

Ausarbeitung

– Algorithmus für Musikverarbeitung –

von René Fischer Matrikel 223358 TU Berlin

Einleitung

Dieses Referat hat zum Ziel, einen Eindruck zu vermitteln, wie Musik aufgebaut ist, wie man sie elektronisch auswerten kann und worauf die Gruppe „Musikauswertung“ achten sollte.

Des Weiteren sind viele Erfahrungen mit dem Thema Musik – sowohl technischer als auch gestalterischer Natur – mit in diese Ausarbeitung eingeflossen, so dass als Ergebnis ein Vorschlag für einen Algorithmus zur Auswertung eines Musiksignals entstanden ist. (Siehe Dokument „Handout“)

Es empfiehlt sich, das Handout parallel zu dieser Ausarbeitung zu betrachten, da dort einige Zusammenhänge schneller grafisch erfasst werden konnten, als es mittels Text möglich wäre.

Aufgabenstellung und Lösungsidee

Die Aufgabe für die gesamte Gruppe besteht darin, eine analoge Schaltung zu entwerfen, die ein Audiosignal in drei Steuersignale für drei verschiedenfarbige Leuchtmittel umzusetzen.

Dabei besteht der Anspruch, dass sich mit ändernden Eigenschaften des Audiosignals auch die Eigenschaften der drei Steuersignale ändern.

Die Steuersignale müssen zudem auch noch ganz speziellen Ansprüchen genügen. Mit Ihnen werden jeweils ein rotes, ein grünes und ein blaues Leuchtmittel angesteuert und in Leuchtdauer und –intensität bestimmt. Die speziellen Ansprüche äußern sich dahingehend, dass der Benutzer – den das Ergebnis wesentlich mehr interessiert als die Technik dahinter – abhängig vom eingespeisten Audiosignal verschiedene Farbmuster, –abfolgen und –stimmungen erwartet.

Damit ist festgelegt, was die Schaltung an Informationen aus der Musik „entnehmen“ muss:

- die Stimmung der Musik
- den Rhythmus der Musik
- Abwechslungsreichtum

Diese drei Punkte sind schnell erklärt: Musik kann verschiedene Wirkungen auf den Menschen haben, aber die grundlegenden Eigenschaften sollte unsere Algorithmus erkennen und umsetzen – so sollte bei schneller beatlastiger (beat – engl. Schlag) Musik sich auch das Licht schnell ändern, aufgeregt bunt sein. Hingegen sollte bei eher ruhigen Passagen (z.B. klassische Musik) sich die Leuchtintensität verringern und nur noch sanfte Farbwechsel zu erkennen sein.

Ein Audiosignal – was ist das eigentlich?

Ein Audiosignal ist in unserem Fall entweder ein Musikstück oder aber Raumgeräusche/ Sprache, aufgenommen durch ein Monomikrofon. Nun, vielmehr ergeben die Informationen aus dem Audiosignal eben diese. Das Signal selbst besteht aus sehr vielen, überlagerten Wechselspannungen, mit verschiedenen Frequenzen im Bereich von 25Hz-20kHz, Amplituden und Dämpfungen.

Jedes Geräusch, das der Mensch wahrnimmt, wurde zuvor durch einen physikalischen Prozess erzeugt – der Schwingung. Dies kann auf die unterschiedlichsten Weisen passieren. Beim Zupfen einer Gitarrensaite beispielsweise (hierbei bestimmt die Zupfstärke die Amplitude, bzw. Lautstärke der Schwingung und die Saitenlänge die Frequenz). Diese Schwingung breitet sich nun im Raum aus und kann durch ein Mikrofon in eine elektrische Wechselspannung mit entsprechendem Amplitudengang und Frequenzgang gewandelt werden.

Es gibt sehr viele Methoden, Schall in elektrische Signale zu wandeln, jedoch würden detailliertere Informationen hierzu zu sehr das Thema verfehlen.

Je mehr Schallquellen bei der Aufnahme vorhanden waren, bzw. je mehr zusammengemischt werden, je mehr Schwingungen überlagern sich in eben diesem Audiosignal, wobei folgendes gilt:

Eine Schwingung einer bestimmten Frequenz kann nur einen bestimmten Amplitudengang besitzen.

