



HOCHSCHULE COBURG

Hochschule für angewandte Wissenschaften Coburg

Fakultät Elektrotechnik und Informatik

Studiengang: Informatik

Bachelorarbeit

Entwicklung einer hardwarebeschleunigten Berechnung der Mandelbrotmenge auf einem FPGA

von

Daniel Kirchner

Matrikelnummer: 02219415

Abgabe der Arbeit: 15.07.2019

Betreut durch: Prof. Oliver Engel, Hochschule Coburg

Zusammenfassung

Im Rahmen dieser Bachelorarbeit wurde eine hardwarebeschleunigte Visualisierung der Mandelbrotmenge auf einem FPGA realisiert. Hierfür werden diverse mathematische und designtechnische Performanceoptimierungen vorgestellt, welche dann in ein paralleles FPGA-Design implementiert wurden. Weiterhin sollen einige Eigenschaften und Besonderheiten der Mandelbrotmenge und von Fraktalen im Allgemeinen aufgezeigt werden.

Das Projekt wurde für das Zybo Zynq-7000 Trainer Board entwickelt, welches über einen VGA-Output die Repräsentation des Fraktals in Form eines 800x600@60Hz Videosignals ausgibt. Zur optimalen Ausnutzung der auf diesem Board gegebenen Ressourcen (DSPs, BRAM) wurde die *Vivado Design Suite* mit dem integrierten IP-Katalog verwendet.

Inhaltsverzeichnis

Zusammenfassung	1
Inhaltsverzeichnis	2
Abbildungsverzeichnis	3
Tabellenverzeichnis	4
Codebeispielverzeichnis	5
Abkürzungsverzeichnis	6
1 Einleitung	7
1.1 Motivation	7
1.2 Aufgabenstellung	7
1.3 Mitgelieferte Skripte	8
2 Technische Grundlagen	9
2.1 FPGAs	9
2.2 VGA-Schnittstelle	10
3 Theoretische Grundlagen	13
3.1 Fraktale	13
3.1.1 Natürlich vorkommende Fraktale	13
3.1.2 Fraktale in der Mathematik	14
3.2 Die Mandelbrotmenge	16
4 Umsetzung	19
4.1 Gewähltes Zahlenformat	19
4.2 Systemüberblick	22
4.3 Komponentenbeschreibung	23
4.3.1 Mandelbrot-Core	23
4.3.2 Mandelbrot-Koordinator	26
4.3.3 Speicher	28
4.3.4 Lookup Tables	29
4.4 Peripherie und funktionale Beschreibung	30

5	Umsetzung des Systems	31
6	Ausblick & Fazit	32

Abbildungsverzeichnis

1	Beispielhafter Aufbau einer 2-Input LUT	9
2	Logikblock, aus [1]	9
3	VGA-Timing/Aufbau für eine 640x480 Auflösung, aus [2]	11
4	Fraktal definierter Baum, Skript: <i>baum.py</i> , abgewandelte Version von [3]	13
5	Blumenkohl, von Rainer Renz	14
6	200 Iterationsschritte für $c = -0.55 + 0.46i$ (rot markierter Punkt)	16
7	Mandelbrotmenge in Graustufe	17
8	Funktionsweise der Funktion <i>fixlen</i>	21
9	Gesamtsystem mit 2 Kernen	22
10	Mandelbrot-Core, schematische Darstellung	23
11	Funktionsweise Mandelbrot-Core, Impulsdiagramm	24
12	Mandelbrot-Koordinator, schematische Darstellung	26
13	Mandelbrot-Koordinator, schematische Darstellung	27
14	Block Memory Generator, IP-Bauteil	28

Tabellenverzeichnis

1	VGA-Werte für ein 800x600@60Hz Signal	12
2	Bereich der Mandelbrotmenge	19
3	Farbtabelle, Übergang von Rot nach Grün	30

Codebeispielverzeichnis

1	Algorithmus zur Berechnung der Iterationszahl eines c	17
2	VHDL-Funktion <i>fixlen</i>	21

Abkürzungsverzeichnis

JAX-RS Java API for RESTful Web Services

1 Einleitung

1.1 Motivation

Die stets wachsende Zahl von Komponenten die auf einem mikroelektronischen Bauteil pro Zeiteinheit untergebracht wird, ist ein Phänomen, welches Gordon Moore schon im Jahr 1965 aufgefallen ist [4]. Die populäre, nach ihm benannte Beobachtung, dass die Anzahl der Transistoren pro integriertem Schaltkreis exponentiell mit der Zeit ansteigt ist allgemein als das Mooresche Gesetz bekannt.

Diese Gesetzmäßigkeit machte es möglich den stets wachsenden Leistungsanforderungen an moderne Hardware gerecht zu werden, indem immer mehr (und komplexere) identische Allzweck-Prozessoren (Kerne) pro CPU verbaut wurden.

Dieses Vorgehen kann jedoch nicht unbegrenzt lange betrieben werden, da die heute verwendeten MOS-Transistoren sich rapide ihren physikalischen Grenzen annähern. Ein besserer Umgang mit dem stetig steigenden Bedarf an Rechenleistung ist die Entwicklung von spezialisierter Hardware, welche zwar nur ein kleines Aufgabenspektrum abdeckt, dies jedoch mit hoher Performanz und Energieeffizienz tut.

Ein Beispiel hierfür ist die moderne Grafikkarte (GPU), welche dem Prozessor Darstellungsrechnungen abnimmt, wodurch dieser mehr Zeit hat, andere Aufgaben zu übernehmen. Die Grafikkarte führt diese Aufgaben mit enorm hohem Durchsatz und niedrigen Berechnungszeiten durch, welche ein herkömmlicher Prozessor alleine nicht erreichen könnte.

