

Interview mit Karlheinz Brandenburg

Wir haben uns nie gegen die Freeware-Szene gewehrt

OBJEKTSpektrum spricht mit MP3-Erfinder Karl-Heinz Brandenburg.



MAINUSCH INTERVIEWT BRANDENBURG

Bei der CD wird ein 16-Bit-Signal auf den Kanälen Links und Rechts 44.000-mal pro Sekunde abgetastet. Daher die 1,4 Megabit pro Sekunde bei der CD. Wie funktioniert nun MP3? Soviel ich verstanden habe, durch Wegfiltern unwichtiger Signale ...

Genauer gesagt ist es die Frage, was sind denn für Frequenzanteile mit welcher Genauigkeit notwendig. Das funktioniert ähnlich, wie die Übersetzung bei uns im Gehör. Da findet auf der Hörschnecke ja auch so etwas wie eine Frequenzanalyse statt. Und die Signale gelangen dann über Nervenfasern ins Gehirn.

Also, wenn in der Gehörschnecke ein Sinneshäuschen durch den Schall stark rumzittert, dann kann man sagen, die Nachbarhäuschen zittern auf jeden Fall etwas mit und deshalb merkte das Gehirn gar nicht, wenn zusätzliche Töne die Nachbarhäuschen zum Zittern bringen würden.

Das ist eine vereinfachte, aber korrekte Beschreibung des Maskierungseffektes. Und das wird bei MP3 in Modelle gegossen, die Leute schon Jahrzehnte vor der MP3 im Rahmen der Psychoakustik erarbeitet hatten. Das war Grundlagenforschung, damals noch ohne sichtbare Anwendung.

Sind Sie auf die Idee gekommen, dieses psychoakustische Verfahren mit dem Weglassen von Frequenzanteilen zu benutzen?

Nicht nur ich: Da lagen eine ganze Menge Dinge in der Luft. Es gab von dem verstorbenen Professor Zwicker aus München ein Buch über Psychoakustik, in dem die Grundideen beschrieben waren. Es gab von Kollegen aus den USA ein Paper, in dem die Idee beschrieben wurde, solche Dinge zu verwenden, um die Qualität von Sprachcodierung zu beurteilen. Es lag viel auf dem Tisch. Erst nach vielen Versuchen kamen wir in die Nähe des MP3-Standards. Eine wichtige Erkenntnis war, mit wie wenigen Bits Genauigkeit wir bestimmte Frequenzanteile codieren

Johannes Mainusch: Hallo Herr Brandenburg, sind Sie ein Vinyltöter? Trifft es Sie, wenn Musiker wie Neil Young MP3 kritisieren?

Karlheinz Brandenburg: Ich fühle mich nicht getroffen. Aber es ist unfair gegenüber all den Audio-Experten und den Tonmeistern, die damals als professionelle Testpersonen mitgearbeitet haben und die Qualität des MP3-Standards für gut genug befanden.

1974 haben Sie angefangen zu studieren, 1982 gab es eine Promotionsstelle und dann haben Sie bis 1989 promoviert. 1991 waren dann wesentliche Elemente Ihrer Promotion in dem Entwurf zum MP3-Standard wiederzufinden. Warum hat das so lange gedauert?

Als wir 1982 starteten, gab es keinerlei Ressourcen. Die Compact Disc (CD) gab es 1982 noch nicht zu erschwinglichen Preisen. Und es gab auch keine Möglichkeit, die Daten in einen Rechner hinein und wieder heraus zu bekommen. Die ersten Versuche machten wir auf einem digitalen Signalprozessor mit Bit-Slice-Technik, der AMD 2900er-Serie, und in meiner Diplomarbeit habe ich auf diesem System eine FFT implementiert ...

eine Fast Fourier Transformation ...

die dann sehr schön schnell funktioniert hat.

Und was konnte man damit machen?

