Richtig im Bilde - Grafikformate im Internet

Oliver Vornberger

Fachbereich Mathematik/Informatik

Universität Osnabrück

Mit der Verbesserung von Kommunikationsbandbreiten im Internet geht eine zunehmende Verwendung von multimedialen Bestandteilen einher. Dieser Artikel stellt die wichtigsten Grafikformate vor, die für 2D-Bilder auf WWW-Seiten geeignet sind.

Da sich nicht alle Effekte von Web-Browsern überzeugend im Schwarzweißdruck darstellen lassen, ist eine Online-Version unter der Adresse `http://www2.informatik.uni-osnabrueck.de/web_pic/paper/index.html` veröffentlicht.

Einleitung

Als Anfang der 90-er Jahre das Internet mit dem *Hypertext Transfer Protocol (HTTP)* um einen neuartigen Dienst erweitert wurde, änderten sich auch die typischen Informationsinhalte der übertragenen Nachrichten: neben rein textueller Information, die bisher bei den Protokollen *email*, *ftp*, *telnet* und *gopher* dominiert hatte, wurden nun auch 2-dimensionale Grafiken fester Bestandteil der im *World Wide Web (WWW)* angebotenen Daten. Eingebettet in Seiten formuliert mit der *Hypertext Markup Language (HTML)*, vervollständigen diese Bilder auf anschauliche Weise den sie begleitenden Text.

2D-Grafiken werden entweder direkt im Rechner erstellt oder nach dem Einscannen einer Fotovorlage mit einem digitalen Bildbearbeitungsprogramm manipuliert. Zu diesem Zwecke gibt es eine Unmenge von Public Domain Tools und auch kommerziell verfügbare Programme. Für die Windows-Welt seien stellvertretend *PaintShop Pro* und *Adobe PhotoShop* genannt; für die UNIX-Welt gibt es die netpbm Library mit Utilities wie `ppmquant`, `cjpeg` und `djpeg`.

Für den Gebrauch im Internet ist es selbstverständlich, daß die Grafikdaten komprimiert werden, denn Übertragungsbandbreiten müssen sparsam genutzt werden. Je nach Vorlage des Ausgangsmaterials und je nach intendierter Anwendung haben sich hierfür folgende Formate zur Verwendung auf WWW-Servern etabliert, erkennbar an ihrer Dateiendung:

- GIF (Graphics Interchange Format)
- JPG (Joint Photographers Expert Group)
- PNG (Portable Network Graphic)
- FIF (Fractal Image Format)
- FPX (FlashPix)
- DJVU (Deja Vu)
- SWF (Shockwave Flash)

Nur zwei dieser Formate, nämlich GIF und JPG werden von allen gängigen Web-Browsern unmittelbar angezeigt. Das PNG-Format wurde erst im Netscape Communicator 4.5 und Microsoft Internetexplorer 4.7 eingeführt; die Formate FIF, FPX, DJVU und SWF erfordern zum Betrachten zur Zeit noch ein Plugin.

In den folgenden Kapiteln werden Stärken und Schwächen aller sieben Formate erläutert und ihre typischen Einsatzgebiete vorgestellt. Vorausgesetzt wird beim Leser die Kenntnis des RGB-Farbmodells, welches Farbinformationen pro Pixel durch drei Parameter für den Rot-, Grün- und Blau-Anteil kodiert.

GIF

Bei den auf WWW-Seiten verwendeten Formaten nimmt GIF den Spitzenplatz in punkto Häufigkeit ein. Insbesondere bei künstlich erzeugten Bildern mit einheitlich gefärbten Farbflächen ist es an Kompaktheit nicht zu schlagen (Bild 1).

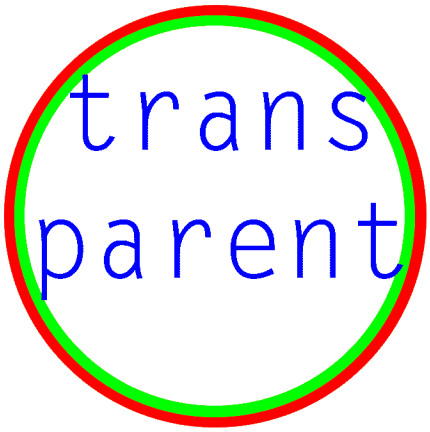


Bild 1: GIF-Bild, 200 × 200 Pixel, teilweise transparent,
unkomprimiert: 116 KB, komprimiert im GIF-Format: 4 KB

Zwei Kompressionsideen tragen zur Datenreduktion bei:

Farbpalette: Statt in jedem Pixel das komplette RGB-Tripel mit 3 Byte = 24 Bit Farbinformation zu speichern, werden für geeignetes p die 2^p wichtigsten Farben in einer Tabelle, genannt Farbpalette, gehalten und über einen p-Bit langen Index referiert. Für $p = 8$ schrumpft der Platzbedarf daher auf ein Drittel.

LZW: Das von *Lempel*, *Ziv* und *Welch* entwickelte und als Patent geschützte Verfahren zur Kompression beliebiger Zeichenfolgen basiert auf der Idee, in einer sogenannten Präfix-Tabelle die Anfangsstücke bereits gelesener Strings zu speichern und wiederholtes Auftauchen derselben Strings durch Verweise in die Tabelle zu kodieren.