Auch andere Hardwarekomponenten, wie die Netzwerkkarte, kryptographische Beschleuniger, oder Soundkarten sind in fast allen Computersystemen verbaut und entlasten den Hauptprozessor. Man spricht auch von heterogenen Computersystemen.

Die im Rahmen dieser Arbeit vorgestellte Hardware soll ein Beispiel für eine derartige heterogene Komponente sein. Auf einem Field Programmable Gate Array (FPGA) soll eine performante und energieeffiziente Visualisierung der sogenannten Mandelbrotmenge realisiert werden. Diese Problemstellung ist auch durch einen ordinären Prozessor lösbar, lastet diesen jedoch enorm aus und ist somit auch sehr energieineffizient.

1.2 Aufgabenstellung

Die Aufgabe dieses Projektes ist es, eine komplett in Hardware stattfindende Berechnung der Mandelbrotmenge durchzuführen und die Ergebnisse über eine VGA-Schnittstelle darzustellen

(ähnlich Bild)

Weiterhin soll die Hardware durch externe Peripherie konfigurierbar hinsichtlich der angestellten Berechnungen sein. So soll etwa aktuell abgebildete Bereich der Mandelbrotmenge oder auch die Farbgebung der Darstellung im laufenden Betrieb geändert werden können.

Hierfür wurde das FPGA-Trainer Board *Zybo Zynq-7000* zur Verfügung gestellt, welches in genauer vorgestellt wird.

1.3 Mitgelieferte Skripte

Im Git-Repository ... sind sämtliche Python-Skripts, die zur Erstellung von selbsterstellten Bildern verwendet wurden, enthalten. Des weiteren gibt die dort enthaltene Datei *readme.md* Aufschluss über nützliche Skripts, die im Rahmen dieser Arbeit verwendet wurden.

2 Technische Grundlagen

Zum besseren Verständnis des Gesamtprojektes sollen in diesem Kapitel einige technische Konzepte erläutert werden.

2.1 FPGAs

Ein *Field Programmable Gate Array* (kurz FPGA) ist ein Schaltkreis, welcher mit Hilfe von Hardwarebeschreibungssprachen konfiguriert werden kann, um beliebig komplexe logische Schaltungen zu realisieren.

Das Grundelement eines solchen Bausteines bilden die sogenannten *Lookup Tables* (kurz LUTs), welche zu einem beliebigen n-bit Input ein 1-bit Output Signal produzieren. Die zugrundeliegende Logiktablelle einer LUT ist hierbei frei programmierbar.

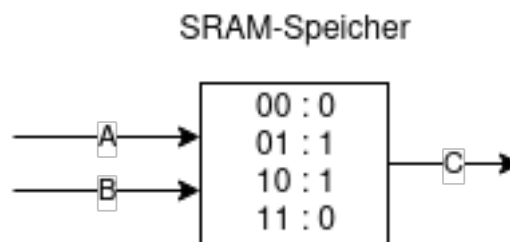


Abbildung 1: Beispielhafter Aufbau einer 2-Input LUT

Eine LUT, welche die Operation $C = A \oplus B$ implementiert ist in Abbildung 1 zu sehen. In dieser wird in einem SRAM-Speicher für jede Inputkombination ein Outputwert hinterlegt, wodurch jede 2-Bit Funktion abgebildet werden kann. In Verbindung mit einem Flipflop¹ bildet eine LUT dann einen sogenannten Logikblock (s. Abbildung 2). [1]

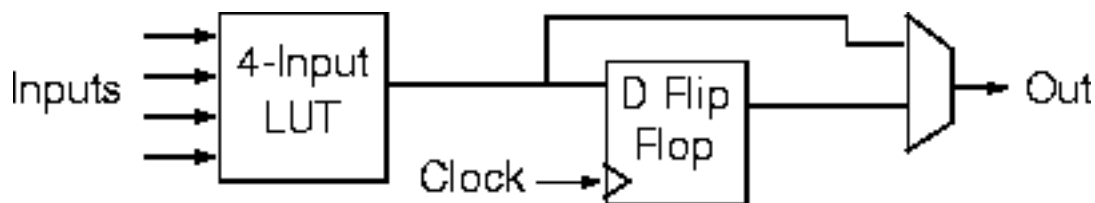


Abbildung 2: Logikblock, aus [1]

Ein FPGA verbindet nun durch ebenfalls konfigurierbare Bussysteme viele solcher Logikblöcke um komplexere Schaltkreise abzubilden.

¹Ein Flipflop ist ein Speicherelement, welches einen einzigen Bit Daten halten kann.

Weiterhin verfügen FPGA-Boards oft noch über ergänzende Hardwarekomponenten, von denen die im Falle dieses Projektes vorhandenen im Folgenden gezeigt werden sollen.

DSP Ein *Digitaler Signalprozessor* (DSP) ist ein fest integrierter Baustein, welcher durch Multiplizierer und Akkumulatoren binäre Algorithmen beschleunigt. So übernimmt dieser etwa grundlegende mathematische Operationen, was dazu führt, dass diese flächen- und energiesparender durchgeführt werden, als bei der Verwendung von LUTs [5, S. 52]. Typische Anwendungsgebiete dieser Bausteine sind Fließkommamultiplikationen, Schnelle Fourier-Transformationen oder einfache Zähler [5, S. 14]. In diesem Projekt wurden die DSPs hauptsächlich aufgrund ihrer 25x18 Bit Multiplizierer verwendet, welche kaskadiert werden können um beliebig Breite Multiplikationen durchzuführen.

