

Experimentieren mit digitalem Ton

Mattias Eisenbarth, Monika Di Angelo

Technische Universität Wien, Institut für Rechnergestützte Automation
mattias.eisenbarth@gmx.at

Die meisten heutigen Handys haben eine mp3-Player-Funktion, die von den Jugendlichen gerne und häufig benutzt wird. Den wenigsten Jugendlichen ist allerdings bewusst, wie Musik (oder Töne und Geräusche i.a.) digitalisiert wird, um am PC, Internet oder in entsprechenden Playern verwendet zu werden. Im Folgenden wird gezeigt, wie (digitaler) Ton entsteht, welche Komponenten der Digitalisierung den Speicherplatz für Musik beeinflussen und wie die Digitalisierung selbst mithilfe von kostenlosen Multimedia-Tools aus dem Internet im Unterricht veranschaulicht werden kann.

1 Vorbemerkung

Der Informatikunterricht bietet eine gute Gelegenheit, um auch fächerübergreifend arbeiten zu können. So können Inhalte des Physik-, Biologie-, Musik-, Sprachunterrichts usw. mit dessen Hilfe näher erörtert und zum besseren Verständnis aufgearbeitet werden. Doch auch für Schülerinnen und Schüler interessante Fragestellungen des Alltags lassen sich im Informatikunterricht beantworten. Das Internet ist voll von Musik, Videos und Bildern, die unterschiedliche Formate besitzen und verschieden aufgebaut sind. Doch eines haben diese Daten, die wir tagtäglich benutzen, allesamt gemein: sie sind digital. Was macht denn nun aus einer Folge von Nullen und Einsen Musik, was macht die Daten zu Bildern? Mit anderen Worten stellt sich die Frage welche Informationen die Daten beinhalten müssen, damit sie als das dargestellt werden können, was sie sind.

2 Der Schall - Entstehung von Tönen und Klängen

Als Schall wird die wellenförmige Ausbreitung von Druckschwankungen in elastischen Medien bezeichnet. Dieser kann mathematisch durch den zeitlichen Schwingungsverlauf der Schallwelle (Frequenz, Amplitude, Schallgeschwindigkeit usw.) beschrieben werden. Sobald eine Schallwelle auf das Ohr eines Lebewesens trifft, kann sie von diesem wahrgenommen werden. Doch nicht jedes Lebewesen hört „gleich gut“. Der menschliche Hörbereich liegt in einem Schallfrequenzbereich von 20 Hz bis 20 kHz und einem Schalldruckbereich von 10^{-5} Pa bis 20 Pa .

Den Schalldruck empfindet der Mensch als Lautstärke. Übersteigt die Lautstärke einer Schallwelle den Druck von 20 Pa , so ist dies für uns bereits schmerzhaft. Da die Angabe des Schalldrucks in Pascal als Richtwert für die Lautstärke wegen des großen Hörbereichs von der Hörschwelle bis zur Schmerzgrenze (1:2.000.000) ungeeignet ist, wurde der Schalldruckpegel eingeführt. Dieser wird in Dezibel [dB] gemessen und ergibt sich aus der logarithmischen Transformation des Schalldruckverhältnisses zwischen dem Druck der gehörten Schallwelle und einem Bezugsschalldruck. Betrachtet man den menschlichen Hörbereich unter dem Blickwinkel des Schalldruckpegels, so liegen die Hörschwelle bei 0 dB und die Schmerzgrenze bei 120 dB . Der Logarithmus in der Schalldruckpegeldefinition ist auf ein

Phänomen der menschlichen Wahrnehmung zurückzuführen, die sowohl im auditiven als auch im visuellen Bereich logarithmisch stattfindet.

Innerhalb des menschlichen Hörbereichs (Hörfläche) befinden sich weitere Teilbereiche, die für die Wahrnehmung – und in weiterer Folge für die digitale Verarbeitung von auditiver Information – wichtig sind. So nehmen wir Musik beispielsweise in einem Bereich von ca. *50 Hz* bis *10 kHz* wahr, Sprache in einem Bereich von rund *180 Hz* bis *5 kHz*. Je nachdem für welche Aufgaben eine computertechnische Verarbeitung von Audiodaten Anwendung findet, müssen Frequenzen, die außerhalb dieser Hörbereiche liegen, nicht mehr berücksichtigt werden. Ein Telefon hat so zum Beispiel die Aufgabe Sprache zu übermitteln – ob während des Gesprächs auch Musik in guter Qualität übertragen werden kann, ist dabei nicht relevant.

