Einleitung

Bei Multimediasystemen geht der Trend zu rein digitalen Systemen. Insbesondere bei der Darstellung von Farbbildern und Videos sind dabei jedoch enorme Datenmengen zu verarbeiten. Eine direkte Bearbeitung (in Realzeit) dieser Datenmengen ist mit den heutigen Computer- und Netzwerksystemen und deren Kapazitäten kaum möglich. Anhand der folgenden Beispiele lassen sich die auftretenden Datenmengen und die damit verbundenen Probleme schnell veranschaulichen.

1. Problemstellung

Bei einer sehr guten Auflösung von 1280x1024 Bildpunkten besteht ein Computerbildschirm insgesamt aus ca. 1,3 Mio. Bildpunkten (*pixel*). Bei einer Echtfarbendarstellung (24 Bit Farbtiefe), was einer Darstellung mit etwa 16,7 Millionen Farben entspricht, sind somit für die Speicherung des Inhaltes eines Computerbildschirm etwa 3,75 MegaByte (MB) erforderlich.

Ein Scanner mit 300 dpi erzeugt beim Scannen einer Din A4 Seite eine Datenmenge von etwa 1 MB multipliziert mit der Anzahl der Bits, die für die Farbdarstellung notwendig sind. Schon bei einer Farbtiefe von 8 Bit, was 256 Farben oder Graustufen entspricht, sind das ca. 8 MB.

Ein Videobild benötigt je nach Fernsehnorm etwa zwischen 750 und 920 kiloByte (kB), woraus bei 25 Bildern pro Sekunde (*frames per second*) ein Datenstrom von 20 bis 22 MB pro Sekunde Videofilm resultiert.

Das Ziel der Datenkompression ist es nun, den benötigten Speicherplatz für ein Bild oder eine Videosequenz zu verringern, d.h. zu komprimieren, um den komprimierten Datenstrom anschließend speichern oder übertragen zu können. Bei der Wiedergabe wird aus dem komprimierten Datenstrom das “ursprüngliche” Bild oder die “ursprüng-

liche” Videosequenz rekonstruiert. Es ist nicht immer möglich, das ursprüngliche Bild originalgetreu zu rekonstruieren und oft genügt es auch, daß der Betrachter den (annähernd) gleichen Eindruck erhält.

Abb. 1: *Prinzip der Datenkompression*

Daraus ergeben sich eine Reihe von Kriterien für die Charakterisierung von Kompressionsverfahren.

Verlustfreie Kompression

Bei der verlustfreien Kompression (*lossless compression*) ist es möglich, aus dem komprimierten Datenstrom das Original vollständig zu rekonstruieren.

Verlustbehaftete Kompression

Bei der verlustbehafteten Kompression (*lossy compression*) ist das aus dem komprimierten Bitstrom rekonstruierte Bild gegenüber dem Originalbild verändert. Ziel ist es dabei, die Verluste so zu gestalten, daß durch das menschliche visuelle Wahrnehmungssystem möglichst wenig davon erkannt wird. Originalbild und Rekonstruktion sind verschieden, sollen aber denselben Eindruck hervorrufen.

Kompressionsgrad

Für die Angabe des Kompressionsgrades oder der Kompressionsrate (*compression ratio*) sind zwei Varianten gebräuchlich.

1. Durchschnittliche Anzahl der Bits pro Pixel
2. Verhältnis der ursprünglichen zur komprimierten Datenmenge

Wird z.B. ein Bild mit einer Auflösung von 256x240 Pixeln und einer Farbtiefe von 8 Bits auf 15kB komprimiert, dann bedeutet dies eine Reduktion der ursprünglichen 8 Bits/Pixel auf 2 Bits/Pixel, was einem Verhältnis der ursprünglichen zur komprimierten Datenmenge von 4:1 entspricht.

2. Historischer Abriß

Die meisten Ansätze bei der Kompression von Bild- und Videodaten basieren auf Techniken, bei denen eine Abbildung der Eingangsdaten in einen Frequenzraum vorgenommen wird. In diesem Frequenzraum werden die Daten dann so bearbeitet, daß es im Laufe der Kodierung zu einer Kompression der Eingangsdaten kommt. Daraus ergibt sich als eine wesentliche Eigenschaft dieser Kodierungstechniken, daß die Darstellung und der Inhalt des Bildes bzw. des Videos unberücksichtigt bleiben. Als Konsequenz aus diesem Ansatz teilen all diese Techniken dasselbe Konzept für die Darstellung von Bildern und Bildsequenzen. Bilder und Bildsequenzen bestehen immer aus den zu kodierenden Grundelementen *Pixel* oder *Pixelblock*. Eine weitere Gemeinsamkeit dieser Techniken ist, daß die Berücksichtigung des menschlichen Sehsystems HVS (*human visual system*) bei der Konstruktion der Kodierer / Dekodierer nur rudimentär ist. Aufgrund dieser Eigenschaften und der Unterschiede zu neueren Ansätzen, von denen die hier zu diskutierende fraktale Bildkodierung einer ist, werden solche Techniken i. a. als *first generation approaches*, als Methoden der ersten Generation bezeichnet.

