

H. Völz
Gedanken zu einem hochdichten Grafik-Code
Einleitung

Bilder werden heute - von seltenen Ausnahmen abgesehen - pixelweise als reine Zeitsignale längs der Bildzeile und dann in Abfolge von Zeilen übertragen. Dabei ist schon lange bekannt, daß dieses Verfahren in vielfacher Hinsicht höchst redundant ist. Das gilt zumindest bezüglich der Ähnlichkeit zwischen den Zeilen und der Fastgleichheit aufeinander folgender Bilder.

In der Rechentechnik gibt es dagegen ein konsequentes Bemühen, Ähnliches bis Gleiches durch wiederholten Aufruf der eventuell leicht modifizierten Routine nur einmal zu speichern. Genau in diesem Sinne ist das Prinzip der Rekursion in dem letzten Jahrzehnt starkem Brauch geworden. Außerdem existiert aber mindestens seit Turing die Erkenntnis, dass alles Berechenbare mittels rekursiver Funktionen ausdrückbar ist.

Dieser Unterschied kann noch durch einen andersartigen, praktischen Widerspruch vertieft werden. Woran kann es wohl liegen, dass man sich in der Regel einen Film nur wenige Male anschaut, ein Musikstück dagegen sehr oft hört? Ist doch der Aufwand bei der Produktion und bei der Speicherung genau entgegengesetzt, nämlich beim Film extrem viel höher. Weiter kommt noch hinzu dass unser Gehirn Schallsignale über sieben und visuelle Signale nur über sechs Neuronen-Schichten leitet. Schallsignale sind also wesentlich komplexer. Folglich stimmt an der z.Z. geübten Technik des Fernsehens und allgemeiner der Grafik-Realisationen auf dem Rechner etwas nicht. Diesen Problemen sei hier aus der Sichtweise neuartiger, allgemeiner Codebetrachtungen nachgegangen.

Die Textcodierung der Sprache

Beim Rundfunk gilt die Regel: Eine Schreibmaschinenseite zu verlesen, dauert etwa zwei Minuten. Für nur mittlere Qualitäts-Ansprüche (Bandbreite 10 kHz und Störabstand 40 dB) folgen aus der Informationstheorie etwa 65 kBit/s oder 8 kByte/s und somit rund ein MByte Speicherkapazität.

Erzeugt wird nun aber heute der Text meist auf dem Computer mit einem Textverarbeitungssystem. Dabei ist er in ASCII-Zeichen codiert und entsprechend Normseite mit 30 Zeilen zu 60 Zeichen werden nur rund 2 KByte benötigt.

Das Verhältnis beider Speicheranforderungen beträgt also fast 1:1000. Natürlich enthalten beide Verfahren nur inhaltlich gleiche Information. Der Unterschied kann über die Begriffe relevant und redundant erklärt werden. In der ASCII-Speicherung fehlen z.B. alle Spezifika des Sprechers, sowohl sein individueller Sprachklang als auch seine bewusst vorgenommene Betonung, die Pausen usw. Solche Besonderheiten werden bei der ASCII-Speicherung als irrelevant betrachtet und gerade hier liegt der Gewinn an Speicherkapazität.

Ähnliches gilt auch beim DTP (desktop publishing) und beim Zeichenerkennen (OCR; optical character recognition). Der Text besteht hiernach aus zwei Komponenten: Den ASCII-Code und dem Schriftfont. Der Schriftfont ist die spezifische (individuelle) Auslegung für den ASCII-Text. Beim DTP wird er hinzugefügt, beim OCR muss er entfernt werden. Das erste ist technisch relativ einfach, das zweite höchst kompliziert. Für den ASCII-Text ist der Schriftfont irrelevant.

Während Relevanz stets stark kontextbezogen ist, gilt dies für Redundanz weitaus weniger. Sie wird meist rein technisch und nahezu perfekt beherrscht. Optimale redundanzarme Codierung erfolgt zum Beispiel mit dem Huffman-Code. Betonte technische Redundanz wird bei fehlererkennenden und -korrigierenden Verfahren verwendet. Sie ist daher im Kontext dieses Beitrages weniger wichtig. Hierzu noch einige Beispiele:

Relevant bedeutet also, dass etwas entsprechend dem interessierenden Umständen wichtig, wesentlich, erheblich ist. Folglich ist Relevanz nicht exakt, eindeutig festzulegen. Vielmehr gehen hier das subjektiven Ermessen und der aktuelle "Zustand" des Empfängers oder Nutzers der Information entscheidend ein. So kann "gut" auf sehr unterschiedliche Weise ausgesprochen werden, ja sogar so, dass es das Gegenteil bedeutet. Es gibt aber auch objektivere Komponenten. Sie sind dann aber eher redundant. Für unsere Schallwahrnehmung ist es unwesentlich, Frequenzen über 20 kHz zu übertragen. Sie sind für uns unhörbar. Für eine Fledermaus sind sie dagegen wesentlich. Die Selbstlaute der deutschen Sprache sind eigentlich redundant. Sie können bei Fehlen problemlos erraten werden und schließlich macht hiervon die Stenografie reichlich Gebrauch.