Das war zunächst so zum Spaß und um zu sehen, was man mit dem System machen konnte. Und das war dann auch gleich etwas, womit man mit ein paar weiteren Dingen Audio-Codierung betreiben konnte. Das war 1983 ein allererstes System, auf das man mit einem Mini-Computer über ein digitales Interface Musik aufspielen konnte. Dann wurde der Rechner mit einem analogen Potenziometer hochgedreht, bis kurz bevor er ausgestiegen ist, und dann konnte man bei etwa 30 Kilohertz Abtastfrequenz Echtzeitverarbeitung von Musiksignalen machen. Und sich das dann gleich wieder anhören. Manche Musik hörte sich darüber sogar schon ganz gut an.

Etwas später bekamen wir einen neuen leistungsstärkeren Mini-Computer, mit dem man Simulationen mit etwas höherer Geschwindigkeit durchführen konnte, um bessere Tonqualität zu erreichen. 1986 hatten wir dann einen Durchbruch mit neuen Varianten, die man als direkten Vorgänger zu MP3 ansehen kann.

können, ohne dass ein Qualitätsverlust hörbar ist. Auch Rauschen muss ich nicht mit 16 Bit Genauigkeit wiedergeben, das geht deutlich sparsamer.

Also, eine Frequenzanalyse eines kleinen Zeitblocks eines Musikstücks. Dann sieht man für diesen Zeitblock die lauten Frequenzen und blendet alles aus, was nicht laut ist. So spart man viele Bits im Signal und das Ohr hört den Unterschied nicht, korrekt?

Ja, das ist die sogenannte Grundhörschwelle. Wichtiger noch ist die Mithörschwelle, die sagt, wie genau ich in der Nähe sehr starker Signale codieren muss. Die Nachbarfrequenzen werden dort eh nicht gehört. Da können Bits gespart werden. Die Aufgabe war, so eine sehr kompakte effiziente Darstellung zu erlangen. Die ursprünglichen Verfahren hatten eine feste Anzahl Bits pro Fragment, und gegen Ende der 80er-Jahre kamen wir dann auf die Idee, das auch fallen zu lassen und adaptiv anzupassen, um eine noch höhere Datenkompression zu erlangen. An dieser Grundaufgabe wurde an verschiedenen Stellen weltweit gearbeitet.

„Wichtiger als die Grundhörschwelle ist die Mithörschwelle“

Was ist denn Psychoakustik?

Psychoakustik ist interdisziplinär, die Methoden sind die einer technischen Beschreibung von Schall und der subjektiven Wahrnehmung von Hörsignalen. Die Parameter dafür bekommt man aus Hörtests.

Also eine Paralleldisziplin zu dem, was Sie geforscht haben?

Psychoakustik gab es in Erlangen nicht. In Deutschland gab es Mitte der 80er-Jahre allerdings besonders viele Universitäten, die Psychoakustik beforchten. Das gab uns am Anfang einen Vorsprung.

1982 Start der Arbeit, 1998 der erste tragbare MP3-Spieler? Und da gab es keinen Heureka-Moment? Den einen Moment mit dem Durchbruch?

Es gab viele Personen und technische Zutat, die da zusammen kamen. Für meine persönliche Geschichte gab es einen Heureka-Moment im Februar 1986. Da hatte ich vorher gerade einen Zwischenbericht für die Deutsche Forschungsgesellschaft geschrieben. Da kam mir die Idee, das Paradigma der „Analysis by Synthesis“ aus der Sprachcodierung auf die Audio-

Codierung zu übertragen und damit die Qualität zu verbessern. Die Idee war, diese iterative Strategie und Huffman-Codierung zu kombinieren.

Das habe ich damals innerhalb einer Woche so weit abgetestet, dass ich wusste, es funktioniert. Dann habe ich das meinem Doktorvater gezeigt, der meinte, das solle ich mir patentieren lassen. Die Kollegen im Labor waren auch angetan. Dieses neue Verfahren hat vormals vorhandene Probleme so elegant umschifft, dass es einem Durchbruch gleichkam. Das neue Verfahren ermöglichte uns, sehr viel schneller neue Varianten zu testen und auf Erfolg zu prüfen.

Haben Sie ein Testmusikstück gehabt?