Als Anfang der 90-er Jahre Standards für die Kompression von Bildern (JPEG, [ISO94b]) und Videos (H.261 [ITU90], MPEG-1 [ISO93]) definiert wurden, stellten diese *first generation*-Techniken die Basis für die Kodierung der Daten dar, obwohl schon damals die Grenzen der Verfahren zu sehen waren. Insbesondere für den Fall der sogenannten *low bitrate*-Kodierung werden die Grenzen dieser Methoden sichtbar. Bei der *low bitrate*-Kodierung geht man für die Speicherung bzw. Übertragung von einem Übertragungskanal mit einer Bandbreite von kleiner 64 kBits/s aus. Dieser Problematik nahm sich 1995 ein neuer Standard H.263 [ITU95] an, der speziell für Video-Konferenzen entwickelt wurde und nicht universell einsetzbar ist.

Mit den sogenannten *second generation*-Methoden zur Bild- und Video-Kodierung versucht man nun die Nachteile und Begrenzungen der ersten Ansätze aufzuheben. Erste Veröffentlichungen zu diesem Thema gab es bereits 1985 [Kunt85], obwohl ein Vorläufer dieses Ansatzes in Form einer modellbasierten Videokodierung bereits 1851 existierte. G. R. Smith demonstrierte damals auf einer Ausstellung in London einen *comic electric telegraph*, einen Telegraphen mit einem aus Metall bestehenden flexiblen Modell des menschlichen Kopfes [Pearson95]. 1983 wurde diese Idee durch Forchheimer und Fahlander [Forch83] wieder aufgegriffen. Basierend auf Bildanalyse- und Syntheseverfahren stellten die beiden Autoren ein System vor, welches mit Hilfe von Computer-Animation Video-Sequenzen übertragen sollte. Auch wenn diese Beispiele zugegebenermaßen exotisch sind, so zeigen sie dennoch grundlegende Eigenschaften für Kodierungsmethoden der zweiten Generation.

Als *second generation approach* zur Kodierung und Kompression von Bildern und Videosequenzen bezeichnet man solche Ansätze, die neuere und effizientere Ansätze zur Darstellung der Bilder und Videos verwenden, als es die frequenzorientierten Methoden der ersten Generation tun. Dabei wird das HVS integraler Bestandteil eines Systems zu Kodieren und Dekodieren. Dieser Aspekt wird auch in der vorliegenden Arbeit berücksichtigt, wenn z.B. verschiedene Verfahren zur Partitionierung von Bildern und deren Auswirkung auf das HVS betrachtet werden.

Verfahren zur Kodierung von Bild- und Videodaten arbeiten generell in zwei Schritten. In einem ersten Schritt werden die Bild- oder Videodaten in eine Folge von Nachrichten zerlegt und im zweiten Schritt ordnet man diesen Nachrichten Codewörter zu. Betrachtet man nun den Unterschied zwischen den beiden Generationen unter diesem Aspekt, so läßt sich sagen, daß die Techniken der ersten Generation vorwiegend versucht haben, den zweiten Schritt zu optimieren, was durchaus der historischen Entwicklung aus der Informationstheorie entspricht. Methoden der zweiten Generation legen dagegen wesentlich mehr Wert auf den ersten Schritt, welcher sich mit der Darstellung der Bild- und Videodaten beschäftigt. Aufgrund der Integration des HVS werden nicht mehr Pixel bzw. Pixelblöcke betrachtet und kodiert, sondern Bildelemente oder Objekte. Wie diese Objekte aussehen und welche Eigenschaften sie besitzen, ist natürlich von Ansatz zu Ansatz verschieden. Diese Vorgehensweise impliziert nämlich, daß ein Bild oder eine Videosequenz für die Bestimmung der Objekte zuerst analysiert und segmentiert werden muß. Das Ergebnis der Analysephase sind die zu kodierenden Bildobjekte. Das Ergebnis des gesamten Systems im Bezug auf die Bildqualität und die Kompressionsrate hängt natürlich sehr stark davon ab, wie gut die gewonnenen Bildobjekte das Ausgangsbild bzw. die Ausgangsvideosequenz beschreiben. Die Gesamteigenschaften eines Systems zu Bildkodierung und Dekodierung, wie z.B. Kompressionsrate, sind somit sehr stark von den analysierten und kodierten Bildobjekten abhängig. Wie in den folgenden Kapiteln für den Fall der fraktalen Kodierung gezeigt wird, ist die Kompressionsrate dabei nur eine Eigenschaft, wenn auch die interessanteste.