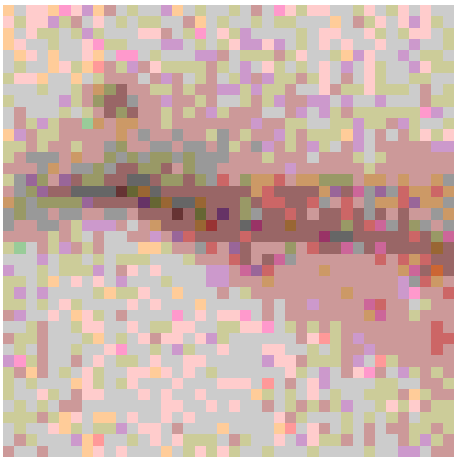
Beide Ansätze zahlen sich insbesondere dann aus, wenn die Vorlage weite Bereiche mit identischer Information enthält, wie es bei computergenerierten Grafiken, wie z.B. Logos, der Fall ist. Zum einen enthält das Bild dann gar nicht die theoretisch verfügbare Zahl von $256^3 \approx 16$ Millionen Farben, sondern nur wenige Dutzend, und kann daher völlig verlustfrei durch eine Palette mit 256 Einträgen dargestellt werden. Zum anderen führen Folgen von identischen Pixelindizes zu kompakten Einträgen in der Präfixtabelle.

Bei der Umwandlung eines True-Color-Bildes in ein Palettenbild mit Hilfe eines Bildbearbeitungsprogramms gibt es zwei Ansätze: Standard oder individuell. Standard bedeutet die Verwendung einer mit 216 repräsentativ gewählten Farben vorbesetzten Farbtabelle (z.B. vom Netscape-Browser). Diese wird eventuell dem gegebenen Bild nicht optimal gerecht, aber die Auswirkungen bei der Präsentation sind vorhersehbar. Individuell bedeutet, eine für die jeweilige Bildvorlage maßgeschneiderte Farbpalette zu vorgegebener Farbzahl 2^p zu konstruieren. Hierzu werden im Median-Cut-Algorithmus die beobachteten Farbhäufigkeiten im RGB-Würfel angeordnet und dieser dann längs eines orthogonal verlaufenden Schnittes (Median) in zwei Subwürfel geteilt (Cut), mit etwa gleichmächtiger Pixelzahl. Diese Partitionierung wird solange fortgeführt, bis 2^p Würfel entstanden sind, die dann jeweils durch einen Index der Farbpalette repräsentiert werden.

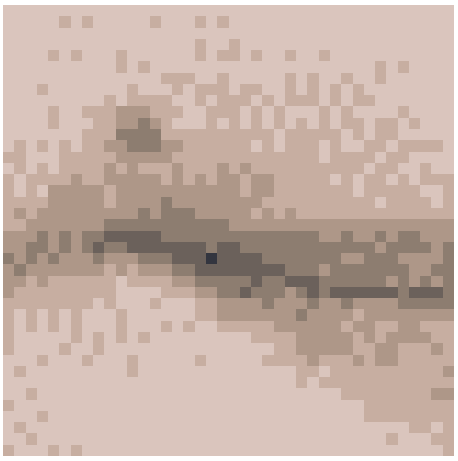
Durch ein sogenanntes *Floyd-Steinberg-Dithering* kann die ohnehin kaum wahrnehmbare Differenz zum Originalbild weiter reduziert werden. Hierzu wird der beim Quantisieren längs einer Scanline verursachte Fehler (d.h. die Differenz zwischen wahrem und approximiertem Farbwert) den nachfolgenden Pixeln vor ihrer Quantisierung zugeschlagen, so daß auf größere Flächen bezogen die Summe der konstruierten Farbwerte mit dem Ausgangsmaterial übereinstimmt. GIF ist streamingfähig; durch ein sogenanntes *Interlacing* wird zunächst nur jede 8. Zeile übertragen und danach jeweils die mittlere Zeile der verbleibenden Lücken.

GIF stößt natürlich an seine Grenzen, wenn es um die naturgetreue Wiedergabe nuancierter Farbverläufe geht. Bild 2 zeigt die entstehenden Probleme anhand einer Hauttönung bei Quantisierung auf 8, 4 und 2 Bit pro Pixel.

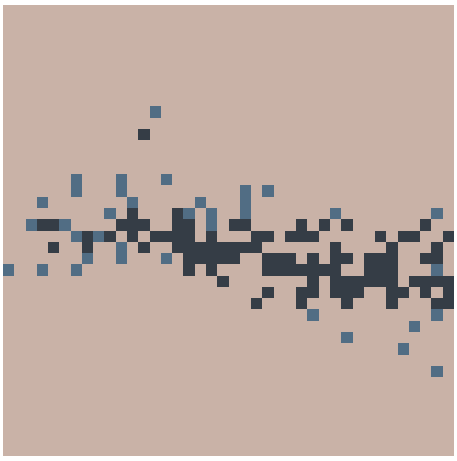
24 Bit pro Pixel,
16 Mill. Farben



8 Bit pro Pixel,
256 Farben



4 Bit pro Pixel,
16 Farben



2 Bit pro Pixel,
4 Farben

Bild 2: Original und quantisierte Versionen einer 40×40 Ausschnittsvergrößerung (Mundpartie), erstellt von einem 1958×2736 True-Color-Bild. Die Zahl der Farben bezieht sich jeweils auf das Gesamtbild. Verwendet wurde der Median-Cut-Algorithmus mit anschließendem Floyd-Steinberg-Dithering.

Das GIF-Format erlaubt die Abspeicherung mehrerer Bilder in einer Datei, welche zudem mit Timing-Information versehen werden können (animated GIF). Hierdurch ist die Erstellung einfacher Animationen möglich, bei denen wenige kleine Grafiken in schneller Abfolge die Illusion eines Films verursachen.

JPG

Wie soeben erläutert, liegen die Stärken von GIF in der Wiedergabe von computergenerierten gleichfarbigen Flächen. Geht es dagegen um die Reproduktion einer fotorealistischen Vorlage, reichen 256 Tabelleneinträge oft nicht aus, sanfte Farbverläufe zu beschreiben. Hier kommt JPEG zu Hilfe, ein Verfahren benannt nach der Gruppe, die es entwickelt hat, der *Joint Photographic Expert Group*, Mitglieder der Standardisierungsgremien *CCITT* und *ISO*.