Eine Schallwelle, die die Form einer einzelnen Sinusschwingung hat, wird Ton genannt. Ein Ton kann in beliebiger Frequenz und Amplitude auftreten. Überlagert man einen Ton (Grundton) mit weiteren Tönen (Obertönen), deren Frequenzen ein ganzzahliges Vielfaches der Frequenz des Grundtones sind, so entsteht daraus ein harmonischer Klang. Wenn man hingegen keinen Grundton mehr erkennen kann und ein Schallereignis weder als Ton noch als Klang einordnen kann, dann wird es als Geräusch bezeichnet. Geräusche sind aperiodische Funktionen, deren Struktur sich im Zeitverlauf ändern kann.

In der Informatik ist die Digitalisierung von analogem Schall von großer Bedeutung. Viele moderne Tonträger speichern Audioinformationen heute ausschließlich im digitalen Format. Bei der Digitalisierung wird das analoge Audiosignal in sehr kurzen, diskreten Zeitabständen abgetastet. Das Verfahren – das Sampling genannt wird – speichert an jedem Abtastpunkt den Wert der Amplitude der analogen Schallwelle. Die Amplitude liegt zu diesem Zeitpunkt bereits in Form einer kontinuierlichen elektrischen Spannung vor, die in einem vorhergehenden Schritt erzeugt werden muss (z.B. mit einem Mikrofon). Ziel der Digitalisierung ist es nun, genügend Samples zu sammeln, um aus dem diskreten digitalen Audiosignal wieder eine kontinuierliche analoge Schallwelle erzeugen zu können. Nach dem Nyquist-Shannonschen Abtasttheorem muss die Abtastfrequenz mindestens doppelt so hoch sein wie die höchste vorkommende Frequenz im analogen Signal, damit dieses wieder annähernd auf seine ursprüngliche Form zurückgerechnet werden kann. Für den gesamten menschlichen Hörbereich bedeutet dies, dass ein analoges Tonsignal mindestens 40.000-mal pro Sekunde abgetastet werden muss, damit wir es später wieder als dasselbe Signal erkennen können. Zum Digitalisieren von Sprache reicht eine wesentlich niedrigere Abtastrate, was auch den Speicherplatzbedarf reduziert.

Das Prinzip der Digitalisierung von analogem Tonmaterial – und in weiterer Folge der Tondatenkompression (Stichwort: mp3) – ist in Zeiten von Musikdownloads, iPods, mp3-Radios usw. für Schülerinnen und Schüler vor allem durch den nahen Bezug zum Alltag interessant. Mehr zur Digitalisierung kann in [EI10] nachgelesen werden.

3 Einsatz von kostenlosen Audio-Tools im Unterricht

Heute existiert bereits eine Fülle an multimedialen Lernumgebungen, die pädagogisch ausgearbeitet sind und eine Menge an Themen für verschiedenste Unterrichtsgegenstände behandeln. Solche Anwendungen sind aus vielerlei Hinsicht für den Unterricht von Vorteil. Leider sind sie aber einerseits nur für ein bestimmtes Unterrichtsgebiet ausgelegt und andererseits in den meisten Fällen nicht kostenlos.

Eine Alternative zu den vorgefertigten Lernprogrammen sind freie Multimedia-Tools aus dem Internet, mit denen man den Theorievortrag über digitales Tonmaterial (aber auch digitale Bilder oder Videos) im Unterricht anschaulich und praktisch unterstützen kann. Als kos-

tenlose Audio-Tools bieten sich hierfür beispielsweise *Audacity* [AU10], *WavePad* [NC10a] oder der *Free Audio Editor* [FR10] an. An Video-Tools sind besonders *Avidemux* [AV10], *VirtualDub* [VI10] oder der *VideoPad Video Editor* [NC10b] zu nennen.