Die Methoden, welche man i.a. als *second generation approaches* bezeichnet, lassen sich in drei große Klassen einteilen:

– Segmentation-based schemes (Bildkodierung mittels Segmentierung)
– Model-based schemes (Modellbasierte Bildkodierung)
– Fractal-based schemes (Fraktale Bildkodierung)

Der besondere Augenmerk dieser Arbeit liegt dabei auf dem dritten Thema, der fraktalen Bildkodierung.

3. Aufbau der Arbeit

Die Beschäftigung mit dem Thema “fraktale Bildkodierung” gliedert sich in verschiedene Themenbereiche, was sich auch im Aufbau dieser Arbeit niederschlägt. Nach dieser Einleitung wird im ersten Kapitel ein Überblick über Verfahren und Techniken aus dem Feld der verlustbehafteten Datenkompression gegeben. Neben einer Erläuterung und Bewertung dieser Verfahren dient dieses Kapitel als Grundlage für eine spätere Einordnung der fraktalen Bildkodierung in den allgemeinen Kontext der verlustbehafteten Datenkompression.

Kapitel 2 besteht aus einer Darstellung und Aufarbeitung von Michael F. Barnsley’s Modell zu fraktalen Bildkodierung, wie er es in mehreren Arbeiten vorgestellt hat, wie z.B. in [Barnsley88a] und [Barnsley93]. Die Darstellung dieses Modells geschieht

dabei unter dem Aspekt einer möglichen Implementierung und diskutiert das Modell am Ende des Kapitels auch unter diesem Gesichtspunkt. Die Aufarbeitung umfaßt eine Vervollständigung der Modelldarstellung unter mathematischen Aspekten und eine Reduktion von Aspekten, die für ein Verständnis oder eine Implementierung des Modells nicht notwendig sind. Unter diesen Gesichtspunkten ist Kapitel 2 eine kompakte Darstellung von Barnsley's Modell zur fraktalen Bildkodierung unter dem Fokus einer möglichen Implementierung, gekoppelt mit einer kritischen Diskussion der Eignung dieses Modells für dieses Vorhaben.

Der Aspekt der Implementierung tritt im anschließenden Kapitel 3 vollständig in den Vordergrund, dessen Ziel die Umsetzung der Ideen von Barnsley in ein lauffähiges System zur Kompression von Bilddaten ist. Kapitel 3 beinhaltet deshalb die Entwicklung eines diskreten Modells zur fraktalen Bildkodierung, mit dessen Hilfe eine Implementierung und dessen Eigenschaften beschrieben werden können. Aufbauend auf den Ideen von Barnsley und der Erkenntnis, daß dessen Modell für eine Implementierung nicht ausreichend ist, wird in diesem Kapitel ein vollständig diskretes Modell zur fraktalen Bildkodierung entwickelt, wobei der Grundgedanke der Implementierung die Entwicklung des Modells maßgeblich trägt und beeinflußt.

In Kapitel 4 dient dieses diskrete Modell zur Darstellung einer Implementierung und deren allgemeinen Eigenschaften, ohne auf eine konkrete Implementierung Bezug nehmen zu müssen. Das Ziel dieses Kapitels ist die Bestimmung des Leistungsspektrums von Systemen zur fraktalen Bildkodierung als Vorbereitung auf eine Einordnung in den allgemeinen Kontext zur verlustbehafteten Datenkompression.

Ergebnisse, die auf eine explizite Implementierung Bezug nehmen, werden im anschließenden Kapitel 5 betrachtet, dessen Inhalt die Darstellung experimentell gefundener Ergebnisse ist. Es erfolgt eine Diskussion der Ergebnisse zur Festigung der in Kapitel 4 getroffenen Aussagen mit dem Ziel der Klassifikation von verschiedenen Verfahren zu fraktalen Bildkodierung. Diese Diskussion mündet in einer Zusammenfassung dieser Arbeit, in welcher die Grenzen und Fähigkeiten fraktaler Bildkodierung aufgrund einer Einordnung in allgemeinen Kontext der verlustbehafteten Datenkompression vorgenommen wird. Basis der Einordnung sind einerseits die Erkenntnisse über die fraktale Bildkodierung aus Kapitel 3 bis 5 und andererseits die Ausführung über weitere Bildkompressionsverfahren aus dem ersten Kapitel. Das Ziel der Einordnung ist eine Standortbestimmung der fraktalen Bildkodierung im Kanon der Verfahren zur verlustbehafteten Datenkompression, weshalb auch ein Vergleich mit diesen Verfahren den Abschluß dieses Unterkapitels bildet.

Den Abschluß dieser Arbeit bildet ein Ausblick auf mögliche Weiterentwicklungen von Methoden zur fraktalen Bildkodierung und deren Einsatz für die Kodierung von Videos. Mit Bezug auf Kapitel 1 und den dort vorgestellten Verfahren zur Videokodierung werden die Schwierigkeiten bei einem Einsatz fraktaler Kodierungsmethoden zur Videokodierung aufgezeigt sowie die Notwendigkeit zur Entwicklung neuer Verfahren für die Entfernung zeitlicher Redundanzen begründet.