Zunächst wird das RGB-Bild in den YUV-Raum transformiert, d.h. die Farbinformation wird verlustfrei durch einen Helligkeitsanteil Y und zwei Farbdifferenzen U und V kodiert. Da das Auge für Helligkeitssprünge sensitiver ist als für Farbdifferenzen, kann man nun die Y-Matrix in der vollen Auflösung belassen und in den U, V-Matrizen jeweils 4 Pixel mitteln (4:1:1 Subsampling).

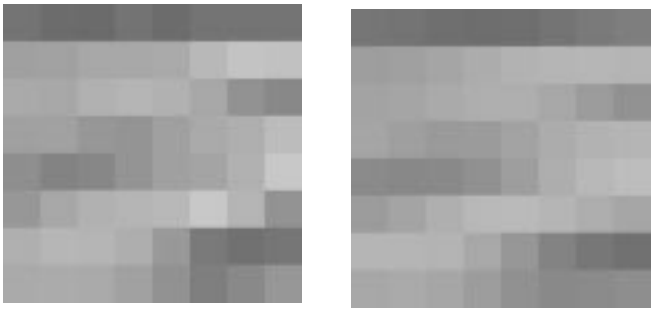
Für je 4 Originalpixel mit insgesamt 12 Bytes werden nun $4 + 1 + 1 = 6$ Bytes benötigt (pro Bildpunkt also $6 \cdot 8/4 = 12$ Bit). Die Reduktion beträgt 50 %.

Nun werden die drei Matrizen in Blöcke mit 8×8 Abtastwerten aufgeteilt. Anschließend durchlaufen die Blöcke folgende Schritte:

Diskrete Cosinus Transformation: Hierdurch wird die 8×8 Ortsmatrix verlustfrei in eine 8×8 Frequenzmatrix umgewandelt. Die Helligkeitsinformation einer Fläche ist nun kodiert als Überlagerung von 64 zweidimensionalen Schwingungen.

Quantisierung: Die errechnete Matrix hat längs einer im Zickzack laufenden Scanline von links oben nach rechts unten Werte abnehmender Größe. Da die Werte rechts unten den hohen, für das menschliche Auge eher unwichtigen Frequenzen entsprechen, werden alle Einträge längs dieser Scanline durch Faktoren zunehmender Größe dividiert.

Komprimierung: Die Folge der quantisierten Koeffizienten wird einer Lauflängenkomprimierung unterzogen und anschließend durch eine empirisch ermittelte Huffman-Tabelle kodiert.



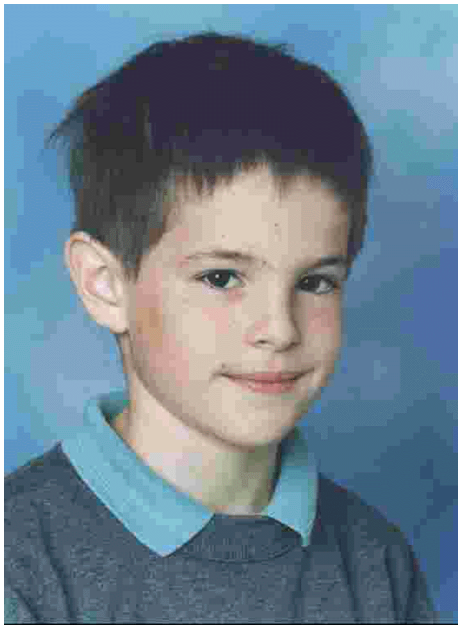
Ausgangsbildmatrix rekonstruierte Bildmatrix

Bild 3: Transformation eines 8×8 Bildausschnittes

Um aus dem komprimierten Bild das Original zu rekonstruieren, werden die Schritte in umgekehrter Reihenfolge und inverser Funktionalität durchlaufen. Bild 3 zeigt die Auswirkungen der drei Phasen angewendet auf eine 8×8 Grauwertmatrix.

Durch die Wahl der Rundungstabelle läßt sich der Tradeoff zwischen Qualität und Kompression beliebig steuern. Ein typisches Farbbild läßt sich auf etwa 5% seiner Originalgröße reduzieren, ohne daß ein menschlicher Betrachter die Detailfehler bemerken könnte. Bei 2% und weniger entstehen deutlich sichtbare Artefakte in Form von einfarbig gefärbten Klötzchen. Bild 4 vergleicht die Auswirkungen zweier Kompressionsgrade an einer Vorlage mit fein abgestuftem Helligkeitsverlauf.

Dateigröße: 7,3 KB
Kompression 22:1



Dateigröße 2,5 KB
Komression 65:1

Bild 4: JPEG-Kompression eines 200×279 True-Color-Bildes, unkomprimierter Speicherbedarf: 163 KB

PNG

Das PNG-Format entstand vor einigen Jahren, als sich das World Wide Web anschickte, zum Marktplatz zu werden und die Firma *Unisys* beschloß, ihr bisher an CompuServe lizenzfrei abgegebenes Patent an LZW zu versilbern. Empört über dieses Ansinnen entschloß sich die Internet-Gemeinde umgehend dazu, ein neues, frei verfügbares Format zu entwickeln. Bei dieser Gelegenheit sollten natürlich auch einige der Unzulänglichkeiten von GIF ausgebügelt werden.

