# Skript Medien 1Technische Grundlagen

Jedes Medium (Text, Bild, Ton, Film, Multimedia, Utensilien, etc.) zeichnet sich durch spezifische Eigenschaften und Eigenheiten aus. In diesem Skript werden ausgewählte technische Grundlagen zu einzelnen Medien zusammengefasst.

## Medium Bild

### Elektronische Darstellung: Pixel- oder Vektorgrafik

Eine Vektorgrafik setzt sich aus Linenelementen zusammen. Stellt man eine Vektorgrafik dar, so wird die Grafik frisch mit allen Linien in der gewünschten Grösse gezeichnet. Eine Vektorgrafik lässt sich daher in beliebiger Grösse darstellen (siehe Abb.1 links). Die Buchstaben der Computerschriften sind als Vektorgrafiken realisiert.



Abb.: Linke Kolonne: Vektorgrafik in verschiedenen Grössen, rechte Kolonne: Pixelgrafik in verschiedenen Grössen.

Eine Pixelgrafik besteht aus einzelnen Bildpunkten oder Pixeln (ähnlich einem Mosaik). Vergrössert man eine Pixelgrafik so wird jeder Pixel ebenfalls vergrössert. Bei einer starken Vergrösserung würde man erwarten, die einzelnen Pixel erkennen zu können. Jedoch versuchen die gängigen Office Programme durch interpolation (verschmieren) ein „besseres“ Ergebnis zu produzieren (siehe Abb.1 rechts).

### Pixelbilder: Grösse, Pixel, Auflösung

Wird eine Pixelgrafik stark vergrössert, kann das Resultat unbefriedigend sein (siehe Abb.1 rechts). Wie findet man heraus, wie viele Pixel eine Grafik haben muss, um „gut“ auszusehen?

Eine Grafik wird in der Regel gedruckt oder auf einem Bildschirm oder Beamer dargestellt. Die meisten Drucker, Bildschirme und Beamer bauen ihrerseits die Bilder, welche sie zeigen aus einzelnen Bildpunkten auf. Nur wenn die Anzahl Pixel der Grafik und des darstellenden Mediums im Einklang sind, wird ein „gutes“ Resultat erreicht.

Bei einer Grafik unterscheidet man

* **Bildgrösse**: Länge und Breite des Bildes in Zentimetern (cm) oder Inch (in), wobei 1 in = 2.54 cm.
* **Anzahl Pixel**: Anzahl Pixel in der Breite mal Anzahl Pixel in der Höhe.
* **Auflösung**: Anzahl Pixel pro Zentimeter oder Anzahl Pixel pro Inch (dpi).

Ein Drucker druckt in der Regel mit 300 dpi, moderne Bildschirme und Beamer schaffen 96 dpi.

**Beispiel**: Wird einer Grafik mit 960 Pixeln in der Breite die Bildbreite 25.4 cm gegeben so weist es eine Auflösung von 96 dpi auf. D.h. auf einem Bildschirm mit 96 dpi Auflösung wird das Bild exakt 25.4 cm breit dargestellt werden. Das gilt jedoch nicht für ein Beamer, welcher grundsätzlich jedes Bild vergrössert. Druckt man dasselbe Bild mit 300 dpi auf einem Drucker aus, so wird es mit einer Bildbreite von rund 8.1 cm ausgegeben.

### Pixelgrafik: Diverse Speicherformate

Warum braucht es verschiedene Speicherformate für Pixelgrafiken?

Der Speicherplatz den eine Pixelgrafik belegt, berechnet sich als

Speicherplatz in Byte = (Anzahl Bildpunkte • Farbauflösung in Bit) / 8 .

Mit 1 Bit Farbauflösung lassen sich Schwarzweissbilder darstellen, mit 8 Bit Bilder mit 256 Farben und mit 24 Bit sogenannte „TrueColor“ Bilder mit 16.7 Mio Farben. Ein TrueColor Bild einer Digitalkamera mit 10 Megapixel hätte somit eine Grösse von 28.6 MByte. Das ist zu gross für viele Anwendungen und zu speicherintensiv für die meisten Kameras. Eine Lösung des Problems besteht darin, die Bilder zu komprimieren. Jedoch gibt es zwei verschiedene Komprimierungsansätze. Beim verlustfreien Komprimieren bleibt die Pixelgrafik unverändert, der Speicherplatz kann dabei reduziert jedoch nicht beliebig klein gehalten werden. Beim nicht-verlustfreiem Komprimieren lässt sich der Speicherplatz im Prinzip beliebig klein machen, man nimmt dabei jedoch einen Qualitätsverlust des Bildes in Kauf.