Block RAM FPGAs verfügen meist über *Block RAM* (BRAM), welcher zur Speicherung binärer Daten dient. Dieser Speicher ist lese-/schreibsynchron, wodurch Inkonsistenzen beim Speicherprozess ausgeschlossen sind [6, S. 11]. Um auf diese Speicherblöcke Zugriff zu erlangen muss der von Xilinx mitgelieferte Baustein "Block RAM Generator" verwendet werden.

IO-Komponenten Zur Kommunikation mit der Außenwelt verfügen FPGAs über diverse Schnittstellen wie z.B. Knöpfe, Schalter, aber auch komplexere Anbindungen wie etwa VGA (s. hierzu Abschnitt 2.2) oder PMOD-Anschlüsse. Diese sind so angebunden, dass ihre Signale direkt in Logikschaltungen von LUTs integriert werden können.

2.2 VGA-Schnittstelle

Eine *Video Graphics Array* (VGA) Schnittstelle wird durch einen Videoübertragungsstandard, welcher erstmals von IBM in ihrer *IBM Personal System/2* Modellreihe verbaut wurde, spezifiziert [7]. Das Darstellungsverfahren verwendet einen 15-poligen Anschluss, um Videosignale in variabler Auflösung und Bildwiederholungsrate zu übertragen.

Hierbei liegen die RGB-Werte eines jeden Pixels als analoge Spannungen an und werden zu bestimmten Zeitpunkten vom Bildschirm ausgelesen. Da der VGA-Standard im Jahre 1987 aufkam, war er ursprünglich noch für Kathodenstrahlröhrenbildschirme (auch Röhrenbildschirme genannt) ausgelegt. Die Elektronenstrahlen dieser Bildschirme konnten sich nicht ohne kurze Verzögerungen über die Anzeigefläche bewegen, was bedeutet, dass der VGA-Standard dies berücksichtigt und dem Bildschirm einige μs für größere Sprünge des Strahls einräumen muss.

Die größten solchen Sprünge finden statt, wenn eine Pixelreihe übertragen wurde und der Strahl an die erste Pixelposition der nächsten Reihe bewegt werden muss oder wenn das gesamte Bild übertragen wurde und wieder zum oberen linken Pixel gesprungen werden muss. Während dieser Pausen werden keine RGB-Werte übertragen, man nennt diese zeitlich gedachten Bereiche auch "Porch"(engl. Vorbau).

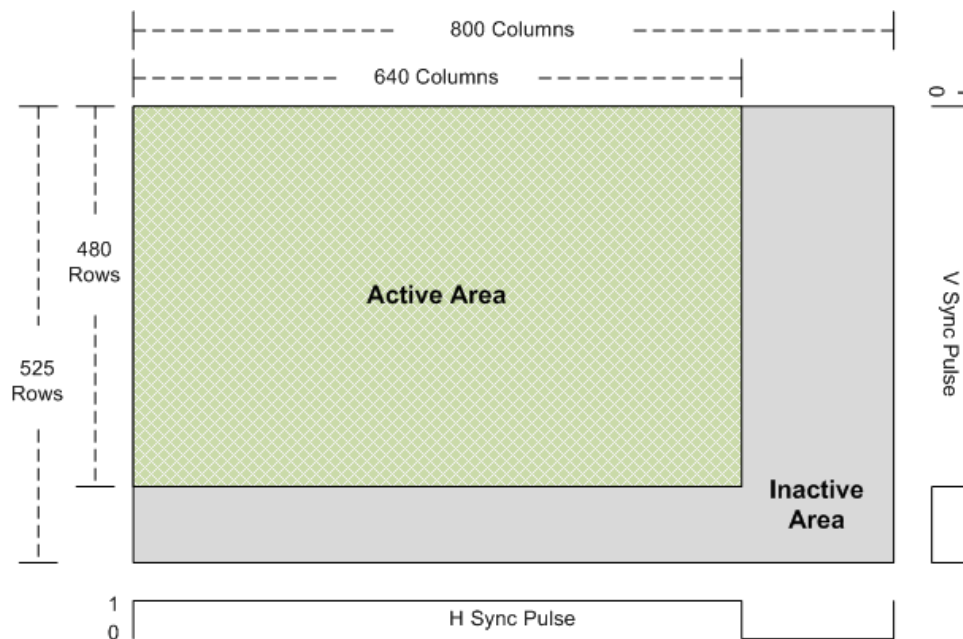


Abbildung 3: VGA-Timing/Aufbau für eine 640x480 Auflösung, aus [2]

Eine 640x480 Pixel Bild baut sich dann wie in Abbildung 3 gezeigt auf: Zuerst werden 640 Pixel RGB Daten für die erste Reihe empfangen, danach kommen 160 Pixel inaktiver Bereich (\Rightarrow Porch). Das Ende dieses Bereiches wird durch das Ansteigen des low-aktiven Signal HSYNC signalisiert. Dieser Vorgang wird nun 480 mal wiederholt, bis das gesamte Bild übertragen wurde. Draufhin wird analog zur horizontalen Synchronisation das VSYNC Signal für 45 Pixel auf auf 0 gesetzt, um den vertikalen inaktiven zu signalisieren. Die Länge des inaktiven Bereiches setzt sich aus Front-, bzw. Backporch und der Länge des Sync-Pulses zusammen. Die Bedeutungen der einzelnen Signale können fügen jedoch dem nötigen Verständnis nicht mehr hinzu, weswegen die Summe dieser Bereiche als Ganzes angesehen werden kann. Wie schon erwähnt ist 640x480 jedoch nicht die einzige Auflösung, die mit einem VGA-Anschluss realisierbar ist. Andere Auflösungen (mit anderen Wiederholungsraten) müssen vom verwendeten Bildschirm unterstützt sein, und werden durch verschieden schnelle Pixelclocks und Porches realisiert.