Herausgekommen ist ein Format zum verlustfreien Abspeichern von Farbbildern und steht damit auch in Konkurrenz zu JPEG steht, welches ja - wie oben beschrieben - ein (informationstheoretisch gesehen) verlustbehaftetes Verfahren darstellt. Die wichtigsten Merkmale lauten:

Farbtiefe: Unterstützt werden True Color in 32 und 24 Bit sowie Palettenbilder mit bis zu

8 Bit langen Indizes.

Streaming: Dateien können seriell gelesen werden und eignen sich daher für Kommunikationsprotokolle, welche Bilder on-the-fly erzeugen wollen.

Interlacing: Ein PNG-Bild kann in verschränkter Weise abgelegt und übertragen werden, so daß es beim Empfänger schrittweise eingeblendet werden kann. Das nach seinem Autor Adam Costello benannte Verfahren *Adam7* strukturiert die Übertragung in 7 Phasen gemäß dem folgenden Pattern, welches in wiederholter Form über den gesamten Bildinhalt gelegt wird:

```
1 6 4 6 2 6 4 6
7 7 7 7 7 7 7 7
5 6 5 6 5 6 5 6
7 7 7 7 7 7 7 7
3 6 4 6 3 6 4 6
7 7 7 7 7 7 7 7
5 6 5 6 5 6 5 6
7 7 7 7 7 7 7 7
```

Hierdurch erscheint ein initiales PNG-Bild beim Empfänger 8 mal so schnell wie im GIF-Format, da der erste Durchlauf nur 1/64 der Information überträgt statt 1/8. Ein typisches PNG-Bild kann nach dem 5. Durchlauf (25 % der Daten) schon gut erkannt werden, beim GIF-Bild dauert es doppelt so lange, nämlich bis zum Ende des 3. Durchlaufs (50 % der Daten).

Transparenz: Über einen sogenannten Alphakanal kann in der gewählten Auflösung eine variable Transparenz von völlig durchsichtig bis völlig opak eingestellt werden. GIF-Bilder dagegen erlauben nur die Extremwerte und keine Abstufungen. Bild 5 soll das Ausblenden der Bildinformation in den Hintergrund hinein zeigen; nicht alle Browser können mit dem Alphakanal schon korrekt umgehen.

Gamma-Korrektur: Der nichtlineare Zusammenhang zwischen beabsichtigter Helligkeit und Spannungsgröße zum Ansteuern der Bildröhre variiert zwischen verschiedenen Geräteherstellern und kann durch eine Korrekturangabe im PNG-Bild normiert werden.

Text: Textuelle Zusatzinformation zum Bildinhalt, wie Autor, Copyrighthinweise und Kameraparameter können in die PNG-Datei aufgenommen werden.

Bild 5: PNG-Bild mit Alpha-Kanal
Auflösung 200×200 Pixel
Dateigröße 89 KB

Bewußt wurde von den Autoren keine Möglichkeit der Abspeicherung mehrerer Einzelbilder zum Zwecke der Animation vorgesehen. Dieses von der GIF-Fangemeinde intensiv genutzte Feature wird sicherlich sehr vermißt. Zur mangelnden Verbreitung trägt jedoch in erster Linie bei, daß bis zum Sommer 98 kein Web-Browser PNG-Bilder direkt anzeigen konnte, sondern immer noch separat zu installierende Plugins erforderlich waren. Es bleibt abzuwarten, ob die Lizenzpflicht für GIF tatsächlich am Markt durchsetzbar ist und somit PNG den weiteren Weg ebnet.

Fraktale Kompression

Viele in der Natur vorkommende Strukturen weisen eine starke Selbstähnlichkeit auf. Beispiele dafür sind Gebirgsformationen, Meeresküsten oder Blätter von Pflanzen. Solche in sich wiederholenden Pattern, die quasi beim Hereinzoomen immer wieder zutage treten, lassen sich ausnutzen, wenn man ein umfangreiches Gebilde durch ein kleines Erzeugendensystem darstellen möchte. Sogenannte Iterierte Funktionensysteme sind z.B. in der Lage, mit wenigen Regeln komplexe und natürlich aussehende Bilder zu generieren. Bild 6 zeigt das Blatt eines Farns, welches durch die wiederholte Anwendung von vier affinen Abbildungen entstanden ist.

Bild 6: Farn mit Selbstähnlichkeit

Nach dieser Vorlage ist die Idee der fraktalen Kompression entstanden. Ausgehend von einer gescannten Bildvorlage werden Regelmäßigkeiten und Selbstähnlichkeiten gesucht und dadurch der Bildinhalt als Fixpunkt eine iterierten Anwendung von Transformationsmatrizen definiert.

In stark vereinfachter Form läuft das Verfahren folgendermaßen ab: Zunächst wird das Bild in sogenannte *Domainblöcke* partitioniert. Zu jedem Domainblock wird nun ein *Rangeblock* gesucht, dessen Abbild unter einer affinen Abbildung dem Inhalt des Domainblocks möglichst nahe kommt. Erforderlich für diesen Schritt ist natürlich die Wahl einer geeigneten Metrik, die den Abstand von zwei Pixelmengen definiert. Affine Abbildungen erlauben einfache geometrische Transformationen wie Drehung, Skalierung und Translation. Je nach zugestandener Zeitvorgabe kann der Suchalgorithmus mehr oder weniger viele Rangeblockpositionen mit mehr oder weniger vielen affinen Abbildungen ausprobieren. Zur Vereinfachung beschränkt er sich auf eine feste Zahl möglicher Transformationsmatritzen sowie feste Rangeblockgrößen, die an den Eckpunkten eines 2-dimensionalen Gitters fester Gittergröße positioniert sind.

