Zusammenarbeit mit dem Institut für Nachrichtentechnik der TU Braunschweig zu einer kompletten Evaluierungsumgebung für Probandentests weiterzuentwickeln.

Literatur

- [1] G. Stoll u. F. Kozamernik, EBU listening tests on Internet audio codecs. EBU Technical Review (Juni 2004), S. 1-24
- [2] Jens Berger, Instrumentelle Verfahren zur Sprachqualitätsbewertung, Modelle auditiver Tests, 1998, Shaker Verlag.
- [3] ITU-R Rec. BS.1116-1 methods for the subjective assessment of small impairments in audio systems including multichannel sound systems