Die in dieser Arbeit verwendete Auflösung von 800x600 Pixeln bei 60Hz benötigt die in Tabelle 1 dargestellten Werte.

aktiver Bereich (horizontal)	Pixel	800
aktiver Bereich (vertikal)	Pixel	600
inaktiver Bereich (horizontal)	Pixel	256
inaktiver Bereich (vertikal)	Pixel	28
Pixelfrequenz	MHz	40

Tabelle 1: VGA-Werte für ein 800x600@60Hz Signal

Die in Tabelle 1 gezeigte Pixelfrequenz von 40 MHz bedeutet, dass alle 25 ns ein neuer RGB-Wert anliegen muss. Dies ist ein wichtiger Aspekt für weitere grundlegende Designentscheidungen.

3 Theoretische Grundlagen

Die dieser Arbeit zugrundelegenden mathematischen Grundlagen und Definitionen sollen in diesem Kapitel näher erläutert werden.

3.1 Fraktale

Der Begriff *Fraktal* wurde vom französischen Mathematiker Benoît Mandelbrot geprägt und leitet sich vom lateinischen Adjektiv *fractus* ab, was „in Stücke gebrochen“ oder „irregulär“ bedeutet. Allgemein ist hiermit entweder eine natürlich vorkommende Struktur mit gewissen Eigenschaften oder eine genau definierte mathematische Menge gemeint. [8, S. 16]

Für ein intuitives Verständnis des Begriffes sollen im folgenden zuerst einige natürlich vorkommende Fraktale gezeigt werden, woraufhin im nächsten Abschnitt eine formale Definition des Begriffes *Fraktal* folgen soll.

3.1.1 Natürlich vorkommende Fraktale

Fraktale besitzen oft selbstähnliche Strukturen, d.h. dass sich die Gesamtstruktur eines Objektes in kleinerem Maßstab immer wieder findet. Ein Beispiel hierfür ist ein fraktal definierter Baum wie er in Abbildung 4 dargestellt ist. Der Baum wird hierbei über ein einfaches rekursives Verfahren definiert, bei dem immer wieder von jedem Teilbaum aus mit einem festen Winkel in jeweils zwei Äste abgebogen wird.

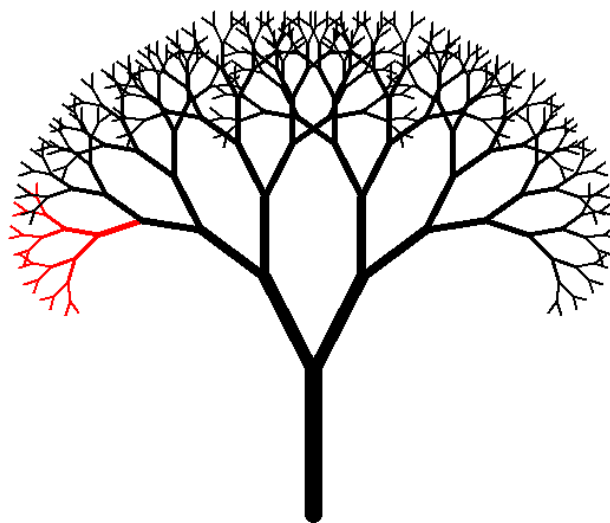


Abbildung 4: Fraktal definierter Baum, Skript: *baum.py*, abgewandelte Version von [3]

Der in Abbildung 4 rot markierte Teilbaum ist in seiner Struktur mit dem Gesamtbaum identisch und besitzt lediglich eine niedrigere rekursive Tiefe.

Der Blumenkohl (Abbildung 5) ist ein weiteres Beispiel für ein natürlich vorkommendes Fraktal. Auch bei diesem wiederholt sich die Gesamtstruktur in kleinerem Maßstab in den Ästen, er besitzt also ein gewisses Maß an Selbstähnlichkeit.



Abbildung 5: Blumenkohl, von Rainer Renz

Auch auf höheren Maßstäben, wie etwa bei Bergen oder ganzen Landschaften können Fraktale Strukturen immer wieder beobachtet werden. Die relativ einfachen Regeln, die diesen Fraktalen zu Grunde liegen machten sich bereits im Jahre 1980 Computergrafiker wie etwa Loren Carpenter zu Nutzen. Die relativ begrenzten Rechnerleistungen zwangen Animatoren zu diesem Zeitpunkt dazu, komplexe Landschaften Bild für Bild von Hand zu zeichnen. Durch Mandelbrot's Arbeit in fraktaler Geometrie inspiriert animierte Carpenter eine Landschaft für den Film *Star Trek II: The Wrath of Khan*, wobei er hierfür fraktale Verfahren verwendete. [9]

Die hierfür angestellten Berechnungen waren so simpel, dass pro Bildpunkt nur etwa 20 bis 40 Minuten Rechenaufwand betrieben werden mussten, was einen großen Fortschritt gegenüber der manuellen Animation darstellte. [10]

3.1.2 Fraktale in der Mathematik

Im Gegensatz zu den in Unterabschnitt 3.1.1 dargestellten Fraktalen, welche sich auf kleinerer Ebene nur wenige male wiederholen, sind Fraktale in der Mathematik bis zu unendlich hohem Detailgrad definiert.

Die formale Definition eines Fraktals lautet hierbei nach Mandelbrot:

Ein Fraktal ist nach Definition eine Menge, deren Hausdorff-Besicovitch Dimension echt die topologische Dimension übersteigt.