Dass heißt, dass Schwingungen mit gleicher Frequenz sich in einem Audiosignal überlagern und entsprechend verstärken, bzw. abschwächen oder gar auslöschen – entsprechend der Phasenlage.

Für die Auswertung eines solchen Audiosignals ist es nun von entscheidender Bedeutung zu wissen, welcher elektronischer Wechselspannung, welche ursprüngliche Schallquelle zuzuordnen ist.

Zurück zu dem Gitarrensaiten-Beispiel:

Vorausgesetzt die Schwingung der gezupften Saite wurde mittels einem Mikrofon aufgenommen und liegt nun als Audiosignal vor. Wenn man nun im Nachhinein bestimmen will welche der sechs Saiten es war, braucht man sich nur zu vergegenwärtigen, dass jede dieser Saiten (aufgrund verschiedener Massen und Federkonstante) verschiedene Frequenzen beim Schwingen erzeugt.

Wenn man das aufgenommene Signal also mit Hilfe eines Oszilloskops betrachtet, kann man die Frequenz ablesen und so eindeutig bestimmen, welche Saite bei der Aufnahme angespielt wurde.

Dabei wird man jedoch auf eine weitere Schwierigkeit stossen. Oberschwingungen. Gerade bei durch Mikrofonen aufgezeichneten Schwingungen gibt es keine monofrequenten Signale. Dass heisst, die angezupfte A-Saite einer Gitarre wird als Audiosignal niemals nur aus einer 440Hz Schwingung bestehen, sondern auch noch überraschend viele andere Schwingungen der verschiedensten Frequenzen – im Idealfall vielfache von 440Hz, die sogenannten Oberschwingungen – enthalten. Die oben angesprochene „eindeutige Identifizierung“ der Saite ist jedoch trotzdem möglich, da sich das Intensitätsmaximum eben bei der Grundschwingung von 440Hz befindet.

Lösungsidee - Fortsetzung

Was charakterisiert ein Musikstück?

Aus der Erfahrung heraus sind es drei entscheidende Punkte, die es zu beachten gibt:

1. Der Rhythmus wird logischerweise durch Rhythmusinstrumente erzeugt. Diese „arbeiten“ zum einen im tieffrequenten Bereich – 25Hz-180Hz – und zum anderen in einem Bereich von ca. 1kHz-4kHz. Der tieffrequente Bereich ist das was man im allgemeinen als „beat“ oder auch Rhythmus versteht, typische einflussstärkste Instrumentierungen sind E-Bass, Bassdrum, große Trommeln, u.a.. In dem zweiten Bereich, sind vor allem Anschlaggeräusche, sogenannte „Clicks“ stark als Rhythmus wahrzunehmen, also beispielsweise beim Schlagzeug die Aufprallgeräusche der Sticks auf die Felle der Trommeln.

2. Die menschliche Stimme findet sich in dem Audiosignal auch an zwei sehr entscheidenden Stellen wieder:

Die Grundtöne lassen sich alle in einem Bereich von 240-ca.450 Hz ausfindig machen, wichtig für den Sprachdurchsatz, aber für die Verständlichkeit am wichtigsten ist der Bereich von 2,5-3,5kHz. In diesem Bereich ist auch das menschliche Ohr am empfindlichsten, so dass die Anpassung von Wiedergabe-(Stimmbänder) und Aufnahmeorgan (Ohren) aneinander bei Menschen zugunsten der Kommunikation angepasst ist.

3. Die Melodieinstrumente, die auch die Brillanz eines Musikstückes zum Grossteil mitbestimmen, finden sich in einem Frequenzbereich von 2-6KHz wieder.