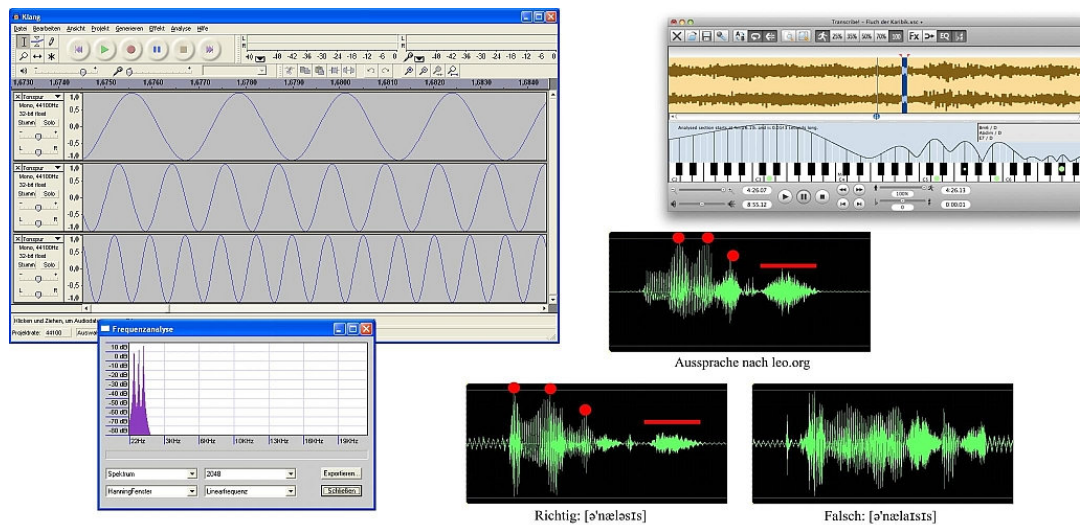


Abbildung 1: Darstellungsmöglichkeiten von Ton mit freien Audio-Tools im Unterricht

Mit den genannten Audio-Tools lassen sich auf einfache Weise digitale Audiodaten grafisch darstellen und manipulieren. So kann Schülerinnen und Schülern der Unterschied zwischen Tönen und Geräuschen präsentiert, die Erzeugung von Klängen und die Darstellung von Musik im Frequenzspektrum (Notensystem) gezeigt, praktische Beispiele zum Audio-Sampling demonstriert, sowie das Aufzeichnen und Visualisieren der Aussprache im Sprachunterricht beigebracht werden. Dies sind nur einige wenige Anwendungsbeispiele von frei verfügbaren Audiowerkzeugen aus dem Internet.

Natürlich setzt die Arbeit mit solchen Tools im Unterricht voraus, dass man sich auch als Lehrer/in auf den Einsatz derartiger Mittel vorbereitet und mit einem entsprechenden Konzept in die Klasse geht. In [EI10] sind einige Beispiele für den Unterricht mithilfe von Multimedia-Tools angeführt, die sich noch beliebig erweitern lassen.

4 Digitalisieren von analogem Ton – ein Beispiel für den Unterricht

Zur Vermittlung des Konzeptes der Digitalisierung von analogem Ton eignet sich besonders das Audio-Tool *Audacity*. Im Folgenden wird ein Unterrichtsbeispiel vorgestellt mit dem sich die Entstehung von Signalfehlern (Aliaskomponenten) durch das Verwenden ungeeigneter Abtastraten gut demonstrieren lässt.

Aliaskomponenten lassen sich nicht nur bei der Digitalisierung von Ton, sondern auch im Alltag im visuellen Bereich beobachten. Die Räder vorbeifahrender Autos können in der Nacht für den Menschen so aussehen, als ob sie sich rückwärts bewegen würden. Dies liegt an der künstlichen Straßenbeleuchtung, die mit Wechselstrom betrieben wird. In Wahrheit leuchtet ein solches Licht nämlich nicht kontinuierlich, sondern schaltet sich (bei Wechselstrom mit 50 Hz) 100-mal in der Sekunde aus. Jedes Mal wenn das Licht wieder eingeschaltet wird „tastet“ unsere visuelle Wahrnehmung den Reifen des vorbeifahrenden Fahrzeuges ab. Die dunklen Phasen der Beleuchtung nehmen wir nicht wahr, sie fehlen unserer Wahrnehmung daher. Im Prinzip wird die kontinuierliche Bewegung des Reifens vom Menschen nur noch in diskreten Schritten wahrgenommen. Wenn sich die Wechselstromfrequenz des Lichts und die Drehzahl des Reifens in einem ungünstigen Verhältnis zueinander befinden, dann

entsteht dieser optische Wahrnehmungsfehler – eine Aliaskomponente. Dasselbe Phänomen beobachten wir bei Ton: Eine Aliaskomponente entsteht, wenn das abzutastende analoge Tonsignal mit einer geringeren als der Nyquist-Shannonschen Abtastfrequenz abgetastet wird.