Um alles Bedürfnissen Rechnung zu tragen, gibt es verschiedene Speicherformate, welche den unterschiedlichen Komprimierungsansätzen Rechnung tragen. Die verbreitetsten Speicherformate für Pixelgrafiken sind im Folgenden beschrieben.

**TIFF (Tagged Image File Format)**

Vorteile: Diese Speichermöglichkeit wird dann angewendet, wenn eine Bilddatei möglichst unverändert (d.h. verlustfrei) aufbewahrt werden soll und wenn auch ältere Programme die Datei öffnen können sollten. TIFF unterstützt 8, 16, 24 und 48 Bit Farbauflösung (für nicht RGB Farben) und erlaubt die Verwendung von transparenten Hintergründen. Es ist im Druckwesen sehr verbreitet.

Nachteile: Das TIFF Format braucht in der Regel viel Speicherplatz, wird nicht von den gängigen Browsern angezeigt (d.h. es eignet sich nicht für das Internet) und erlaubt keine Animationen.

**GIF (Graphics Interchange Format)**

Vorteile: Gängiges Format, um Grafiken im Internet abzulegen. Im Unterschied zu JPEG wird verlustfrei komprimiert. GIF unterstützt animierte Bilder und erlaubt, einen transparenten Bildhintergrund zu verwenden.

Nachteile: GIF kennt nur die Farbtiefen 1, 2, 4 und 8 Bit, d.h. es können nicht mehr als 256 (beliebige) Farben dargestellt werden. Das macht GIF für Fotos ungeeignet. Der im GIF verwendete Kompressionsalgorithmus aus dem Jahre 1989 ist relativ schlecht.

**JPG oder JPEG (Joint Photographic Experts Group)**

Vorteile: JPG komprimiert nicht-verlustfrei und braucht wenig Speicherplatz. Die Resultate sind jedoch erstaunlich gut, da der kompressionsbedingte Qualitätsverlust dem menschlichen Auge angepasst ist. Das Speicherformat eignet sich daher gut für das Internet. JPG hat immer eine Farbtiefe von 24 Bit. Für Bilder mit 24 Bit Farbtiefe sind JPG Ergebnisse in akzeptabler Qualität mit weniger Speicherplatz realisierbar als mit GIF, PNG oder TIFF. Die Qualität kann zwischen 0% und 100% eingestellt werden.

Nachteile: JPG reduziert die Qualität der Bilder (auch die Qualität 100% führt zu einem Qualitätsverlust) und eignet sich daher nicht für Bilder, an denen irgendwann einmal etwas nachträglich geändert werden soll (mehrmaliges Abspeichern im JPG Format verstärkt den Verlust). Bei Bildern, welche geringere Farbtiefen als 24 Bit aufweisen (z.B. bei schwarzweiss Bildern oder Grafiken) brauchen PNG und GIF oft weniger Speicherplatz als JPG. JPG erlaubt keine Transparenz und keine Animation.

**PNG (Portable Network Graphics)**

Vorteile: PNG ist das modernste der aufgelisteten Verfahren, es eignet sich für Internetbilder sowie für Fotos. Es ist wie TIFF ein verlustfreies Verfahren, komprimiert in der Regel besser als GIF und unterstützt 1, 2, 4, 8, 16, 24 und 48 Bit Farbtiefe (für nicht RGB Farben). Auch transparente Hintergründe werden unterstützt.

Nachteile: PNG braucht für TrueColor Bilder in der Regel mehr Speicherplatz als JPEG, da es verlustfrei Komprimiert. PNG erlaubt im Gegensatz zu GIF keine Animation. Es werden weniger Farbformate unterstützt als bei TIFF (z.B. CMYK wird nicht unterstützt).

### Arbeiten mit Ebenen

Wir möchten in der Grafik auf der linken Seite in Abb.2 das Wort „Test“ mitsamt dem schwarzen Hintergrund verschieben. In einer herkömmlichen Pixelgrafik ist das nur mit grossem Aufwand möglich. Der Trick besteht nun darin, das Bild aus mehreren Ebenen aufzubauen, welche wie einzelne Folien übereinander gelegt werden und bei der Durchsicht das gewünschte Bild ergeben (siehe Abb.2). Um das fertige Bild zu erhalten werden alle Ebenen miteinander vereint. Solange die Ebenen einzeln im Speicher vorhanden sind, kann jede einzelne Ebene leicht verändert oder die Ebenen gegeneinander leicht verschoben werden.



