

## Vergleich des Messverfahrens PEAQ mit Hörversuchen in Produkttests

Jonas Fischer<sup>1</sup>, Gregor Feneberg<sup>2</sup>, Gerhard Krump<sup>3</sup>

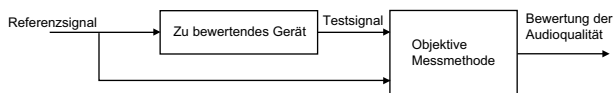
<sup>1</sup> Müller-BBM GmbH, 82152 Planegg, Deutschland, Email: Jonas.Fischer@MuellerBBM.de

<sup>2</sup> BBM Testlab GmbH, 82152 Planegg, Deutschland, Email: Gregor.Feneberg@BBM-Testlab.de

<sup>3</sup> Hochschule Deggendorf, 94469 Deggendorf, Deutschland, Email: gerhard.krump@fh-deggendorf.de

### Motivation und PEAQ

Bei vergleichenden Produkttests können viele Eigenschaften der getesteten Geräte abgemessen, berechnet oder durch etablierte Messverfahren objektiv bestimmt werden. Das Sammeln dieser Daten kann von einer Person relativ schnell erledigt werden. Etliche Produkte aus der Unterhaltungselektronik sind allerdings noch auf zeit- und ressourcenaufwändige subjektive Tests angewiesen, um in eine Rangfolge einsortiert werden zu können. Subjektive Größen wie beispielsweise Bild- oder Audioqualität von Geräten wie Fernseher, Kopfhörer oder Lautsprecher müssen von erfahrenen Probanden miteinander verglichen und subjektiv bewertet werden. Da diese Tests gleich mehrere Personen für längere Zeit binden und durch ihre subjektive Natur für Außenstehende eventuell nicht problemlos reproduzierbar sind, wäre eine Objektivierung wünschenswert. Im Bereich der Audioqualität von Consumergeräten konnte so etwas bis jetzt nicht praktisch umgesetzt werden. Um diese Lücke etwas kleiner zu machen, wird die Eignung des PEAQ-Verfahrens zur Bestimmung dieser Qualität untersucht. Das Akronym „PEAQ“ steht hierbei für „Perceptual Evaluation of Audio Quality“, was übersetzt soviel wie „Beurteilung der wahrgenommenen Audioqualität“ bedeutet. Ursprünglich wurde das PEAQ-Verfahren zur Untersuchung von Audiocodern in digitalen Kommunikations- und Übertragungsbereichen entwickelt. [1] Das Messverfahren gibt dabei idealerweise Messwerte aus, die denen eines subjektiven Hörtest nach der ITU-Empfehlung ITU-R BS.1116 entsprechen. [2] Dabei wird das aus einem Referenzsignal erzeugte Testsignal mit dem Referenzsignal verglichen. (Abbildung 1) Es besteht die Wahl zwischen einem „Basic“- und einem „Advanced“-Verfahren, die sich in Geschwindigkeit und Genauigkeit unterscheiden.



**Abbildung 1:** Basiskonzept für objektive Messungen, nach [1]

PEAQ basiert auf der Anwendung von verschiedenen psychoakustischen Grundlagen. Es werden beispielsweise lineare und nichtlineare Verzerrungen, der Abstand des Nutzsignals zu den Maskierungsschwellen, die harmonische Struktur sowie Veränderungen in der Modulation des Signals gemessen. Diese Daten werden über einzelne MOV (Model Output Values), also numerische Ausgangswerte, die zu einer objektiven Kenngröße zusammenge-

fasst werden, ausgegeben.

### Geräte- und Musikstückauswahl

Als ungefähren Querschnitt aus den in Produkttests hauptsächlich untersuchten Gerätenarten werden zwölf Geräte ausgewählt, darunter sechs Kopfhörer, drei Fernseher sowie drei Mobiltelefone, deren interne Lautsprecher als Beispiel für mindere Audioqualität dienen sollen. Drei Musikausschnitte werden zunächst verwendet, die einen weiten Bereich von verschiedenen Klangcharakteristiken abdecken sollen, ein instrumentales Klavierstück, ein rein A-Capella vorgetragener Liedausschnitt sowie ein instrumentales Fusion-Jazz-Werk.