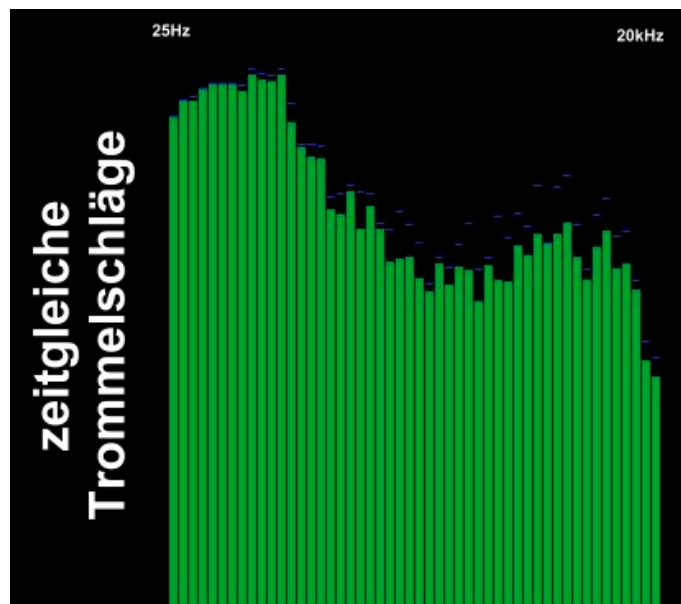
Natürlich sind diese Angaben stark verallgemeinert und würde man alle nicht genannten, trotzdem hörbaren Frequenzbereiche aus einem Musikstück entfernen würde sich dieses auch nicht mehr gut anhören, jedoch geht es hierbei um eine Auswertung.

Auswertung vieler verschiedener Musikstücke

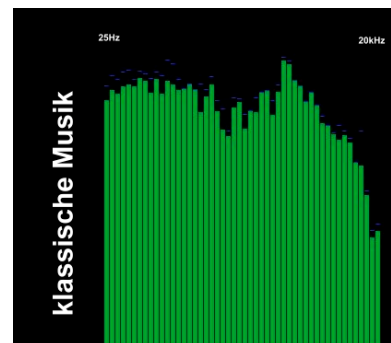
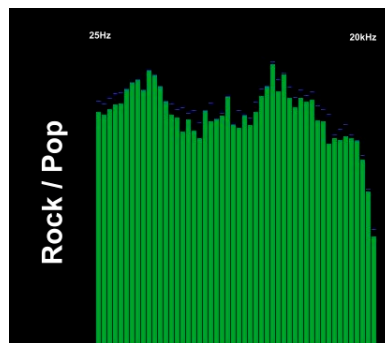
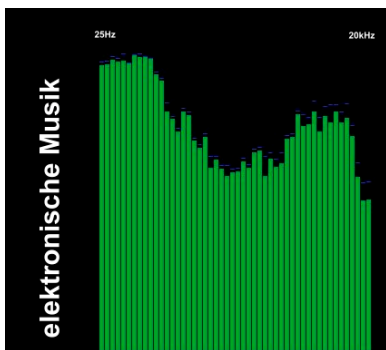
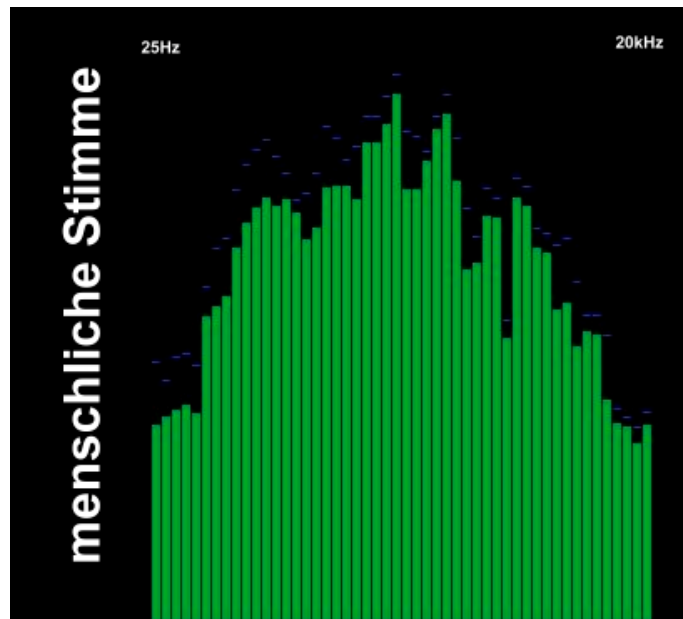
Eine Möglichkeit ein Audiosignal zu visualisieren, wurde bereits oben schon erwähnt. Man schaut sich das ganze mittels Oszilloskop an. Meistens kommt man damit jedoch nicht allzu weit, da man als Mensch nicht schnell genug ist, sämtliche Informationen zu verarbeiten, die einem dieses Gerät auf seinem Display aufzeigt.