Wir hatten 20 Sekunden lange Musikstücke. Mehr ging damals nicht auf den Speicher unseres ersten Computers. Etwa 5 Megabyte Daten auf eine große Platte. Um die Daten ausgeben zu können, mussten wir eigene Treiber schreiben. Es hat etwas gedauert, bis wir Musik direkt auf den Computer bringen und wieder anhören konnten. Dann gab es den Versuch, automatisiert die Maskierungsfunktionen aus dem Gehör nachzubilden, um so einen Hinweis zu bekommen, wo man genauer hinhören muss. Ein Durchlauf auf unserem MicroVAX II-Computer für ein Musikstück von 20 Sekunden, also encodieren, decodieren und Messverfahren laufen lassen, dauerte 4 Stunden.

Das heißt, ein Fehler kostete einen den halben Tag?

Wir verwendeten daher immer kurze Stücke. Für die Dissertation gab es dann einen richtig langen Lauf, der 6 Wochen CPU-Rechenzeit brauchte.

So haben Sie also eine Dissertation über etwas geschrieben, was man hören muss. Hatten Sie das Ergebnis bei der Dissertation mit.

Tatsächlich habe ich Ende 1988 beim ersten Treffen der MPEG Audio Gruppe [Anm. d. Red. MPEG=Moving Picture Experts Group] einen schweren Koffer mit dem Signalprozessor DSG 56000, Hardware und Lautsprechern mitgenommen und damit den Leuten die Ergebnisse vorgeführt. Danach habe ich das System mit nach Hause genommen und hatte somit den ersten MP3-Spieler zu Hause. Dann habe ich den Rest des Wochenendes codiertes Material gehört und war erschrocken, wie viel doch nicht so gut geklungen hat.



Prof. Dr.-Ing. Dr. rer. nat. h.c. mult. Karlheinz Brandenburg

Alter: 62

Hatte 1988 das erste MP3-System zu Hause (für ein Wochenende)

Anzahl Titel: sieben – vier Doktoren, zwei Diplome und eine Professur

Arbeitgeber: 1. TU-Ilmenau, 2. Fraunhofer-Institut für Digitale Medientechnologie IDMT

Studium: Elektrotechnik und Mathematik, Universität Erlangen-Nürnberg

Google-Hits für „Karlheinz Brandenburg“: ungefähr 73.900

Und wo steht diese Kiste heute?

Ein solches System steht heute im Deutschen Museum in Bonn.

Und dann sagt der Professor, nun mach mal ein Patent darauf.

Ja, der Anmeldetag war dann im August 1986. Dieses erste Patent, da hat mein Doktorvater gesagt, das ist Ihres, und damit lief das auch auf meinen Namen.

Ist das denn dann nach den späteren Patenten noch etwas wert gewesen?

Ja natürlich, die bauen alle aufeinander auf. So funktionieren Patente, dass immer die ganzen verschiedenen Ideen, die gebraucht werden, um etwas zu bauen, da dokumentiert sind. Es ist der Normalfall, dass Dutzende bis Hunderte Patente im Laufe der Zeit angemeldet werden.

Also, die Altersversorgung ist gesichert?

Ja, die Altersversorgung ist gesichert. Patente laufen leider aus, dieses 1986er Patent ist schon lange ausgelaufen. Und auch für die anderen Patente ist es so, dass nach deutschem Gesetz die Erfinder einen Anteil an den Lizenzentnahmen bekommen. Das war bei uns so, dass alle beteiligt wurden, die irgendwie beteiligt waren. Also bei MP3 eine Gruppe von etwa 40 Personen mit durchaus unterschiedlichen Anteilen. Für mich hat sich das so ausgezahlt, dass ich meine eigene kleine Venture-Capital-Firma habe. Und dass ich nach wie vor keinen Porsche oder solche Sportwagen fahre, weil das nicht mein Stil ist.

Aber wenn es sein müsste, dann könnte ich mir den leisten.

„Ich könnte ich mir einen Porsche leisten, aber das ist nicht mein Stil“

Auf dem Bild, das Sie uns mitgeschickt haben, sind acht Leute drauf. In Wikipedia werden nur sechs Leute erwähnt. Was stimmt denn nun?