Abb.: Arbeiten mit Ebenen

Grafiken mitsamt allen Ebenen lassen sich in Gimp mit dem Gimp-eigenen Format XCF (eXperimental Computing Facility) abspeichern. Das Format braucht relativ viel Platz, da es eine Grafik mit allen Gimp-spezifischen Einstellungen wie Ebenen, Masken, Filtern, etc. abspeichert.

Der Gimp Kurs <http://www.lernnetz-sh.de/kmLinux/doc/Gimp-Einfuehrung/saug-gimp/index.html> führt Ebenen im Kapitel „Ebenen“ ein und zeigt im Kapitel „Die erste Animation“, wie man mit Hilfe von Ebenen eine einfache Animation erstellen kann.

## Medium Ton

### Elektronische Darstellung: Digitalisieren

Ton breitet sich in Form von Schallwellen im Raum aus. Bei der Digitalisierung werden die Wellen mit Hilfe eines Zeitrasters und eines Lautstärkerasters in einzelne Abtastwerte zerlegt (siehe Abb.3). Diese Abtastwerte werden in einem geeigneten elektronischen Format abgespeichert.



Abb.3: Digitalisierung von Ton; horizontale Achse: Zeit; vertikale Achse: Lautstärke, je stärker die Abweichung von der Mittellinie (Null-Linie), desto lauter das Signal. Jedes Kästchen stellt einen Abtastwert dar.

Die Feinheit des Zeitrasters wird als Samplerate bezeichnet und bestimmt zu wie vielen Zeitpunkten pro Sekunde die aktuelle Lautstärke gemessen wird. Die Samplerate bestimmt den Tonhöhenumfang, welcher aufgezeichnet werden kann. Eine hohe Samplerate bedeutet einen grossen Tonhöhenumfang. Musik CDs werden mit einer Samplerate von 44‘100 Zeitpunkten pro Sekunde aufgezeichnet, was ungefähr dem Tonumfang entspricht, die ein junger Menschen wahrnehmen kann.

Die Feinheit des Lautstärkerasters wird als digitale Auflösung bezeichnet und typischerweise in Bit angegeben. Eine hohe digitale Auflösung bedeutet eine grosse Dynamik. Musik CDs werden mit einer digitalen Auflösung von 16 Bit aufgezeichnet, d.h. es können 216= 65536 verschiedene Lautstärkestufen wiedergegeben werden.

### Digitaler Ton: Diverse Speicherformate

Wie bei den Bildern können digitale Töne je nach Anwendung verlustfrei oder nicht-verlustfrei komprimiert werden. Die gängigsten Speicherformate für digitale Musik sind im Folgenden näher erklärt.

**WAV (Waveform Audio File)**

Vorteile: In der Regel nicht komprimiertes Speicherformat. Liest man eine Musik CD in einen Computer ein, so werden die Musikdaten üblicherweise als WAV Speicherformat dargestellt (Musik CDs verwenden das Speicherformat „Red Book“). Eignet sich für Musik in hoher Tonqualität und Aufnahmen, welche weiter bearbeitet werden.

Nachteile: Das Speicherformat braucht viel Speicherplatz.

**MP3 (Moving Picture Experts Group audio layer 3)**

Vorteile: Das Patent geschützte Format beinhaltet eine nicht-verlustfreie Kompression der Musikdaten mit dem Ziel, nur vom Menschen bewusst hörbare Audiosignale zu speichern. Die Bitrate gibt an, wie viel Speicherplatz pro Sekunde zur Verfügung gestellt wird, um die digitalen Töne abzuspeichern. Eine hohe Bitrate bedeutet eine hohe Qualität. MP3 eignet sich dafür, digitale Musik in einer ansprechenden Tonqualität mit wenig Speicherplatz abzuspeichern. Im Internet ist MP3 mit einer Bitrate von 128 kbit/s gebräuchlich, welche rund einen elftel des Platzes einer CD Aufnahme braucht und dennoch eine für viele akzeptable Qualität liefert.

Nachteile: Bei einer sehr grossen Kompression (kleine Bitrate) tritt ein hörbarer Verlust der Tonqualität ein, speziell bei kurzen Tönen (z.B. Kastagnetten).

**OGG (OGG Vorbis)**

Vorteile: Der Begriff „Ogg“ ist einem Computerspiel und der Begriff „Vorbis“ einem Roman entliehen. Das patentfreie Format OGG Vorbis gilt bezüglich Tonqualität speziell bei tiefen und mittleren Bitraten als überlegen gegenüber dem MP3 Format.