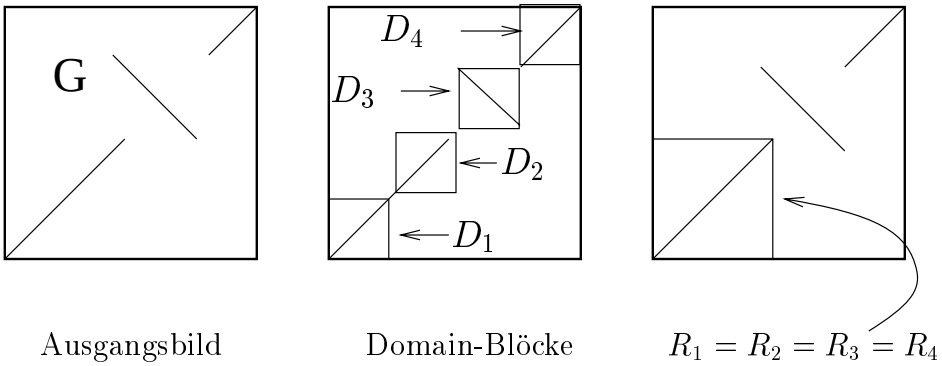
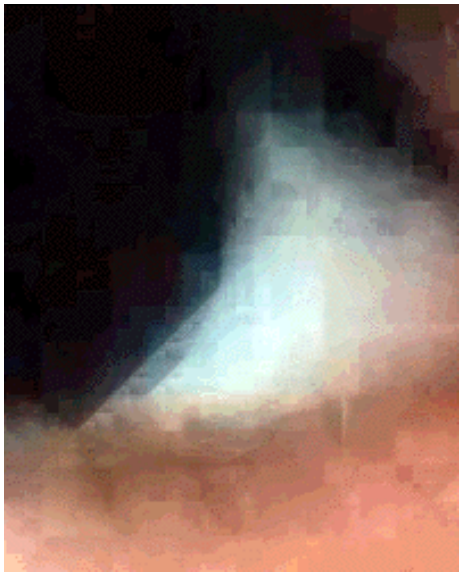


Bild 7: Beschreibung der Vorlage G durch 4 affine Transformationen

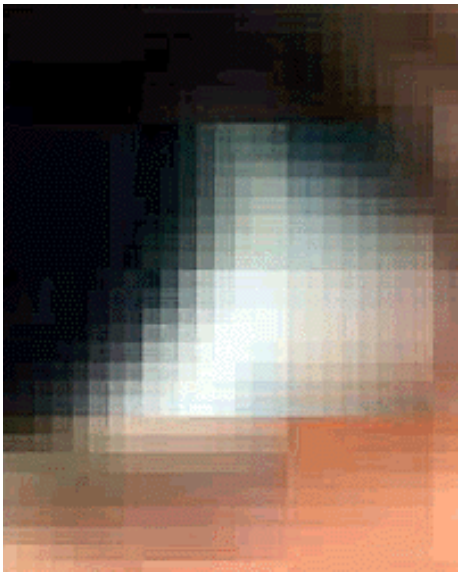
Bild 7 zeigt, wie die Vorlage G zunächst mit 4 Domainblöcken überdeckt wird und dann der Inhalt jedes Domainblocks durch eine affine Transformation des zugehörigen Rangeblocks

dargestellt wird (in diesem Beispiel sind zufällig alle Rangeblöcke identisch).

Bei der Dekompression eines fraktalen Bildes werden nun, beginnend mit einem zufälligen Bildinhalt, die Transformationsmatritzen solange angewendet, bis sich keine Änderung gegenüber der Vorgängeriteration mehr ergibt. Gewisse mathematische Randbedingungen, die bei der Wahl der Matritzen einzuhalten sind, garantieren, daß es einen Fixpunkt gibt, auf den das Verfahren konvergiert.



Fraktal-Bild
Kompressionsrate 72:1



JPEG-Bild
Kompressionsrate 71:1

Bild 8: Vergleich Fraktale Kompression versus JPEG
anhand eines vergrößerten Ausschnitts (Augenpartie) einer True-Color-Vorlage, eingescannt auf 696×924 Pixel

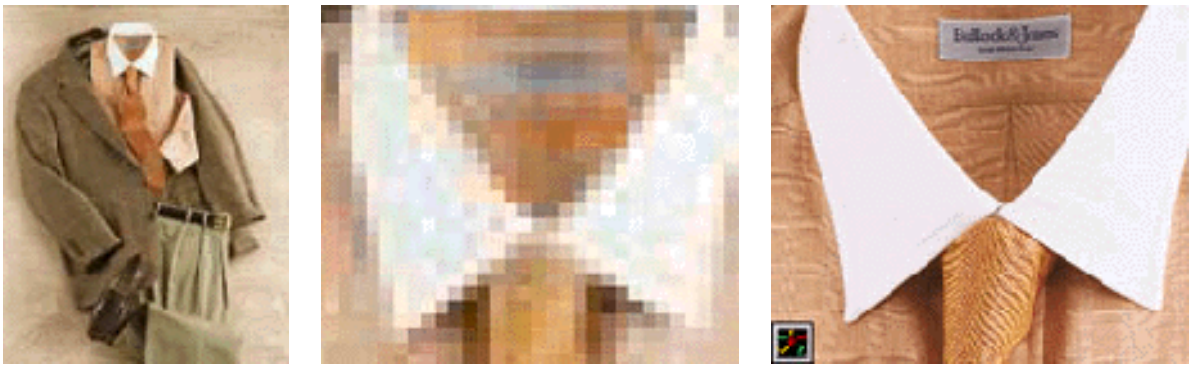
Das Verfahren sieht auf den ersten Blick wie eine typisch akademische Spielerei aus, bei der ein recht eleganter Algorithmus zwar (verblüffenderweise) eine gegebene Vorlage reproduzieren kann, aber ohne praktischen Nutzen scheint. Weit gefehlt ! Die entscheidende Überlegenheit zu den Verfahren GIF, PNG und JPEG liegt in der Auflösungsunabhängigkeit. Nach dem Digitalisieren der Vorlage ist der Bildinhalt gegossen in eine Menge von Transformationsmatritzen. Diese können bei ihrer iterierten Anwendung Bilder jeder gewünschten Auflösung erzeugen, da sie jedes Bildstück nicht durch endlich viele Pixel beschreiben, sondern durch mathematische Beziehungen. Daher läßt sich aus demselben fraktalen Bild sowohl ein briefmarkengroßes Thumbnail als auch ein hochwertiges Poster generieren. Insbesondere kann man in ein fraktales Bild beliebig hereinzoomen, ohne daß die gefürchteten Klötzchen entstehen. Bild 8 zeigt, daß das generierte Bild qualitativ besser sein kann als die eingescannte Vorlage, weil die durch das Digitalisieren entstandenen Rasterpunkte im Fixpunkt der Transformationen weggemittelt werden.