Zur Durchführung der Übung startet man ein *Audacity*-Projekt und erzeugt drei Tonspuren (Menü *Projekt*). Die Standardsamplerate von Audacity beträgt 44,1 kHz, die Standardsamplegröße 32 bit (float). Eine CD-Audio verwendet vergleichsweise 44,1 kHz als Samplerate und eine Samplegröße von lediglich 16 bit. Die Tonqualität für diese Übung ist somit mehr als ausreichend. Im nächsten Schritt ändert man die Samplerate jeder Tonspur (Klick auf den „Unten-Pfeil“ bei der jeweiligen Tonspur → *Samplerate einstellen*). Am besten eignen sich Extrembeispiele indem man die Abtastrate der ersten Tonspur auf 96 kHz setzt, was die erforderliche 40-kHz-Samplerate für den menschlichen Hörbereich bei weitem übersteigt. Die Abtastraten der zweiten und dritten Tonspur werden auf 44,1 kHz (CD-Audio) bzw. 8.192 Hz gesetzt.

Nun fügt man jeder Tonspur einen Sinuston mit maximaler Amplitude und beliebiger Dauer hinzu (Menü *Generieren*). Die Frequenz eines jeden eingefügten Tones wird auf 10.000 Hz festgelegt. Um den 10-kHz-Ton korrekt zu digitalisieren, ohne dabei eine Aliaskomponente zu erzeugen, ist eine Abtastrate von mehr als 20 kHz notwendig (Nyquist-Shannon Theorem). Demnach dürften nur die erste und die zweite Tonspur den 10-kHz-Ton richtig wiedergeben. Wenn man mithilfe der Zoomfunktion von *Audacity* weit genug in die drei Tonspuren hineinzoomt (siehe Abbildung 2), kann man das Signal auch optisch gut erkennen. Dabei kommen die diskreten Abtastpunkte des digitalen Signals zum Vorschein. Je höher die Samplerate ist, desto näher liegen die Punkte beieinander. Verbindet man die Abtastpunkte miteinander, so erhält man wieder ein kontinuierliches analoges Signal.

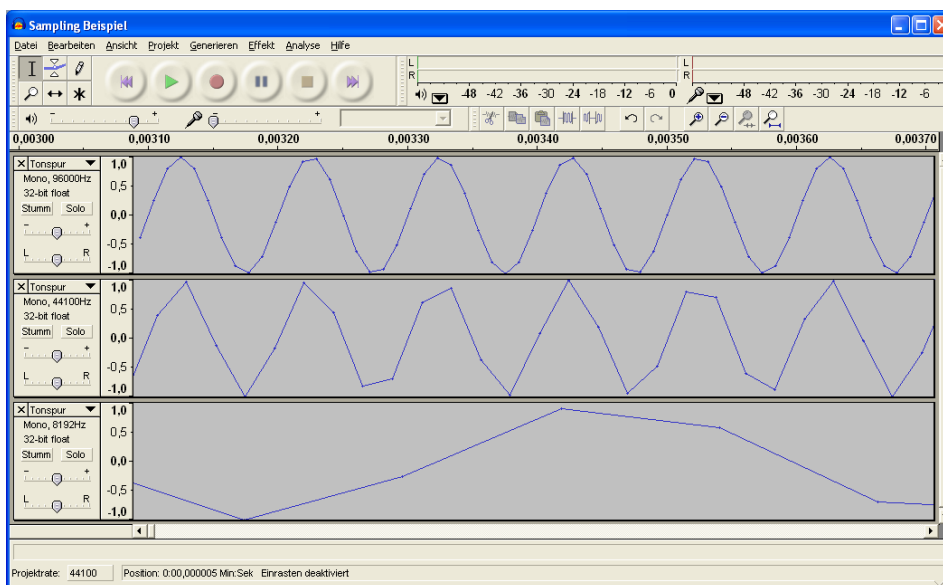


Abbildung 2: Demonstration der Aliaskomponente (erzeugt mit *Audacity*)

Den Schülerinnen und Schülern kann das Phänomen der Aliaskomponente sowohl optisch als auch akustisch demonstriert werden. Wie man bereits anhand der Darstellung der Tonspuren erkennen kann, liefert eine Samplerate von 96 kHz (Tonspur 1) die genaueste Annäherung an die ursprüngliche 10-kHz-Sinusschwingung. Obwohl die zweite Tonspur auf den ersten Blick hin nicht mehr wie eine richtige Sinusschwingung wirkt, stimmt die Periodendauer der

Schwingung mit der von Tonspur 1 gut überein. Auch die Amplitudenausschläge sind annähernd gleich groß. Ganz anders sieht dies jedoch bei Tonspur 3 aus, die als einzige nicht dem Nyquist-Shannon Theorem folgt. Die Abtastrate liegt hier unter der notwendigen doppelten Schwingungsfrequenz. Bei dem Ton in Tonspur 3 handelt es sich im Großen und Ganzen noch immer um eine Sinusschwingung. Ihre Frequenz stimmt aber nicht mehr mit dem Ursprungssignal überein. Die schlecht gewählte Abtastrate hatte eine Störfrequenz (Aliaskomponente) zur Folge.

