

Beschreibung der von Sprachkodierern hervorgerufenen Beeinträchtigungen mittels Impairment-Faktoren

Sebastian Möller, Institut für Kommunikationsakustik, Ruhr-Universität Bochum, moeller@ika.ruhr-uni-bochum.de
Jens Berger, T-Nova Deutsche Telekom Innovationsgesellschaft mbH Berkorn, Berlin, jens.berger@telekom.de

1. Zusammenfassung und Einleitung

Auf Telefonübertragungsstrecken (ISDN, mobil, IP-basiert) werden zur Datenreduktion in zunehmendem Maße Sprachkodierer mit Bitraten von 32-4 kbit/s eingesetzt. Die hierdurch hervorgerufenen Beeinträchtigungen wirken sich perzeptiv sehr unterschiedlich auf die Sprachübertragungsqualität aus. Zwar wurde in verschiedenen Untersuchungen die Multidimensionalität des Problems gezeigt (vgl. z.B. Bappert & Blauert, 1993); dennoch ist besonders im Anwendungsfall der Netzwerkplanung eine eindimensionale Beschreibung der Qualitätsbeeinträchtigung von Vorteil, um eindeutige Aussagen über optimal zu wählende Netzkonfigurationen machen zu können.

Zu der Beeinträchtigung durch Kodierer treten weitere Störungen, welche sich auf die vom Benutzer erfahrene Gesamtqualität negativ auswirken. Dies sind in traditionellen Netzen z.B. lineare Frequenzverzerrungen, Rauschen (Leitungsrauschen sowie Hintergrundgeräusche am Sende- und Empfangsort), Sprecher-Echos, Hörer-Echos, nicht-optimales Rückhören, oder Verzögerungszeiten (in der Konversations-Situation). In mobilen und paketorientierten Netzen kommen auch weitere nichtlineare und/oder zeitvariante Störungen, z.B. durch Übertragungsfehler, Sprach-Pausen-Detektion, *comfort noise injection*, etc. hinzu. Für die optimale Planung von Kommunikationsnetzen ist daher die alleinige Betrachtung einer einzigen Störungsart nicht ausreichend; vielmehr muss die Gesamtstörung quantifiziert werden, um bestmögliche Planungswerte für die später vom Benutzer erfahrene Gesamtqualität und letztendlich die Akzeptanz des Systems zu erhalten.

In diesem Beitrag wird die Beschreibung von Störungen durch Impairment-Faktoren analysiert, welche für die Qualitätsplanung von Telefonnetzen neuerdings als einzige Planungshilfe international empfohlen werden (ITU-T Rec. G.113, 2001). Die von Kodierern hervorgerufenen Beeinträchtigungen werden dabei durch sog. Equipment-Impairment-Faktoren (*Ies*), beschrieben. Sie sind als inkrementeller Wert der durch Sprachkodierer hervorgerufenen Beeinträchtigung definiert, welcher zu anderen Beeinträchtigungen (z.B. durch Rauschen, Echos, etc.) additiv ist. Die hier beschriebenen Untersuchungen belegen jedoch, dass die Additivität nicht in allen Fällen (z.B. bei Kodier-Kaskaden) gegeben ist. Es wurde eine neue Methode entwickelt, welche Impairment-Faktoren für neue Sprachkodierer in Bezug zum bestehenden Bewertungssystem verankert. Experimentelle Untersuchungen zur Verifikation der Methode werden vorgestellt. Sie zeigen, dass sich hiermit Beeinträchtigungen durch Kodierer für die Netzwerkplanung adäquat quantifizieren lassen.

2. Impairment-Faktor-Prinzip

Das Ziel des Impairment-Faktor-Prinzips ist es, die Auswirkungen von mehreren gleichzeitig auftretenden Störungen auf die vom Benutzer erfahrene Gesamtqualität zu quantifizieren. Da Störungen, die in modernen Netzen vorkommen, unterschiedliche perzeptive Dimensionen ansprechen, ist eine Dimensionalitäts-Reduktion notwendig. Diese erfolgt durch die Definition einer eindimensionalen Skala, der sog. *transmission rating scale*.