- Benoît Mandelbrot, aus [8, S. 27]

Die hier erwähnte Hausdorff-Besicovitch Dimension ist ein Maß, welches einem beliebigen geometrischen Raum zugeordnet werden kann, wobei die Dimension hier keine natürliche Zahl sein muss. In vereinfachter Form ermittelt sich die Hausdorff-Dimension folgendermaßen:

Man betrachte die Anzahl Kugeln (oder Kreisen) N mit Radius R , die nötig sind, um eine Punktmenge vollständig abzudecken. Geht nun R gegen 0 werden immer mehr Kugeln benötigt, um die Punktmenge abzudecken. Beobachtet wird nun in welcher Relation N zu R wächst, mit Hilfe der Formel:

$$D = - \lim_{R \rightarrow 0} \frac{\log N}{\log R}$$

wobei D die Hausdorff-Dimension ist. Betrachtet man etwa eine Linie der Länge 1, kann diese zunächst mit $N = 1$ Kreisen des Radius $R = 1$ abdecken. Halbiert man nun R sind doppelt so viele Kreise nötig um die Linie abzudecken. Allgemein lässt sich sagen, dass in diesem Fall N umgekehrt proportional zu R wächst. Drückt man nun N in Abhängigkeit von R aus erhält man für die Dimension einer Kurve:

$$D = - \lim_{R \rightarrow 0} \frac{\log \frac{1}{R}}{\log R} = 1$$

Analog werden etwa bei einem Rechteck $1/R^2$ Kugeln zur Abdeckung benötigt, wenn R gegen 0 läuft. Die Dimension eines Rechtecks ist daher:

$$D = - \lim_{R \rightarrow 0} \frac{\log \frac{1}{R^2}}{\log R} = 2$$

Bei den hier gezeigten Formen ist die Hausdorff Dimension nicht höher als deren topologische Dimension ¹ Nimmt man jedoch nun ein Fraktal, wie etwa das Sierpinski-Dreieck herbei, hat dieses oft einen gebrochenen Dimensionwert, in diesem Fall:

$$\frac{\log 3}{\log 2} \approx 1.585$$

Zusätzlich zu der bereits gezeigten Definition gilt: Jede Menge, die einen nichtganzzahligen Dimensionwert hat, ist ein Fraktal [8, S. 27].

¹s. hierzu <https://www.math.tu-cottbus.de/~froehner/sonstiges/skripte/node9.html>

3.2 Die Mandelbrotmenge

Nachdem nun die formale Definition eines Fraktales bekannt ist, soll im folgenden Mandelbrotmenge erläutert werden.

Es gilt: Teil der Menge sind alle komplexen Zahlen c , für die die Iteration

$$z_0 = 0$$

$$z_{n+1} = z_n^2 + c$$

nicht divergiert. Im folgenden sind Iterationen für einige c gezeigt:

$$c_1 = 1; z = 2 \rightarrow 5 \rightarrow 26 \rightarrow 677$$

$$c_2 = -1; z = 0 \rightarrow -1 \rightarrow 0 \rightarrow -1$$

$$c_3 = 1 + 1i; z = 1 + 3i \rightarrow -7 + 7i \rightarrow 1 - 97i \rightarrow -9407 - 193i$$

Für c_1 und c_3 sieht man schnell, dass der Betrag dieser Zahlen divergiert, während es bei c_2 klar ist, dass z nur zwischen 0 und -1 wechselt. Also ist c_2 Teil der Menge, während es c_1 und c_3 nicht sind. Jedoch ist die Divergenz nicht für alle Werte so einfach festzustellen. Man betrachte folgendes c :

$$c_4 = -0.55 + 0.46i; z = 0.45 - 0.04i \rightarrow -0.34 + 0.05i \rightarrow -0.68 + 0.11i \rightarrow -0.09 + 0.29i$$

Hier ist es zunächst unklar ob eine Divergenz stattfinden wird, es müssten erst viele Iterationsschritte angestellt werden, um eine Aussage treffen zu können. Nach 200 Schritten ergibt sich die in Abbildung 6 gezeigte Struktur, bei der man erkennen kann, dass z gegen einen bestimmten Punkt konvergieren zu scheint.

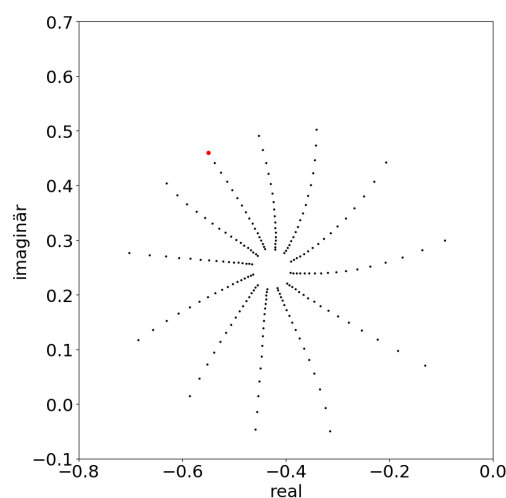


Abbildung 6: 200 Iterationsschritte für $c = -0.55 + 0.46i$ (rot markierter Punkt)

Allgemein gilt, dass ein c nicht zur Mandelbrotmenge gehört, sobald für eine Zahl in einer Iteration gilt ²:

$$|z_n| > 2$$

Zählt man nun die Iterationen, die benötigt wurden, um festzustellen, ob eine Zahl divergiert (oder eben konvergiert) und ordnet Bereichen von Iterationszahlen verschiedene Graustufen zu, erhält man ein Bild wie es in Abbildung 7 zu sehen ist.

Als Iteration wird fortan der in Zeile 4-7 dargestellte Pseudocode aus Code 1 bezeichnet.