Nachteile: Nicht alle Geräte unterstützen das OGG Vorbis Format.

### Tonbearbeitung: Stimmverständlichkeit erhöhen

Tonaufnahmen menschlicher Stimmen werden in der Regel nach der Aufnahme tontechnisch bearbeitet, um die Stimmverständlichkeit zu erhöhen. Dies gilt speziell dann, wenn die Stimme mit Musik unterlegt wird. Nicht unüblich sind folgende Bearbeitungsschritte, welche genau in der folgenden Reihenfolge durchgeführt werden müssen, um ein optimales Resultat zu erreichen. Die einzelnen Schritte können z.B. mit der Audiosoftware Audacity umgesetzt werden.

1. **High Pass Filter** (oder auch low-cut Filter)

Tonhöhen (Frequenzen) welche tiefer als ein zu bestimmender Frequenzgrenzwert sind werden abgeschnitten, alle höheren Frequenzen werden durchgelassen. Stellt man den Grenzwert auf 80 Hz ein, wird die menschliche Stimme durchgelassen und störende Mikrofongeräusche im tieffrequenten Bereich (z.B. Brummen) rausgefiltert.

Tipp: Ist in jedem Fall zu empfehlen.

1. **Equalizer**

Ein Equalizer erlaubt es verschiedene Tonhöhen (Frequenzen) anzuheben oder abzuschwächen. Ein Anheben bestimmter Frequenzbereiche kann die menschliche Stimme verändern. In der Regel werden einzelne Frequenzen nur um 2-4 dB angehoben (dB ist ein Lautstärkemass), damit die Natürlichkeit der Stimme erhalten bleibt:

* 200 Hz – 300 Hz (Männer), 500 Hz – 800 Hz (Frauen): Mehr Bauch, d.h. die Stimme tönt voller aber ist eventuell auch schlechter zu verstehen.
* 6 kHz – 8 kHz: Betont die Höhen und sorgt für bessere Textverständlichkeit.
* 11 kHz: Bringt die Stimme aber auch die Atemgeräusche mehr in den Vordergrund.

Es kann sein, das eine Stimme für sich voll klingt, aber im Zusammenspiel mit Musik eher dünn erscheint. Das kann man verbessern, indem man die Lautstärken der Frequenzen zwischen 200 Hz – 300 Hz in der Musikaufnahme um ein paar dB absenken, damit mehr Platz für die Stimme bleibt.

Tipp: Ist nicht in jedem Fall erforderlich.

1. **Kompressor**

Den Lautstärkenunterschied von leisen zu lauten Passagen in einem Tonsignal bezeichnet man als Dynamik. Speziell die menschliche Stimme weisst eine grosse Dynamik auf. Das kann dazu führen, dass einzelne Stimmlaute oder -passagen untergehen. Ein Kompressor reduziert die Dynamik des Tonsignals und führt dazu, dass eine Stimme klarer herausgehört werden kann und stärker in den Vordergrund rutscht.

Tipp: Die Default Einstellungen des Audacity Kompressors liefern meist vernünftige Resultate.

1. **Pitch**

Es ist möglich die Tonhöhe (Pitch) eines Tonsignals nachträglich zu ändern – bei genügend Geduld kann man den Gesangsvortrag eines unsicheren Sängers Ton für Ton korrigieren. Idealerweise wird bei der Veränderung der Tonhöhe das Originaltempo nicht zerstört.

Tipp: Ist nur selten sinnvoll. In Audacity kann z.B. eine Anpassung der Tonhöhe einer längeren Gesangspassage zu leichten Veränderungen im Tempo führen.

1. **Delay**

Beim Delay werden mehrere Kopien eines Tonsignals leicht versetzt und üblicherweise mit immer kleinerer Lautstärke zum Originalsignal hinzu gemischt. So kann ein Halleffekt wie in einer grossen Halle erzeugt werden, welcher die Stimme voluminöser und wuchtiger erscheinen lassen kann. Je nach Geräusch kann man ab einer Delay time (Zeit zwischen den Echos) von mehr als 20 ms (Millisekunden) die Echos einzeln wahrnehmen, was in der Regel unerwünscht ist. Die menschliche Sprache verträgt auch längere Delay times, das hängt jedoch auch ab vom Decay amount (um soviele dB wird jedes folgende Echo leiser).

Tipp: Ist zu empfehlen. Für eine menschliche Stimme kann ein Delay mit einem Decay amount von 12 dB, einer Delay time von 0.07 Sekunden und 5 Echos eventuell zum gewünschten Resultat führen.