Zwei neuartige Codes

Es sollen jetzt zwei Fakten unterschieden werden: Ein Objekt und seine rechnerinterne (technikrelevante) Modellierung. Zwischen beiden muss eine Vermittlung (Abbildung) erfolgen. Diese Vermittlung erfolgt je nach Richtung durch Objekt- und Struktur-Code. Sie seien im Weiteren definiert:

Ein Objekt-Code bildet das Objekt auf eine Zahl ab. Es wird so durch die Zahl bzw. ein aus ihr abgeleitetes technisches Signal ersetzt. So steht z.B. die ISBN-Nummer für ein Buch. So kann männlich durch 0 und weiblich durch 1 codiert werden. Der o. g. ASCII-Code ist ein international vereinbarter Objekt-Code für Buchstaben, Ziffern und Zeichen. Das Verfahren der Objekt-Codierung ist also jenem Prinzip ähnlich, wie wir für einen Gegenstand, ein Tier, einen Menschen usw. einen Namen zu seiner Benennung verwenden. Mit dem Objekt-Code werden oft auch Klassen von Objekten bezeichnet. Dann ist viel von der Individualität des einzelnen Objektes irrelevant. Beim Buchstaben "A" interessiert nicht sein Schriftfont. Beim Code für "Hund" interessiert dann weder seine Rasse noch Farbe.

Ein Struktur-Code ist in schlicht die Umkehrung des Objekt-Codes. Ausgehend vom Zahlenwert des Codes wird hier meist mit technischen Mitteln ein Objekt erzeugt. Aus dem ASCII-Code "65" wird ein "A" mit dem gerade ausgewählten Schriftfont gedruckt.

Der genetische Code ist aus dieser Sicht ein Struktur-Code für das jeweilige Lebewesen. Er besitzt jedoch noch zusätzliche Eigenschaften. Er "konstruiert" nämlich nicht nur das Lebewesen sondern hält es darüber hinaus auch ständig lebensfähig.

Für Struktur-Codes existieren in diesem Falle folglich funktionelle und generierende, strukturierende Komponenten.

Die Bedeutung der Zeichenmanipulation

Als die Menschheit die Sprache erfand, diente sie nur der direkten Kommunikation. Die mit ihr getroffenen Aussagen entsprachen direkt der zugeordneten Handlung. Mit der Entwicklung der Schrift entstand eine neue Qualität. Die schließlich so eingeführten schriftlichen Buchstaben-Zeichen ermöglichen es, mit den Buchstaben, Worten und Texten frei zu manipulieren. So konnte jemand in aller Ruhe experimentell eine große Vielfalt von Textvarianten erproben und für eine spätere "Rede" festhalten. Mit wachsender Erfahrung ließen sich dabei immer mehr komplizierte, raffinierter Strukturen erzeugen. Es wurde möglich, sowohl analytisch als auch konstruktiv in völlig neuer Art zu "denken". Selbst hierarchische Strukturen waren relativ einfach zu bewältigen. Als Beispiel sei auf dem kompliziert aufgebauten, sehr langen Anfangssatz in "Michael Kohlhaas" verwiesen.

Zeichenmanipulation ist also eine sehr kreative Methode. Sie setzt aber Objektcodes voraus.

Objekt-Code für die Musik

Musik ist das zweite technische Gebiet mit einem international genormten und allgemein verwendeten Objektcode, nämlich 1982 wurde MIDI (musical instrument digital interface) von mehreren Gremien beschlossen. Schon weit vorher war die europäische Notenschrift der Buchstabenschrift für Sprache recht ähnlich. Auch sie gestatte es, musikalische Strukturen frei und bequem zu manipulieren. Ohne sie hätte wohl kaum die polyphone Musik entstehen können. Mit dem MIDI-Code entstand dann aber auch wie mit dem ASCII-Code die Möglichkeit viele Probleme auf dem Rechner zu übertragen und die Speicherung hoch dicht zu gestalten. Auch hierbei existieren wieder zwei Komponenten:

1. Tonhöhe, Tondauer und Lautstärke sind über genormte Zahlen ähnlich dem ASCII-Code definiert und
2. An die Stelle der Schriftfonts treten Klangbänke.