In der Praxis ist es üblich, sich eine maximale Iterationszahl zu setzen, nach der einfach angenommen wird, dass ein gegebenes c zur Mandelbrotmenge gehört.

Eine Funktion zur Berechnung der Iterationszahl eines Punktes c der Mandelbrotmenge wird in Code 1 gezeigt.

Code 1: Algorithmus zur Berechnung der Iterationszahl eines c

```
1 funktion mandelbrot(c) :  
2   z = 0  
3   für iteration in 0 bis maxiter:  
4       wenn betrag(z) > 2:  
5           return iteration  
6       sonst:  
7           z = z*z + c  
8   return iteration //maximale Iteration erreicht
```

Diese Funktion würde nun für jeden Pixel (bzw. einem dazugehörigen c Wert) aufgerufen werden, um ein Bild wie in Abbildung 7 darzustellen.

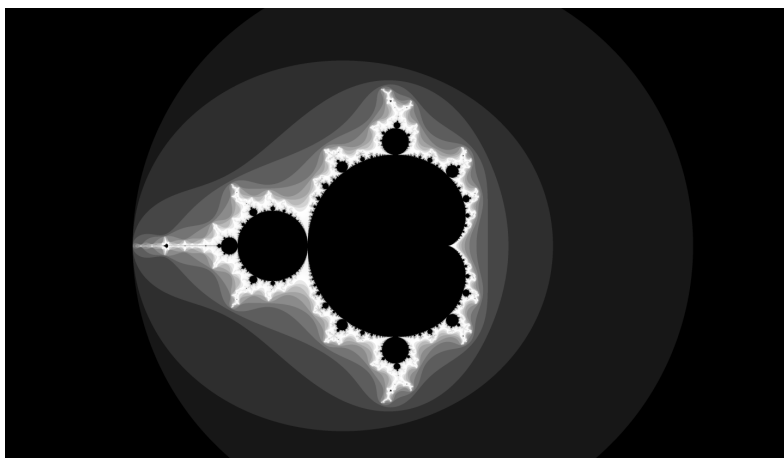


Abbildung 7: Mandelbrotmenge in Graustufe

²vgl. hierzu: [11]

Die eigentliche Faszination der Mandelbrotmenge liegt in ihrem Detailgrad, welcher beim Vergrößern des Randes der Menge (in Abbildung 7 weiß dargestellt), sichtbar wird. Es sind diverse (oft selbstähnliche) Strukturen zu sehen, wobei kein eindeutig visuelles Muster vorhanden zu sein scheint. Da mit einem Vergrößern des Detailgrades präzisere Berechnungen und höhere maximale Iterationszahlen einhergehen, liegt ein enormer Rechenaufwand vor. Einige besonders sehenswerte Teile der Mandelbrotmenge sind im dargestellt.

4 Umsetzung

In diesem Kapitel soll gezeigt werden, wie die in Abschnitt 1.2 formulierte Aufgabenstellung umgesetzt wurde. Hierfür wird zuerst in Abschnitt 4.1 das für die zugrundeliegenden Berechnungen verwendete Zahlenformat vorgestellt. Danach wird das in Abschnitt 4.2 gezeigte Gesamtsystem in den folgenden Abschnitten in seinen Einzelteilen erklärt.

4.1 Gewähltes Zahlenformat

Die Beobachtung eines komplexzahligen Punktes hinsichtlich seiner Zugehörigkeit zur Mandelbrotmenge setzt ein sehr präzises Datenformat voraus, damit sich auch bei hohen Vergrößerungsfaktoren Rechenungenauigkeiten nicht visuell bemerkbar machen.

Da die Mandelbrotmenge vollständig in dem in Tabelle 2 gezeigten Bereich liegt, muss das verwendete Datenformat nur einen kleinen Teil seiner Bits für den ganzzahligen Teil der verschiedenen Berechnungen verwenden.

Minimum Realteil	-2
Maximum Realteil	0.5
Minimum Imaginärteil	-1.25i
Maximum Imaginärteil	1.25i

Tabelle 2: Bereich der Mandelbrotmenge

Der große Teil des Bitvektors sollte also für die Darstellung der Nachkommastellen verwendet werden. Die Sprache VHDL liefert im Paket *IEEE_numeric_std* den Typ *signed*, welcher dazu dient vorzeichenbehaftete Ganzzahlen in einem beliebig breiten Bitvektor darzustellen.¹

Da die Rechenoperationen auf binärer Ebene für eben solche Festkommazahlen und Ganzzahlen (jeweils mit Vorzeichen) identisch sind, ist der *signed*-Typ gut zur Verwendung geeignet. Intern rechnet das FPGA also mit großen Ganzzahlen, welche jedoch eigentlich Festkommazahlen repräsentieren.

Hierbei gilt es, die Breite der *signed* Zahlen so festzulegen, dass zum einen eine maximale Präzision erreicht wird und zum Anderen die mit ihnen durchgeführten Multiplikationen von den eingebauten DSPs durchgeführt werden können (s. Abschnitt 2.1).

Durch Testen ergab sich, dass Multiplikationen mit mehr als 40 Bit durch LUTs umgesetzt

¹s. hierzu die Definition in: [12]

werden (statt durch DSPs), weswegen eben diese 40 Bit sich als ideale Breite herausstellten.

Da der Betrag einer Komplexen Zahl am Anfang einer Iteration nicht größer als 2 sein kann (dies wird stets geprüft), ist das größte Multiplikationsergebnis $2 * 2 = 4$, weswegen $\log_2 4 + 1 = 3$ Bits für den Vorkommateil der Zahl benötigt werden.