Ausgangspunkt der Beschreibung unterschiedlicher Störungsarten sind instrumentell messbare Parameter der Übertragungsstrecke, z.B. frequenzgewichtete Indizes für Abschwächungen (*loudness ratings*) und Rauschpegel (A- oder psophometrisch bewertet), Verzögerungszeiten, Bitfehlerraten oder Paketverlustraten, etc. Aus einem oder einer Kombination mehrerer dieser Parameter wird dann ein bestimmter perzeptiver Effekt bezüglich seiner Auswirkungen auf die

Gesamtqualität quantifiziert. Solche perzeptiven Effekte sind z.B. die nicht-optimale Lautheit der wahrgenommenen Sprache, Sprecher-Echos, nicht-optimales Rückhören, etc., und die Quantifizierung erfolgt auf der oben erwähnten *transmission rating scale*, in Form sog. Impairment-Faktoren. Die Transformationsregeln von instrumentellen Parametern auf Impairment-Faktoren wurden durch Vergleich mit den Ergebnissen auditiver Tests bestimmt. Sie bilden in ihrer Gesamtheit das E-Modell, welches z.B. bei Möller und Raake (2000) umrissen wird und in ITU-T Rec. G.107 (2000) standardisiert ist. Im E-Modell werden alle Impairment-Faktoren zu einer Gesamtstörung addiert und vom Qualitätswert der idealen Standard-Verbindung subtrahiert. Somit lässt sich ein Gesamtqualitäts-Index errechnen, welcher vorrangig zur Planung von modernen Telekommunikationsnetzen eingesetzt wird.

Bei der Beschreibung von Sprachkodierern besteht zunächst der Wunsch, die hervorgerufenen Beeinträchtigungen ebenfalls auf der *transmission rating scale* zu erfassen. Dies erfolgt durch die bereits erwähnten *Ies*, welche im E-Modell zu den Impairment-Faktoren anderer Beeinträchtigungen hinzu addiert werden. Es ergibt sich jedoch das Problem, dass Kodierer bislang nicht durch einfache, instrumentell messbare Parameter beschrieben werden können. Daher müssen andere Wege zur Bestimmung der *Ies* eingeschlagen werden.

Zwei Ansätze sind dabei bislang hauptsächlich verfolgt worden. Der erste besteht aus einer Glass-Box-Beschreibung der menschlichen Hörwahrnehmung. Eingangs- und Ausgangssignal des betrachteten Kodierers werden – nach einer Vorverarbeitung zur Beseitigung perzeptiv irrelevanter Signalmodifikationen – auf eine perzeptiv motivierte Beschreibungsebene transformiert. Auf dieser Ebene werden dann Ähnlichkeiten oder Abstände zwischen den beiden Signalen berechnet. Diese können in einen Qualitätsschätzwert transformiert werden, wie er sich in einem auditiven Test als mittleres Urteil der Versuchspersonen ergeben würde. Verschiedene solcher instrumentellen Qualitätsschätzer wurden in der Vergangenheit diskutiert (z.B. Berger, 1998; Hansen, 1998; Hauenstein, 1997).

Die zweite Möglichkeit besteht in einem Black-Box-Ansatz: die gesamte Kodier-Dekodier-Strecke wird, inklusive möglicher Übertragungsfehler, als eine *black box* betrachtet, für die ein Equipment-Impairment-Faktor in einem auditiven Test bestimmt wird. Hierzu wurde eine neue Methode erarbeitet, welche sicherstellt, dass die so bestimmten *Ies* sich in das bestehende System aus Impairment-Faktoren, wie es durch das E-Modell für viele etablierte Kodierer und für Kombinationen mit weiteren Störungsarten vorgegeben ist, einfügen. Diese Methode ist als neue ITU-T Rec. P.833 (2001) standardisiert worden und wird im Folgenden beschrieben.

3. Berechnung von Impairment-Faktoren aus den Ergebnissen auditiver Tests

Zur Bestimmung von Equipment-Impairment-Faktoren aus den Ergebnissen eines Hörversuchs ist es notwendig, die hervorgerufene Qualitätsbeeinträchtigung so zu quantifizieren, dass sich eine zu anderen Störungen konsistente Beschreibung ergibt. Für bereits etablierte Kodierer ist dies durch das E-Modell sichergestellt; eine detaillierte Untersuchung der verschiedenen Störungsarten findet sich z.B. bei Möller (2000). Bei neuen Kodierern, für die bislang noch keine *Ies* bekannt sind, lässt sich die quantitative Einordnung durch Vergleich mit Referenzkodierern erreichen.