Nach diesem Prinzip kann ähnlich wie bei DTP ein im MIDI-Code gespeichertes "Klanggebilde" mittels unterschiedlicher Klangbänke äußerst verschiedenartig, z.B. als Violine, Akkordeon, Sopranstimme, Chor usw. erklingen. Auch die hohe Verdichtung der von Musik erfolgt ähnlich wie beim Text mit ASCII-Code. Für ein vielstimmiges Musikstück von etwa einer Stunde Dauer genügen wenige KByte, während die direkte Schallaufzeichnung der Klänge viele MByte erfordert.

Bei der Musik ist darüber hinaus, noch mehr als bei Texten eine hierarchische Struktur zu erkennen. Beim ersten Hören wird meist nur die unmittelbare und einfache Oberfläche gehört. Sie führt schrittweise zum Erkennen und Wiedererkennen der Rhythmen, Melodien, Themen und Motive. Erst danach, also beim erneuten Hören (oder bei einem erfahrenen Hörer) wird ein gut konstruiertes Werk detaillierter rezipiert.

Gibt es einen Objekt-Code für Bilder und Film?

In der Rechentechnik werden heute hauptsächlich zwei Grafik-Typen unterschieden: Vektor-Grafik und Pixel-Grafik:

Bei der Pixelgrafik werden die einzelnen Bildpunkte auf dem Bildschirm unmittelbar oder leicht verdichtet gespeichert. Auf alle Fälle wird aber immer wieder zum direkten Pixelbild zurückgekehrt. Mit den üblichen Codierungsmethoden (insbesondere Run-length-Code) sind nur Verdichtungen auf maximal etwa ein Drittel möglich. Da hierbei keine Manipulation auf spezielle Bild-Objekte möglich ist, kann hier auch von keinem Objekt-Code gesprochen werden.

Die Vektorgrafik genannt) lässt sich u.a. an einer Linie bzw. bei einem Kreis erklären. Es werden dann nicht mehr die zugehörigen Pixel, sondern nur der Anfangs- und Endpunkt, bzw. der Mittelpunkt und Radius gespeichert. Dies bedeutet eine erheblich größere Redundanzminderung. Auch sind mit solchen Prinzipien Manipulationen möglich. Genau aus dieser Sicht läge eine Objekt-Codierung vor. Deshalb wird hierbei zuweilen auch von Objekt-Grafik gesprochen. Es gibt hierfür sogar eine weltweite Normung GKS = Grafisches Kern-System. Es ist gut geeignet für die technisch-wissenschaftliche Konstruktion, versagt jedoch bei den "üblichen" Bildern, wie etwa Landschaften und Portraits. Hier sind derartige, durch die Geometrie bestimmte Elemente unwesentlich. Erst in allerletzter Zeit zeichnen sich neuartige Ansätze ab. Sie wurden vor allem von Barnsley stark vorangetrieben und machen in Übereinstimmung mit den Ausführungen zu Text und Klang stark von der Rekursivität und zwar von einer Umkehrung der bekannten fraktalen Methoden Gebrauch.

In diesem Sinne - wenn wahrscheinlich auch aus anderem Gesichtspunkten heraus - ist es nun Barnsley offensichtlich gelungen, den fraktalen Prozess invers zu gestalten. Aus gegebenen Bildern zieht er die selbstähnlichen Anteile heraus und kann sie so zu wenigen Parametern von Gleichungen verdichten. So werden komplexe Bilder auf sehr wenige Daten komprimiert. Aus diesen Daten kann dann jederzeit das Ursprungsbild wieder nach der fraktalen Rekursion erzeugt werden. An dieser Methode ist nicht nur die hohe Verdichtungsrate sondern auch die Manipulierbarkeit der Objekte über Änderung der Parameter von großer Wichtigkeit. Es könnte

also durchaus dieses Prinzip zu einem Grafik-Code führen, der ähnliche Universalität und Leistungsfähigkeit in sich trägt, wie der ASCII- und MIDI-Code. Damit rückt eine Frage in den Mittelpunkt der Betrachtungen:

Warum ist dieses Prinzip so kompliziert, während wir andererseits so (scheinbar) einfach und effektiv optisch rezipieren können?