Das höchstwertige Bit ist für das Vorzeichen reserviert, weswegen 36 Bits für den Nachommateil bleiben. Hiermit ist die Genauigkeit im Dezimalsystem ungefähr 11 Stellen:

$$36 * \log_{10} 2 \approx 10.84$$

Die betragsmäßig kleinste darstellbare Zahl ist damit:

$$0000,0000...001_2 = 2^{-36} \approx 1.455_{10} * 10^{-11}$$

Die betragsmäßig größte darstellbare Zahl ist also:

$$0111,1111...111_2 = 8 - 2^{-36} \approx 7.99999999998$$

Die Zahl wird im Einerkomplement dargestellt, was bedeutet, dass bei negativen Zahlen alle Bits invertiert werden. So ist die kleinste darstellbare Zahl:

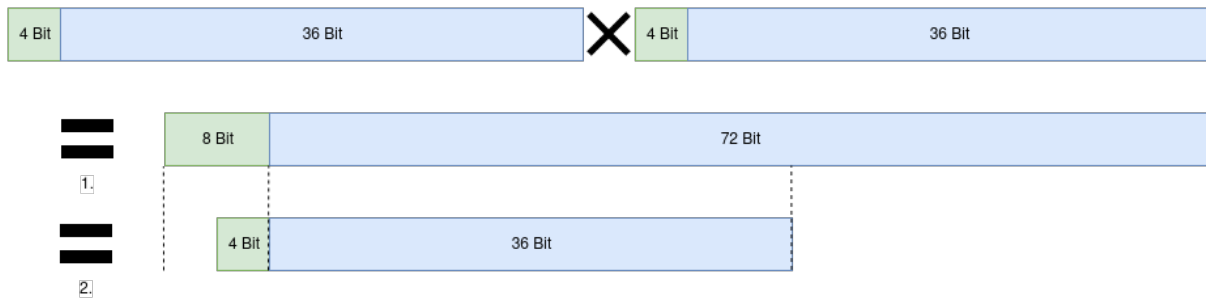
$$1000,0000...000_2 = -8 + 2^{-36} \approx -7.99999999998$$

Selbst bei einer 134217728-fachen Vergrößerung (27 Zooms um Faktor 2) ergeben sich mit dieser Präzision noch keine visuell wahrnehmbaren Rechenfehler.

Um noch tiefere Zooms zu realisieren müsste entweder eine variable Breite der einzelnen Datenworte realisiert werden, oder der grundlegende Algorithmus zur Berechnung der Mandelbrotmenge geändert werden. Hierfür würde sich der relativ neue Algorithmus von K.I. Martin (s. [13]) eignen, welcher größtenteils unabhängig von aktueller Zoomtiefe und ohne die Verwendung von beliebig präzisen Datentypen auskommt [13].

Da die Umsetzung dieses Algorithmus jedoch um einiges komplexer als die in dieser Arbeit verwendete Methode ist und die aktuelle Präzision zufriedenstellend ist, wird von dessen Implementierung abgesehen, jedoch ist die Bitbreite der einzelnen Datenworte in der Datei *constants.vhd* konfigurierbar.

Bitbreite bei Multiplikationen Da bei der Multiplikation zweier 40-Bit-Zahlen das Ergebnis 80 Bit lang ist, muss dieses wieder in ein 40-Bit Format zurückgeführt werden. Dies geschieht mittels einer dafür entwickelten VHDL-Prozedur *fixlen*, welche in der Datei zu finden ist.


Abbildung 8: Funktionsweise der Funktion *fixlen*

Die Funktionsweise dieser Prozedur ist in Abbildung 8 dargestellt. Die aus der Multiplikation in Schritt 1. entstehende 80 Bit zahl besteht aus 8 Bits Vorkommateil und 72 Bit Nachkommateil. Um diese Zahl nun wieder in das normale Format zurückzuführen werden die ersten 4 Bit des Vorkommateils und die letzten 36 Bits des Nachkommateils abgeschnitten (Schritt 2.). Hierbei verliert das Ergebnis der Multiplikation an Präzision, was jedoch unvermeidbar ist, da sich die Länge der Datenworte sonst bei jeder Multiplikation verdoppeln würde. Umgesetzt wird dieses Verfahren in einer VHDL-Funktion *fixlen* (Code 2), welche durch die Konstante *N_BITS* parametrisierbar ist.

Code 2: VHDL-Funktion *fixlen*

```

1  --N_BITS sei hier 16
2  function fixlen (input : signed(N_BITS*2-1 downto 0)) return signed is
3      variable ret : signed(N_BITS-1 downto 0);
4  begin
5      ret := input(N_BITS*2-1-4 downto N_BITS*2-4-N_BITS);
6      return ret;
7  end fixlen;
8
9  a <= "1111000011110000";
10 b <= "0000111100001111";
11
12 a <= a*b; --Fehler!
13 a <= fixlen(a*b);

```

Da das Ergebnis der 16-Bit Multiplikation in Zeile 10 32 Bit breit ist, kann dieses nicht dem 16-Bit Wert *a* zugewiesen werden. Stattdessen muss dieses Ergebnis wie in Zeile 11 gezeigt wieder auf eine Breite von 16 Bit gekürzt werden.

Umwandlung zwischen Zahlensystemen Die Verfahren zur Umwandlung zwischen dem im Rahmen dieser Arbeit vorgestellten Zahlensystem und dem Dezimalsystem soll im folgenden aufgezeigt werden.

4.2 Systemüberblick

In diesem Abschnitt soll ein Überblick über das entwickelte Gesamtsystem gegeben werden, während in den folgenden Abschnitten die einzelnen Komponenten erläutert werden.