Ausgangspunkt der Methode sind die Ergebnisse eines Hörtests, welcher neben dem betrachteten Kodierer auch weitere Referenzko-

dierer enthält. Die Referenzen sollten perzeptiv der betrachteten Beeinträchtigung ähnlich sein. In Rec. P.833 sind hierzu zwei Tabellen definiert: eine Listen von 14 Kodierern unterschiedlicher Prinzipien (ADPCM, LD-CELP, CS-ACELP, RPE-LTP, VSELP) und eine zweite Liste von 10 Kaskadierungen dieser Kodierer mit dem betrachteten neuen Kodierer. Die Testergebnisse liegen üblicherweise als MOS vor, das heißt als Mittelwerte über die Versuchspersonen-Urteile auf einer 5-stufigen ACR-Skala. Je nachdem, ob der ungestörte Fall oder auch Übertragungsfehler betrachtet werden sollen, werden die Ergebnisse in 3-5 Schritten zu $I_{e,s}$ transformiert.

Im ersten Schritt werden die MOS-Werte auf die *transmission rating scale* transformiert. Dies geschieht mit Hilfe der im E-Modell definierten, S-förmigen Kurve:

$$\begin{aligned} &\text{für } MOS = 1.0: && R = 0 \\ &\text{für } 1.0 < MOS < 4.5: && MOS = 1 + 0.035 \cdot R + R \cdot (R - 60) \cdot 7 \cdot 10^{-6} \\ &\text{für } MOS \geq 4.5: && R = 100 \end{aligned}$$

Im Ergebnis liegt für jede Testbedingung (neuer Kodierer oder Referenz) ein Beeinträchtigungswert auf der *transmission rating scale* (R) vor. Aus diesen Werten können nun erste experimentelle $I_{e,s}$ -Werte ($I_{e,sub}$) bestimmt werden, indem die Standard-Verbindung mit log-PCM nach G.711 als Referenz mit $I_e = 0$ definiert wird ($I_{e,sub} = R(G.711) - R(\text{Testbedingung})$).

Im zweiten Schritt wird für alle Referenzkodierer – für die ja bereits etablierte $I_{e,exp}$ definiert sind – eine Ausgleichsgerade durch die Wertpaare ($I_{e,sub}; I_{e,exp}$) gelegt, mit dem Kriterium einer minimalen quadratischen Abweichung:

$$I_{e,sub} = a \cdot I_{e,exp} + b$$

Mit Hilfe dieser Interpolationsgeraden kann für den betrachteten neuen Kodierer bereits ein stabiler $I_{e,exp}$ -Wert aus dem entsprechenden $I_{e,sub}$ berechnet werden. Durch Mittelung über drei verschiedene Eingangspegel wird die Stabilität dieses Wertes erhöht.

Im dritten Schritt wird der so bestimmte neue I_e -Wert auf seine Additivität zu den I_e -Werten anderer Kodierer überprüft. Hierzu werden für alle Kodier-Kaskaden der zweiten Referenztabelle aus P.833 die entsprechenden ($I_{e,sub}; I_{e,exp}$)-Paare in das Diagramm der Ausgleichsgeraden eingetragen. Wenn die Additivität erfüllt ist, so gruppieren sich die Paare um die Ausgleichsgerade herum. Größere Abweichungen von der Geraden werden vermerkt und ihre Ursachen untersucht.

Die Schritte 4 und 5 sind eine Wiederholung der Schritte 2 und 3, diesmal allerdings für alle Testverbindungen mit Übertragungsfehlern. Auf analoge Weise lässt sich dabei für jeden Kodierer und für jede Fehlercharakteristik ein eigener I_e -Wert bestimmen. Untersuchungen zeigen, dass die Beeinträchtigung sowohl vom verwendeten Kodierer als auch von der Häufigkeit und der zeitlichen Verteilung der Fehler sowie von angewandten Korrekturalgorithmen abhängt.

4. Evaluierung der Methode

Zur Evaluierung der Methode wurden bei der T-Nova zwei Hörversuche durchgeführt. Die dabei neben den in P.833 vorgeschriebenen Referenzen verwendeten Kodierer sind aus der Literatur bekannt, sodass ein Vergleich der experimentell abgeleiteten mit den etablierten I_e -Werten erfolgen kann. Der Vergleich ist in Abb. 1 und 2 in Form sog. Scatter-Plots illustriert. Auf der linken Seite sind jeweils die Ergebnisse für die 14 Referenzverbindungen und die Ausgleichsgerade gezeigt, auf der rechten Seite die Ergebnisse für die hieraus bestimmten „neuen“ I_e -Werte. Letztere sollten sich um die Mitteldiagonale gruppieren.