Die Auswirkungen der verschiedenen Abtastraten können in einem weiteren Schritt auch akustisch gezeigt werden. Das *Control Panel* jeder Tonspur besitzt einen *Stumm*-Button. Mit diesem kann man jede beliebige Spur bei der Wiedergabe ausblenden. Am besten man deaktiviert zu Beginn alle Tonspuren und startet die Wiedergabe durch Drücken des *Play*-Buttons. Nun aktiviert man abwechselnd eine Tonspur nach der anderen, sodass stets nur ein Ton alleine zu hören ist. Obwohl alle drei Tonspuren grafisch nicht ident sind, klingen Tonspur 1 und 2 für unser Ohr gleich. Lediglich Tonspur 3 klingt auch akustisch anders. Dies liegt an der falschen Rückrechnung des digitalen in ein analoges Signal, die aus der zu gering gewählten Samplerate hervorgeht.

Neben dem selbst erzeugten, simplen Beispiel, lässt sich die Auswirkung der Abtastfrequenz auf die Qualität von Audiodaten auch noch anders demonstrieren. Dazu lässt man den Schülerinnen und Schülern ein beliebiges Musikstück (am besten in CD-Audio-Qualität) in ein neues *Audacity*-Projekt laden. Für den Beginn des Experiments sucht man eine geeignete Stelle im Lied, an der viele Hoch- und Tieftöne bzw. viele unterschiedliche Klänge zu hören sind und startet dort die Wiedergabe der Musikdatei.

In *Audacity* lässt sich vor jeder Wiedergabe die Projektrate ändern. Sie bestimmt jene Abtastfrequenz, die beim Abspielen einer Audiodatei tatsächlich angewandt wird und kann am entsprechenden Button im linken, unteren Teil der Benutzeroberfläche eingestellt werden. Vor jeder Wiedergabe verringert man nun die Projektrate und hört, ob sich die wiedergegebene Tondatei ändert. Da der menschliche Hörbereich bis zu einer Schallfrequenz von 20 kHz reicht, dürfte man nach dem Nyquist-Shannon Theorem bis zu einer Projektrate von mehr als 40 kHz keinen Unterschied beim Abspielen der Audiodatei feststellen können. Erst nach dem Unterschreiten dieser Grenze fallen höhere Töne nach und nach weg. Sobald die Projektrate die doppelte Höhe der größten Frequenz im Musikstück unterschreitet wird ein merkbarer Qualitätsverlust auftreten.

Da die Musikwahrnehmbarkeit des Menschen in einem höheren Frequenzbereich liegt als die Sprachwahrnehmbarkeit, wird zuerst die Musik falsch wiedergegeben werden, jedoch Gesprochenes länger richtig erscheinen.

Literatur und Referenzen

- [AU10] Audacity Projektwebseite, <http://audacity.sourceforge.net/>, 24.05.2010.
- [AV10] Avidemux Webseite, <http://www.avidemux.org/>, 24.05.2010.
- [BR00] Brice R., „Music Engineering. The Electronics of Playing and Recording“, Newnes, Paris, 2000, S. 7ff.
- [EI10] Eisenbarth M., „Audio und Video im Bildungswesen“, Magisterarbeit, Technische Universität Wien, Wien, 2010.
- [FR10] Free Audio Editor Produktwebseite, <http://www.free-audio-editor.com/>, 24.05.2010.
- [NC10a] NCH Software, WavePad Webseite, <http://www.nch.com.au/wavepad/>, 24.05.2010.
- [NC10b] NCH Software, VideoPad Webseite, <http://www.nchsoftware.com/vidiopad/>, 24.05.2010.
- [ST99] Steinmetz R., „Multimedia-Technologie. Grundlagen, Komponenten und Systeme“, 2. Auflage, Springer Verlag, Heidelberg, 1999, S. 24ff.
- [VI10] VirtualDub Projektwebseite, <http://www.virtualdub.org/>, 24.05.2010.