Fraktal komprimierte Bilder haben sich in den Web-Browsern bisher nicht durchsetzen können. Zum einen, weil der Komprimierungsaufwand deutlich höher ist als bei GIF und

JPEG. Zum andern kann der Vorteil der Auflösungsunabhängigkeit erst dann voll ausgespielt werden, wenn ein analoges Foto (hinter einer analogen Linse), ohne den Umweg über das digitale Abtasten direkt in die beschreibenden Transformationsmatritzen umgewandelt werden könnte.

FlashPix

Es wird noch einige Jahre dauern, bis die Zukunftsvisionen der fraktalen Kompression Wirklichkeit werden. Solange wollte die kalifornische Firma LivePicture, Inc. nicht warten. Ihre Lösung der Auflösungsunabhängigkeit heißt FlashPix und wurde mit den heute möglichen Mitteln auf völlig unspektakuläre Weise gelöst. Die Idee geht zurück auf eine gemeinsame Anstrengung der Firmen Kodak, Microsoft und Hewlett Packard bei der Suche nach einem neuen Bildformat und ist recht einfach: Das im Web-Browser zu präsentierende Bild wird in mehrfachen Versionen unterschiedlicher Auflösung auf dem Server gehalten und jeweils der vom Klienten angeforderte Teil in passender Größe übertragen. Typischer Einsatzbereich sind Online Shops, die ihr Warenangebot auf Web-Seiten in einer Weise präsentieren wollen, daß der Kunde ausgehend von einer Totalansicht des Objekts fast beliebig nah heranzoomen und somit alle Details inspizieren kann. Wegen der Verfügbarkeit der abzubildenden Waren für den Shop-Anbieter ist auch die Gelegenheit für die Erstellung hochauflösender Bilder durch die Verwendung einer professioneller Kameraausrüstung gegeben.



Gesamtansicht JPEG-Ausschnitt FlashPix-Ausschnitt

Bild 9: Auswirkung der FlashPix-Technologie in einem Online-Shop

In der einfachsten Form erfordert dies beim Klienten keine weiteren Vorkehrungen, da der LivePicture Image Server den angeforderten Bildteil als JPEG-Bild liefert. Welcher Teil eines bereits im Browser geladenen Bildes in erhöhter Auflösung angefordert wird, kann durch die Koordinaten des Mausklicks bestimmt werden. Auf diese Weise führt der Betrachter einen Zoomvorgang aus, der ohne die übliche Rasterung abläuft. Bild 9 demonstriert den Effekt. Daß auf dem Server ggf. große Datenmengen vorgehalten werden müssen, kümmert den Klienten wenig; er profitiert vielmehr von kurzen Kommunikationszeiten, da immer nur der Teil übertragen wird, den er benötigt.

Die komfortablere Version von FlashPix verwendet ein Plugin oder setzt einen Java-fähigen Browser voraus, der ein leichtgewichtiges Applet vorweggeschickt bekommt. Vorteil dieser

beiden Ansätze ist eine höhere Funktionalität beim Betrachten der Bilder wie z.B. Vor- und Zurückzoomen sowie Verschieben des Betrachterausschnitts.

Neben der multiplen Auflösung, die ja bereits auf der Kodak Foto CD Verwendung fand, gehören folgende Features zum FlashPix-Format:

Kompression: Für jede Auflösungsstufe wird das Bild nach dem JPEG-Verfahren kodiert. Es besteht auch die Möglichkeit, ohne Kompression zu speichern.

Text: Textuelle Zusatzinformation kann, ähnlich dem PNG-Ansatz, gespeichert werden. Diese Daten sind strukturiert anhand der Schlüssel, wie z.B: Quelle, Urheber, Inhalt, Kameraeinstellungen, Filmcharakteristika, Scancharakteristika.

Farbräume: Neben unkalibrierten RGB-Werten gibt es z.B. noch monochrom und Photo YCC

Transformation: FlashPix-Bilder halten neben den Rohdaten einige wenige Filter- und Transformationsmatritzen vor, die jeweils nur auf den zur Zeit sichtbaren Teil angewendet werden müssen. Beispiele sind Translation, Rotation, Skalierung, Schärfe, Kontrast, Helligkeit und Farbbalance. Hierdurch reduziert sich der Hauptspeicherbedarf eines Bildverarbeitungsprogramms und gleichzeitig erhöht sich seine Bearbeitungsgeschwindigkeit, da die Masse der Bilddaten unbearbeitet auf der Platte verbleiben können.