Das ist ein dauerndes Hin und Her, wer da nun tatsächlich ins Kernteam gehört und wer sonstige Beiträge geliefert hat. Ich mag da immer eher eine längere Liste aufzählen. Da gehörten ja auch noch wissenschaftliche Mitarbeiter dazu. Vorangetrieben hat das maßgeblich Professor Gerhäuser, damals noch Dr. Gerhäuser. In seiner Doktorarbeit entstand der gro-

ße Signalprozessor, den ich anfangs programmierte. Es gab viele Spezialisten für einzelne Teile des Codes. Ich war in der Arbeitsgruppe derjenige, der die übergreifende Klammer bildete.

Wann war denn der Moment, als Sie ahnten, das wird richtig groß?

Das waren mehrere Situationen, ich wurde beispielsweise 1988 gefragt, was denn aus den Dingen, die ich an der Uni mache, so werden kann. Damals sagte ich, entweder verstaubt meine Dissertation in der Bibliothek oder das wird ein Standard, denn Millionen Menschen verwenden. Dass ich damit um einen Faktor 1000 daneben lag, wusste ich damals nicht.

Also 1988 hatten Sie schon eine Idee vom zukünftigen Erfolg, aber richtig erahnten Sie es noch nicht?

Ja, denn in den frühen 90er-Jahren, als der Standard fertig war und wir aus dem Profi-Bereich die ersten Anwender hatten, ist der MPEG-1 Audio Layer 2 [=MP2, Anm. d. Red.] mit ersten Consumeranwendungen voraus gezogen. Wir mit MP3 bekamen hier und da öffentlichen Spott ab, da die Zahl der Geräte, die unseren MP3-Standard verstanden, immer noch zweistellig war. Das Dumme war, die Spötter hatten recht.

1993 wechselte ich von der Universität zum Fraunhofer-Institut für Integrierte Schaltungen IIS und war als Abteilungsleiter für das wirtschaftliche Wohlergehen des vergrößerten Teams verantwortlich.



Ich treffe heute immer noch Kolleginnen und Kollegen aus der Zeit, die sagen, sie haben damals nicht verstanden, wieso ich so optimistisch war. Dann verbreiterte sich so Stück um Stück die Zahl unserer Anwender. Die erste Softwarefirma war Marcomedia. Etwas später kam der Kontakt zu Microsoft und die erste Windows-Lizenz wurde 1996 ausgegeben. 1997 startete die Diskussion über nicht autorisierte Verbreitung von MP3-codierter Musik. Es gab damals die ersten FTP-Sites, über die Musik zum Download bereitstand.

Der MP3-Encoder wurde dann von einem Australier geknackt und veröffentlicht ...

Zu knacken war da nicht viel. Wir waren wohl ein bisschen blauäugig. Wir hatten für die Encoder für professionelle Zwecke einfach ein Interface verwendet, das von Microsoft für solche Software vorgegeben war. Der Australier hat dann mit einer gestohlenen Kreditkartennummer dieses Softwarepaket eingekauft, sich das genauer angeschaut, ein neues Benutzerinterface geschrieben, unsere Encoder-DLL dazugepackt und das auf einem Umweg über Schweden auf amerikanischen Servern bereitgestellt. Sein Paket hatte ein Readme, darin stand: „Das ist Freeware, Danke an Fraunhofer“.

Das war ja dann nicht Freeware.

Nein. Wir wurden das auch gefragt und wir haben dann allen gesagt, es handelt sich um gestohlene Software. Solche Dinge sind aber letztendlich nicht abzustellen.

Hat das nicht letztlich der Verbreitung des Standards geholfen?

Im Nachhinein kann man sagen, die Bedingungen bei MPEG waren FRAND [Anm. d. Red. *Fair, Reasonable and Non-Discriminatory*]. Uns war es eigentlich



MP3-Team 1987



verboten, den MP3-Encoder den Einen kostenlos zu geben und von Anderen Geld zu verlangen. Das hat der Australier uns dann abgenommen. Das heißt, wir konnten immer klar sagen, das war nie frei. Und wir haben uns nie massiv gegenüber der Freeware-Szene aufgestellt. Für uns waren die professionellen Nutzer und die Firmen interessant, die Chips und Geräte bauten.