Ich meine, es liegt einfach daran, dass wir nicht gewohnt sind rekursiv zu denken. Das Prinzip der Rekursivität ist eigentlich nur in ausgewählten den Fachkreisen von Informatikern bekannt. Nur wenige von diesen können mit der Rekursivität gut umgehen. Noch weniger verstehen dann, was invers rekursiv beinhaltet. Dies schließt aber nicht aus, dass unser Gehirn (Sehzentrum) solche Prozesse hocheffektiv realisiert. Dies würde bedeuten, dass wir uns intensiv mit dieser Problematik auseinandersetzen müssen, um künftig effizient mit Bildinformation umzugehen. Hieraus folgt dann aber auch, dass schon als klassisch geltende Aussagen ihre Gültigkeit verlieren dürften. Eine solche Aussage wäre, dass wir wahrscheinlich nicht 90 % unserer Information optisch aufnehmen, sondern weitaus weniger, vielleicht nur 20 oder 30%. Alle Abschätzungen hierzu erfolgten ja nur über die Pixelstruktur, nämlich die Anzahl unser Stäbchen und Zäpfchen im Auge. Weiter wäre zu folgern, dass Bilder mit einer weitaus geringeren Komplexität als man bisher annahm, zu beschreiben sind. So wird zumindest verständlicher was Therapeuten festgestellt haben: ein Blinder lernt wesentlich leichter und besser abstrakte Begriffe (wie z.B. Gott) als ein Taubstummer. Vielleicht gehört hierzu auch, dass der Mensch nur aktiv Klänge (Sprache, Gesang, Geräusche) aber keine Bilder ausgeben (d.h. auch manipulieren) kann. Andererseits scheint für uns das Bilddenken äußerst wichtig zu sein. Hierauf ist ja die rechte Hirnhälfte, die nicht der Sprache fähig ist, weitgehend spezialisiert. Schließlich spielt das Bildgeschehen in den Träumen eine wichtige Rolle. Schließlich muss noch erwähnt werden, dass mit der Lösung dieser Problematik eventuell noch neue Fragen bezüglich des Zeitablaufs von Bildern, also für Film und Fernsehen auftreten können.

Mögliche Folgen für die Animation

In der Animationstechnik steht heute (immer noch) der Begriff fotorealistisch im Vordergrund. Ganz anders ist es in der Literatur und Musik. Ihre künstlerischen Leistungen liegen gerade darin, dass sie uns ungewöhnliche Welten aufzeigen. Es ist geradezu eine Aufgabe der Literatur, uns mögliche bis unwahrscheinliche, manchmal auch unmögliche Welten zu zeigen. Hierdurch können wir individuelle Vergleiche mit der von uns erlebten Welt anstellen oder Erkenntnisse gewinnen, die kaum auf einem anderen Weg zu erreichen sind. Pragmatisch betont, zeigt uns die Literatur auf ungewöhnliche Erlebnismöglichkeiten auf. So verweisen auch Gleichnis, Parabel, Fabel und Märchen oft auf etwas Anderes als ihre direkte Aussage. Besonders oft ist dies auch in der Zukunftsliteratur (u.a. St. Lem: Gast im Weltraum und Sterntagebücher) genutzt. Aber auch die "skurrile" Literatur zeigt uns neue traumhafte Welten (u.a. Kafka, Capek). Besonders wirksam ist dies z.B. in der Erzählung von John Collier "Grüne Gedanken" gezeigt. Hierbei nehmen die Blüten einer fleischfressenden Orchidee immer die Form des Kopfes des gefressenen Lebewesens an.

In der Musik ist der Bezug zu "natürlichen" Klängen sogar die Ausnahme. Programm-Musik ist sogar wegen dieses Anspruch oft unbeliebt. Musik schafft also fast immer "unnatürliche" Welten.

Aus diesen Analogien folgt unmittelbar, was ein effizienter Grafik-Code für die Bildgestaltung in Film, Fernsehen und Animation leisten könnte, ja müsste. Im Gegensatz zur bisherigen Technik müssten sich relativ leicht und effektiv "neuartige", verfremdete, künstlerische Bilder und Zeitabläufe gestalten lassen. Hierbei muss, damit die Rezipierbarkeit gewährleistet ist, ein geschichtlicher Prozeß ablaufen. Die Verfremdung und Neugestaltung der Bilder muss schrittweise erfolgen, wobei zumindest zunächst immer noch Bezüge zur Wirklichkeit gegeben sein müssen. Schrittweise werden dann immer mehr verfremdete und höher komplexe Bild und Bildfolgen als Kulturentwicklung möglich. Dies ist aber etwas völlig anderes als die Bildverarbeitung, mit z.B. Falschfarben, geometrischen Verzerrungen, Äquidensiten, Zeitlupe, Zeitraffer, Zeitsprünge, Zeitumkehrungen oder hochdichte Schnitte. Vielleicht liegt gerade deshalb der Reiz der Bilder von M. C. Escher in den unmöglichen Raumstrukturen und in der Selbstähnlichkeit mit der bis ins unendlich kleine führenden Wiederholung.