Wenn man jedoch das Oszilloskop durch einen geeigneten Rechner ersetzt und mittels entsprechender Software die gesammelten Informationen auswertet und visualisiert kommt man schneller zu einem für uns geeigneten Ergebnis.

Ich habe viele hundert Musikstücke vom Rechner analysieren und grafisch darstellen lassen. Dabei habe ich das hörbare Spektrum in 50 Frequenzbänder eingeteilt und je Musikstück, bzw. Audiodatei eine Grafik erstellen lassen, die den Durchschnitt jedes Bandes über das gesamte Stück gemessen darstellt.



Auf der rechten Seite sieht man eine Momentanalyse zweier zeitgleich aufgenommener Trommelschläge. Dabei ist auch gut zu erkennen, dass obwohl es lediglich zwei Anschläge auf Trommeln sind, das so entstandene Audiosignal sehr komplex ist. Das zweite Beispielbild zeigt ein Audiosignal einer männlichen Stimme. Wäre es eine (typische) weibliche Stimme, würde das Maximum etwas weiter in den Bereich der höheren Frequenzen verschoben sein.



Die einzeln analysierten Musikstücke habe ich dann noch nach Stil sortiert und von allen Musikstücken eines Stils jeweils noch mal den Durchschnitt gebildet.

Wenn man nun das gesamte beobachtete Spektrum in die drei Bereiche Bass, Mitten, Höhen unterteilt lassen sich folgende Aussagen die verschiedenen Musikstile betreffend treffen:

- Elektronische Musik:** stark ausgeprägte Tief- und Hochfrequenzanteile, verhältnismäßig schwach vertretene Mitten, insgesamt starke Unterschiede in den Momentanmessungen durch Rhythmusbetonung
- Rock/ Pop:** Betonung im Mitteltonbereich, teilweise starke Rhythmusaussprägung in Tiefmitten und Bassbereich
- Klassische Musik:** tieffrequente Grundmuster, Reichtum an Hochmitten durch Melodieinstrumente

Zusammenfassung

Nach all diesen Betrachtungen merken wir, dass 3 Auswertungen für 3 Steuersignale nicht genügen, wenn man Rhythmus, Stil und Stimmung der Musik auf jede Farbe aufschlüsseln möchte ohne am Ende bei jedem Musikstück die gleiche Farbmischung zu erhalten.

Das eigentliche Audiosignal muss also in zwei identische Signale gesplittet werden.

Dann muss jeder dieser beiden Signalwege dreifach ausgewertet werden.

Signalweg 1 wird den Stil der Musik auswerten, indem das dort das gesamte Spektrum in drei Bereiche zerlegt wird und die entsprechenden Farben von Bass, Mitte und Höhe angesteuert werden.

Im Signalweg 2 werden die 3 Charakteristiken eines jeden Musikstücks ausgewertet.

Dabei werden der Rhythmus, die Grundtöne der Stimme sowie die Melodie auf jeweils eine der drei Farben gemischt.

Mit diesen theoretischen Überlegungen und der vorgeschlagenen Zusammenmischung der ausgewerteten Signale würden sich folgende Lichtmuster ergeben:

Vorgeschlagene Signalmischung (für Details siehe auch Handout):

Signalweg ROT: $1a + 2c$; Durchschnitt Bass & Momentanwerte Melodie

Signalweg GRÜN: $1b + 2a$; Durchschnitt Mitte & Momentanwert Beat

Signalweg BLAU: $1c + 2b$; Durchschnitt Höhen & Momentanwert Grundtöne

Elektronische Musik: Grundfarbe rot-blau; Charakteristik: rhythmusabhängige bunte peaks

Rock/ Pop: Grundfarbe grün-blau; Charakteristik: rote und blaue Impulse

Klassische Musik: Grundfarbe rot; vorwiegend bläuliche Farbwechsel