### Aufnahme- und Hörtestverfahren

Die Aufnahmen finden in einem reflektionsarmen Halb-Freifeldraum statt, der dank Raum-in-Raum-Konstruktion sehr hohe Schalldämmwerte zur Außenwelt erreicht, so dass auch leisere Geräte wie beispielsweise Handylautsprecher ohne störende externe Einflüsse aufgenommen werden können. Für Kopfhöreraufnahmen wird ein Kunstkopf mit ITU-T 3.3 Ohrtypen nach ITU-T P.57, für die restlichen Aufnahmen ein rauscharmes 1-Zoll-Messmikrofon verwendet. Bei der Aufnahme der Kopfhörer muss darauf geachtet werden, dass diese fest und passend am Kunstkopf sitzen. Ein zu locker auf beziehungsweise eingesetztes Ohrstück verändert den Gesamtklang beträchtlich. Der Abstand des Messmikrofons zu den Geräten mit Lautsprecher beträgt 50 cm. Die Aufnahmen werden vor dem weiteren Vorgehen in Stereodateien mit identischen Rechts-/Linkskanälen umgewandelt. Zudem werden eventuell vorhandene Trittschallanteile entfernt und die Testfiles auf annähernd gleiche Lautstärke gebracht. Der anschließende Blind-Hörtest mit den aufgenommenen Testdateien wird angelehnt an das bei Produkttests üblichen Prozedere, es wird auf einer Skala von 5,5 (sehr gut) bis 0,5 (mangelhaft) in Schritten von 0,5 bewertet. Grund für diese Entscheidung ist die Erfahrung der Probanden in der Handhabung dieser Bewertungsweise sowie die grundsätzlich leichtere Durchführbarkeit als beispielsweise ein Hörtest nach ITU-R BS.1116 oder BS.1534.

### Erste Ergebnisse

Erste Vergleiche von subjektiven mit objektiven Noten ergeben bereits vernünftige Korrelationen, wobei die Basic-Version des PEAQ-Verfahrens, wie in Abbildung 2 gezeigt, besser abschneidet als die Advanced-Version.

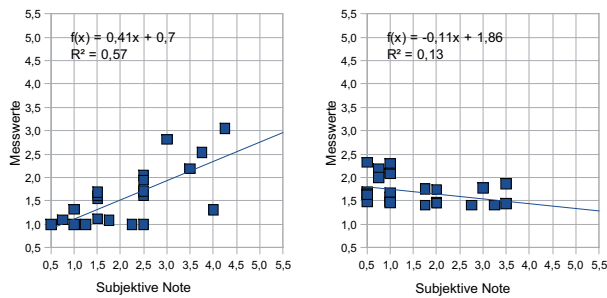


Abbildung 2: Korrelationsdiagramme des Vokalbeispiels, links: Basic-Version, rechts: Advanced-Version

### Modifikation der Messergebnisse

Da sich das Bestimmtheitsmaß der Messungen im Optimalfall an den Wert 1 annähern sollte, sind noch Überlegungen hinsichtlich der Optimierung der Ergebnisse anzustellen. Die Messung des Vokalstücks „Sid’s Boys Choir“ gibt mit einem Bestimmtheitsmaß von 0,57 im Vergleich mit der subjektiven Note das beste Ergebnis heraus, weshalb dieses Musikstück für die folgenden Untersuchungen in den Fokus genommen wurde. Da in den subjektiven Bewertungen Geräte mit fehlenden Bassfrequenzen weitaus niedrigere Bewertungen bekommen als in der objektiven Bestimmung, scheint die PEAQ-Implementierung im verwendeten Messgerät diesen Parameter nur unzureichend einfließen zu lassen. Ein manuelles Einrechnen der unteren sowie der oberen Übertragungsgrenzfrequenz, die mit Unter- beziehungsweise Überschreiten einer -20-dB-Grenze festgelegt wurde, konnte den Zusammenhang zwischen subjektiver Note und objektivem Messergebnis auf ein Bestimmtheitsmaß von 0,88 anheben. Eine zusätzliche Zuhilfenahme einer einzelnen MOV konnte das Bestimmtheitsmaß schließlich auf 0,9 erhöhen. (Abbildung 3)