Wie ist das heute, gibt es heute „freie“ MP3-Encoder?

Es gibt Versionen, die frei verteilt werden.

Es ist also eine Grauzone?

Ja, da gibt es eine Grauzone. Die wird irgendwann vorbei sein, wenn die Patente auslaufen. Dazu geben wir keine offiziellen Statements ab, denn es ist etwas komplizierter, da alle Patente länderbezogen sind und es daher keine einheitlichen Regeln gibt. Im Windows-Betriebssystem ist MP3 im Übrigen bereits mit dem Betriebssystem lizenziert.

Wie ist es denn im Linux-Umfeld? Gab es die Überlegung, das als Open Source bereitzustellen?

Open Source heißt ja nicht patentfrei. Es gab schon früh die Überlegungen, die Adaption des Standards durch Beispiel-Sourcen zu unterstützen. Aus einer dieser Beispiel-Sourcen wurde der spätere LAME-Encoder geschrieben [Anm. d. Red.: LAME ist ein Open-Source-Projekt, das sich selbst als Entwicklungsprojekt zur Unterstützung des MP3-Audioformats versteht.]. Der ursprüngliche LAME war ein Patch auf fehlerhafte MP3-Sourcen.

So gegen Ende der 90er-Jahre hat MP3 sich dann stark verbreitet, nach 16 Jahren Forschung. Ist Geduld eine Ihrer Tugenden?

Eher Zähigkeit. Vor etwa 15 Jahren schrieb ein Journalist in der Wochenzeitung „Die Zeit“ einen Satz, den ich nun selbst immer gern verwende: „Der Erfolg des Teams in Erlangen war eine Kombination aus Vision, Überstunden und Sturheit.“

„Der Erfolg des Teams war eine Kombination aus Vision, Überstunden und Sturheit“

Wie ist es dann mit Ihnen weiter gegangen?

Auf Umwegen und über Professor Gerhäuser kam das Ansinnen zu mir, dass das Land Thüringen in Ilmenau ein Fraunhofer-Institut ansiedeln möchte. Das war damals die Chance, neu starten zu können mit einer Anschubfinanzierung, wie sie es damals in den Zeiten des Internet-Hypes für ein Start-up gegeben hätte, und trotzdem in Verbindung mit den alten Freunden in Erlangen zu bleiben. Und diese Zusammenarbeit funktioniert immer noch. Heute bin ich an der TU Ilmenau als Professor angestellt und Teil meines Arbeitsvertrages ist es, das Fraunhofer-Institut für Digitale Medientechnologie IDMT voranzubringen.

Was beschäftigt Sie heute?

Zurzeit ist mein Steckenpferd, genauer herauszufinden, wie Gehör und Gehirn funktionieren, wenn es um räumliches Hören geht. Der klassische Kunstkopf hat beispielsweise das Problem, dass er

nicht immer funktioniert. Wenn man den Kopf dreht, dann nimmt unser Gehör die Schallwellen gefiltert auf. Wenn man sich das näher anschaut, das sind die sogenannten head-related transfer functions, dann sind die Peaks und Täler ohne Weiteres 20 Dezibel auseinander. Unser Gehirn ist darauf eingespielt, dass ich sogar ein Ohr verschließen kann und immer noch räumlich höre.

Was kann man mit diesen Erkenntnissen machen?

Die perfekte räumliche Klang-Illusion. Weiterhin interessiert uns, was man mit vielen Lautsprechern machen kann. So wäre es zum Beispiel interessant, eine Lautsprecherwand im Großraumbüro zu installieren, um damit störende Geräusche wegzufiltern.

Herr Brandenburg, vielen Dank für das Gespräch.

Das Interview führte ...



Dr. Johannes Mainusch

(johannes.mainusch@kommitment.biz)

Berater für Unternehmen, die Bedarf im Bereich IT, Architektur und agiles Management haben. Dr. Mainusch ist seit 2012 Mitglied der OBJEKTSpektrum-Redaktion.