LivePicture eröffnet mit seinem FlashPix-Plugin auch die Möglichkeit zum interaktiven Betrachten von 360-Grad Panoramen. Bei der Herstellung solcher Bilder ist der Ausgangspunkt eine Folge von Einzelschnappschüssen, die mit konventioneller Kamera im Kreis erstellt wurden. Nach dem Digitalisieren werden mit dem Werkzeug *LivePicture PhotoVista* die Aufnahmen, unter automatischer Berücksichtigung der überlappenden Bereiche, zu einem länglichen Schlauch im JPEG-Format zusammengeschweißt. Bild 10 zeigt die aus einer Sequenz von 12 Einzelbildern generierte JPEG-Datei. Vorteil der Verwendung dieses gängigen Bildformats ist natürlich die Möglichkeit, mit jeder Bildverarbeitungssoftware weitere Manipulationen an dem Rundblick durchführen zu können.

Bild 10: JPG-Vorlage für 360-Grad-Panorama

PhotoVista erzeugt neben dem JPEG-Bild noch eine gleichnamige Datei mit der Endung `ivr`. Die dort gespeicherte Information benutzt das Plugin zur perspektiv richtigen Umrechnung des JPEG-Bildes in ein interaktiv drehbares Panorama. Als multimediale Zugabe können der Datei noch Klanginformation und Animationen beigefügt werden.

Unter der Adresse `http://www.livepicture.com/zr/solutions/advanced/content.html` finden sich mehrere Beispiele für Bilder im FlashPix-Format.

DjVu

In den Forschungslaboratorien des amerikanischen Kommunikationsspezialisten AT&T wurde in diesem Jahr ein neues Bildformat ausgebrütet, welches die Grafikindustrie revolutionieren könnte: DjVu, gesprochen *deschah wü*. Hiermit ist der aus dem französischen stammende Begriff *deja vu* (schon mal gesehen) gemeint, und dieser Name kennzeichnet das typischen Einsatzgebiet dieses Formats: Abspeichern von bereits existierenden Postern, Büchern, Landkarten, etc. Die Überlegenheit von DjVu zeigt sich insbesondere bei Vorlagen, in denen Schrift und Zeichnungen auf farbigem Hintergrund plaziert sind, wie dies bei historischen Dokumenten der Fall ist. Mit dieser Mischung von hochfiligranen Mustern innerhalb von weichen Farbverläufen hat JPEG immer besonders zu kämpfen, denn die diskrete Cosinus-Transformation kannscharfe Kontrastsprünge nur dann in erträglicher Qualität abbilden, wenn die Koeffizienten der hochfrequenten Anteile nicht zu sehr quantisiert wurden, was aber nur einen geringen Kompressionsgrad bedeutet.

DjVu löst dieses Problem durch eine Separierung der Vorlage in Hintergrund- und Vordergrundinformation sowie durch spezielle darauf angewandte Kompressionsverfahren. Für den Hintergrund mit seinen kontinuierlich und meistens unscharf verlaufenden Farben wird eine sogenannte *IW44 wavelet* Kompression eingesetzt, die ähnlich wie JPEG durch Überführung der Ortsmatritzen in den Frequenzraum versucht, die Bildinformation als Überlagerung weniger Grundschwingungen darzustellen. Die Kodierung der scharfkantigen und meist einfarbigen Vordergrundanteile geschieht separat davon. Insgesamt kommen drei patentierte Verfahren zum Einsatz: *JB2*, eine Zwei-Ebenen-Bildkompression, *ZP-coder*, ein adaptiver arithmetischer Entropie-Kodierer sowie die *IW44* Maskierungstechnik zum Einsparen von Bits auf dem Teil des Hintergrunds, der durch Text abgedeckt ist.

Gesamtinformation des DjVu-Bildes

Vordergrundanteil des DjVu-Bildes

Bild 11: strukturierte Bildinformation durch Ebenen unterschiedlicher Auflösung

Bild 11 zeigt, daß es nicht nur auf Bildern mit aufgedrucktem Text gelingen kann, feine detailreiche Strukturen von einem eher groben Hintergrund zu trennen.

Bild 12 zeigt das Potential von DjVu im direkten Vergleich mit der JPEG-Technik anhand einer eingescannten Vorlage aus einem Versandhauskatalog. Deutlich zu sehen sind die scharfkantigen Schriftzüge, die JPEG bei diesen Kompressionsraten nicht mehr auflösen kann. Im Mittel sind folgende DjVu-Dateigrößen zu erwarten: Das Digitalisieren einer typischen farbigen DIN-A-4-Magazin-Seite mit 300 DPI (dots per inch) ergibt 25 MB unkomprimierte Daten und dagegen nur zwischen 40 und 70 KB DjVu-Daten. Dies ist etwa 5 - 8 mal besser als JPEG. Beim Digitalisieren einer schwarz/weißen Faxvorlage entstehen um etwa 3 - 8 mal kleinere Dateien als mit dem bekannten *CCITT-G4*-Algorithmus.

Ausschnitt aus JPEG-Bild
45 KB, Kompression 532:1

Ausschnitt aus DJVU-Bild
41 KB, Kompression 584:1

Bild 12: Vergleich JPEG versus DJVU anhand eines Posterscans auf 2460 × 3252 Pixel
Platzbedarf der unkomprimierten Datei: 22,8 MB

Bei Verwendung des Plugins ist der Benutzer in der Lage, Vordergrund und Hintergrund separat anzeigen zu lassen. Außerdem besteht die Möglichkeit eines interaktiven Zooms, da durch die AT&T-Technologie immer nur der in Realzeit errechnete relevante Ausschnitt für den Betrachter angezeigt wird.

Unter der Adresse `http://djvu.research.att.com/home_mstr.htm` finden sich mehrere Beispiele für Bilder im DjVu-Format.