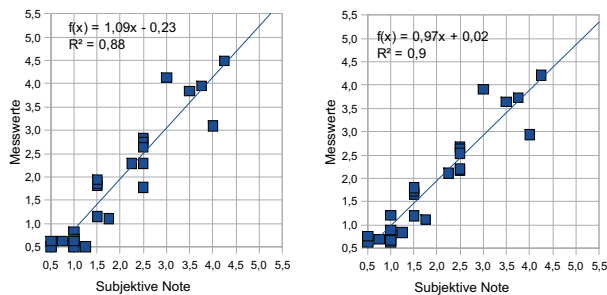


Abbildung 3: Korrelationsdiagramme, Subjektive Note / Mit Frequenzgang modifizierter PEAQ-Wert (links), Subjektive Note / Mit MOV modifizierter PEAQ-Wert (rechts)

### Zusammenfassung

Zusammenfassend lässt sich sagen, dass das PEAQ-Verfahren, etwas entgegen der eigentlichen Bestimmung, durch die Ergänzung eines Spektralkriteriums mit guten Korrelationen für Lautsprechermessungen verwendet werden kann. Zwar sind die Unterschiede innerhalb der einzelnen Geräteklassen noch nicht perfekt bestimmbar, doch die absolute Einteilung der Qualität ist als gelungen zu bezeichnen. Der Umfang der getesteten Geräte,

von hochwertigen kabelgebundenen Kopfhörern bis zu bestenfalls als lärmend zu bezeichnenden Handylautsprecher, wird gut wiedergegeben und bleibt unterscheidbar. (Abbildung 4)

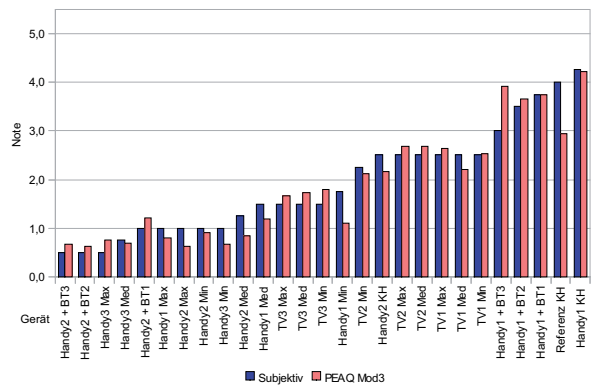


Abbildung 4: Sämtliche subjektive und objektive Ergebnisse

Ein Punkt, der die praktische Umsetzung des PEAQ-Verfahrens für diesen Einsatzzweck eventuell erschwert, ist der relativ hohe Aufwand, der für die Aufnahme der Testdateien betrieben werden muss. Neben der Einhaltung der rein technischen Voraussetzungen wie beispielsweise Dateiformat, Abtastungsrate, Kanalanzahl und Auflösung gilt es zusätzlich auf ein möglichst hoch eingeregelt Signal zu achten, das zudem möglichst wenig Verzögerung in Relation zur Originaldatei besitzen sollte. Die Gefahr bei hoch eingeregelt Signalen ist das Auftreten von Verzerrungen bei Pegelspitzen. Diese Spitzen sind oftmals nicht vorhersehbar oder nicht direkt am Anfang eines wiedergegebenen Musikstücks, was oftmals mehrfache Aufnahmevorgänge mit sich zieht.

Es bestehen momentan noch Einschränkungen hinsichtlich der praktischen Anwendbarkeit des Verfahrens. Für eine erhöhte Aussagekraft müssen erst weitere Untersuchungen mit weiteren Geräten und Testsignalen, darunter eventuell auch synthetisch erzeugte Audiodateien, angestellt werden und auch auf längere Sicht noch parallele Hörtests vorgenommen werden. Der Ersatz von Hörvergleichen ist noch nicht erreicht, aber die Unterstützung der subjektiven Eindrücke durch objektive Messungen dürfte in Kürze möglich sein.

Für eine Optimierung des Verfahrens bietet es sich an, nach weiteren unterstützenden Messungen zu suchen, die parallel zur Aufnahme der Testdateien vorgenommen werden können. So ist eine genaue Bestimmung des Frequenzgangs sicherlich förderlich für das Gesamtergebnis.

### Literatur

- [1] ITU-R Rec. BS.1387-1: Method for objective measurements of perceived audio quality. International Telecommunications Union, 2001.
- [2] ITU-R Rec. BS.1116-1: Methods for the subjective assessment of small impairments in audio systems including multichannel sound systems. International Telecommunications Union, 1997.