Die linken Grafiken zeigen deutlich, dass durch die Interpolation eine testabhängige Normalisierung erzielt wird. Die Ausgleichsgeraden unterscheiden sich hauptsächlich bezüglich des Achsenabschnitts, was einen unterschiedlichen experimentellen Bias vermuten lässt. Auf der rechten Seite gruppieren sich die nach P.833 ermittelten $I_{e,s}$ relativ gut um die Mitteldiagonale. Die experimentellen $I_{e,sub}$ sind fast ausschließlich geringer als die definierten Werte. Dies ist durch einen Sicherheitsabstand zu erklären, welcher bei der

Definition der Referenz- I_e -Werte veranschlagt wurde. Die beobachteten Abweichungen treten meistens bei den selben Kodier-Kaskaden auf. Dies deutet an, dass einfache Additivität bei Kaskadierung nicht ausreicht, sondern dass eine Maskierung der Beeinträchtigungen auftreten könnte. Glücklicherweise scheint diese Maskierung „gutarig“ zu sein; ein Aufschaukeln der Beeinträchtigungen ist nicht zu beobachten. Dadurch wird eine gewisse Planungssicherheit für Telefonnetze erhalten. Im Sinne einer ökonomischen Qualitätsplanung sind jedoch weitere Untersuchungen angebracht, um die perzeptiven Effekte kaskadierter Kodierer genauer zu erfassen. Hierzu könnte man sich auch der instrumentellen Qualitätsschätzer (Glass-Box-Ansatz) bedienen.

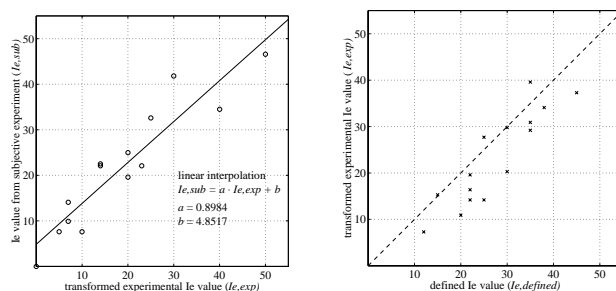


Abb. 1: Scatter-Plot der Ergebnisse von Test 1.

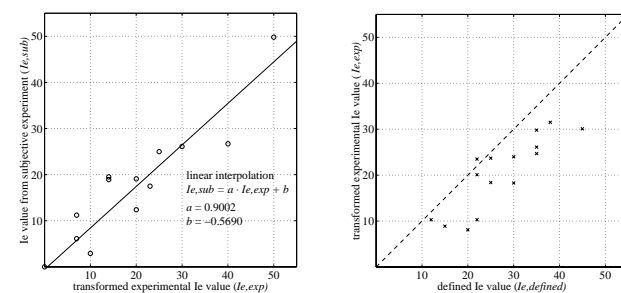


Abb. 2: Scatter-Plot der Ergebnisse von Test 2.

Danksagung:

Die vorliegende Arbeit entstand im Rahmen einer Kooperation zwischen dem IKA, Bochum (Prof. J. Blauert, PD. U. Jekosch) und der T-Nova, Berlin (H. Klaus).

Literatur:

- Bappert, V., Blauert, J. (1994). *Auditory Quality Evaluation of Speech-Coding Systems*, acta acustica 2, 49-58.
- Berger, J. (1998). *Instrumentelle Verfahren zur Qualitätsschätzung - Modelle auditiver Tests*, Dissertation, Christian-Albrechts-Universität, D-Kiel.
- Hansen, M. (1998). *Assessment and Prediction of Speech Transmission Quality with an Auditory Processing Model*, Dissertation, Carl-von-Ossietzky-Universität, D-Oldenburg.
- Hauenstein, M. (1997). *Psychoakustisch motivierte Maße zur instrumentellen Sprachgütebeurteilung*, Dissertation, Christian-Albrechts-Universität D-Kiel.
- ITU-T Rec. G.107 (2000). *The E-Modell, a Computational Model for Use in Transmission Planning*, International Telecommunication Union, CH-Geneva.
- ITU-T Rec. G.113 (2001). *Transmission Impairments*, International Telecommunication Union, CH-Geneva.
- ITU-T Rec. P.833 (2001). *Methodology for Derivation of Equipment Impairment Factors from Subjective Listening-Only Tests*, International Telecommunication Union, CH-Geneva.
- Möller, S. (2000). *Assessment and Prediction of Speech Quality in Telecommunications*, Kluwer Academic Publishers, USA-Boston.
- Möller, S., Raake, A. (2000). *Planung von Telefon-Sprachqualität: Unterschiedliche Modellansätze*, in: Fortschritte der Akustik – DAGA 2000: Plenarvortr. u. Fachbeitr. d. 26. Dtsch. Jahrestg. f. Akust., D-Oldenburg, Dtsch. Ges. Akust., D-Oldenburg, 208-209.