Flash

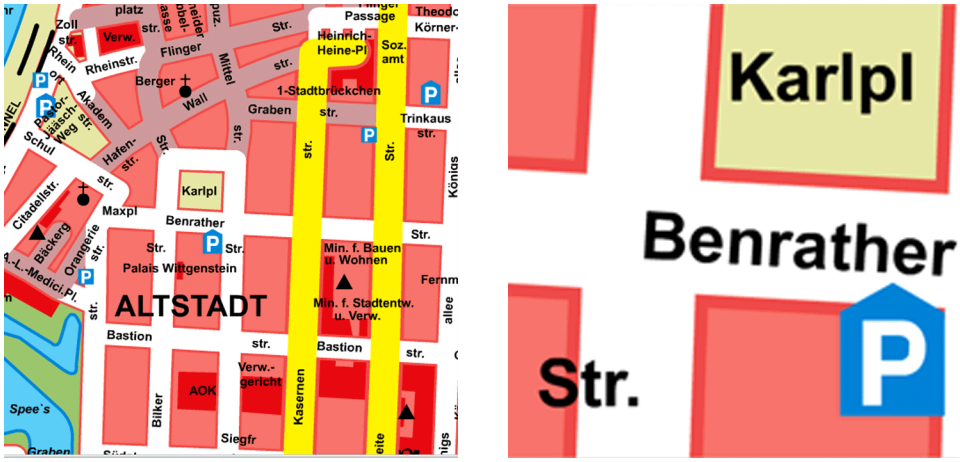
Alle bisher vorgestellten Formate sind pixelorientiert. Das bedeutet, die dargestellten Bilder sind ursprünglich duch Einscannen einer Vorlage mit fest vorgegebener Auflösung entstanden und das Ziel von kleinen Dateigrößen wurde durch trickreiche Verfahren erreicht, die Pixelinformationen reduzieren, ohne den visuellen Eindruck wesentlich zu verändern.

Eine völlig andere Ausgangslage ergibt sich, wenn die präsentierten Bilder computergenerierte Daten darstellen, da nun eine kompakte Repräsentation im Rechner gewählt werden kann, aus der das externe Erscheinungsbild erst errechnet wird. Für geometrisch strukturierte Vorgaben bietet sich das Vektorformat an, da hier mit einer Handvoll von Koordinatenangaben die Plazierung von elementaren Objekten wie Linien, Rechtecken und Kreisen möglich ist. Auch Kurven sind durch Bezier- und Splines-Techniken über wenige Parameter zu spezifizieren. Als Resultat kann sich der Betrachter in eine Vektorgrafik beliebig hineinzoomen, ohne daß Rastereffekte auftreten.

Die Firma Macromedia hat im Jahre 1996 das Format *Flash* auf den Markt gebracht und zur Unterstützung einer schnellen Verbreitung kurz darauf auch veröffentlicht. Damit sind Entwickler von Grafikprogrammen in der Lage, in ihre Applikationen das Format zu inte-

grieren. Flash wird im Bundle mit Netscape Communicator und Microsoft Internet Explorer ausgeliefert und pro Tag etwa 150.000 mal per Download von der Macromedia Homepage geholt.

Bild 13 zeigt durch zwei Schnappschüsse eine sinnvolle Anwendung des Flash-Formats. Ein Stadtplan ist sowohl in der Totalen als auch bis auf Straßengröße runtergebrochen vom selben Datensatz präsentierbar. Der an diesem Beispiel beobachtete Skalierungsunterschied von etwa 1:6 könnte durch ein pixelorientiertes Programm nur unter Bereitstellung riesiger Dateien bewerkstelligt werden.



Gesamtansicht des Stadtplans

Teilansicht des Stadtplans

Bild 13: Rasterfreies Zoomen bei Dateien im Flash-Format durch Vektorgrafik

Ein weiterer Vorteil des Vektorformats liegt in seiner Animationsfähigkeit. Während ein Film basierend auf Pixelgrafik für ein sich bewegendes Objekt eine Sequenz von Einzelbildern explizit abspeichern muß, reicht es nun, allein den räumlichen Bewegungspfad des Objektbezugspunkt geometrisch zu beschreiben. Flash-Dateien sind streaming-fähig, d.h. der Player kann mit dem Abspielen beginnen, bevor die komplette Szene übertragen worden ist. Die Möglichkeit der Einbindung von Audiofiles soll hier nur am Rande erwähnt werden.

Zusammenfassung

Eins ist sicherlich deutlich geworden: **das** ideale Grafikformat für das Internet gibt es nicht. Dazu sind die Anforderungsprofile der Anwender zu verschieden. GIF hat seine Berechtigung für kleine, animierte Logos mit wenigen, klaren Farben. JPEG bietet sich an für Landschafts- und Personenfotos mit wenig Details. PNG soll verlustfrei True-Color-Bilder ins Netz bringen; ob da ein Bedarf besteht ist unklar. Die auflösungsunabhängige Fraktale Kompression ist sicherlich ihrer Zeit noch ein wenig voraus. FlashPix ist eine praktikable Lösung für Online Shops, die auch ihren Kunden mit Browsern der älteren Generation die Illusion des Zooming in Warendetails bieten wollen. DjVu ist sicherlich ein großer Durchbruch insbesondere für die Kompression von Bild/Schrift-Mixturen. Das Vektorformat Flash schließlich kann in sehr

kompakter Weise geometrische Strukturen kodieren, die sowohl in der Totalen als auch im Detail betrachtet werden sollen. Für welche Variante der Anbieter sich auch entscheidet: Ein Mehrwert an Informationsgehalt gegenüber rein textbasierten Daten ist es allemal. Es bleibt zu hoffen, daß die entwickelten Formate vom Anwender im richtigen Maße zur Visualisierung seines Anliegen verwendet werden.