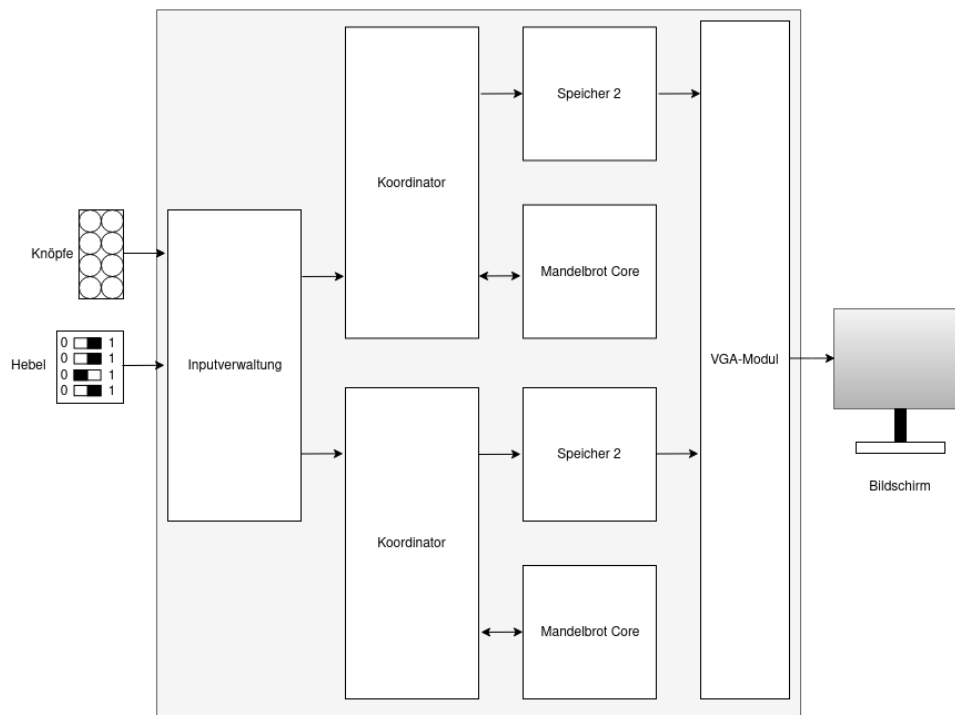


Abbildung 9: Gesamtsystem mit 2 Kernen

In Abbildung 9 ist der konzeptuelle 2-kernige Aufbau des Gesamtsystems dargestellt.

Die Darstellung erfolgt über eine VGA-Schnittstelle, welche durch das VGA-Modul bedient wird. Dieses liest stetig Pixelwerte aus den vorhandenen Speicherblöcken aus und generiert aus diesen ein gültiges VGA-Signal.

Der Speicher wird wiederum von den sogenannten *Koordinatoren* gefüllt. Jeder Koordinator verwaltet die Berechnungen eines *Mandelbrot-Cores* und schreibt dessen Rechenergebnisse in seinen eigenen Speicher.

Die Koordinatoren erhalten Informationen wie z.B. den aktuell darzustellenden Bereich von der Inputverwaltung, welche die externen Peripheriegeräte wie Hebel und Knöpfe verwaltet.

4.3 Komponentenbeschreibung

Die genaue Funktion und Umsetzung aller genannten Module wird in den folgenden Abschnitten gezeigt. Alle Signale der vorgestellten Bauteile sind high-aktiv, wenn dies nicht der Fall ist, wird es explizit deklariert. Zudem sind alle Inputs/Outputs **fettgedruckt**, um sie von Variablen in Gleichungen unterscheiden zu können.

4.3.1 Mandelbrot-Core

Der zentralste Baustein des Chips ist der Mandelbrot-Core (s. Abbildung 10), welcher für ein gegebenes c einen Iterationswert errechnet.

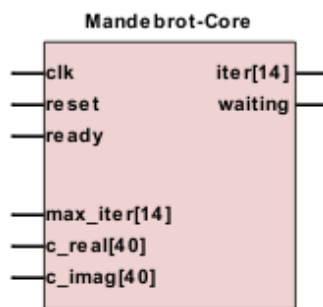


Abbildung 10: Mandelbrot-Core, schematische Darstellung

Als Input erhält ein Mandelbrot-Core zunächst die üblichen Clock- und Resetsignale (**clk**, **reset**). Des weiteren liegt der zu untersuchende Punkt c an den Eingängen **c_real** (Realteil) und **c_imag** (Imaginärteil) an. Die maximale Zahl von Iterationen, nach der angenommen wird, dass ein Punkt nicht zur Mandelbrotmenge gehört wird über das Eingangssignal **max_iter** übermittelt. Sobald über das Signal **ready** angezeigt wird, dass **c_real**, **c_imag** und **max_iter** stabil anliegen beginnt der Kern den Iterationsvorgang.

Hierbei wird jeden Zyklus von **clk** eine Iteration abgeschlossen.

Über den Ausgang **waiting** wird angezeigt, dass ein endgültiges Ergebnis am Ausgang **iter** anliegt und der Kern bereit für neue Inputs ist. Das Ergebnis kann Werte zwischen (jeweils inklusive) 0 (falls $|c| > 2$) und **maxiter** annehmen.

Abbildung 11 zeigt ein Impulsdigramm der In- und Outputs eines Mandelbrot-Cores. Zu bestimmten ist in diesem Fall die Iterationszahl für den Punkt $-0,7487 - 0,7487i$, wobei die maximale Iterationszahl 100 sein soll. Diese Werte werden an die Inputs **c_real**, **c_imag** und

max_iter angeleg. Sobald diese stabil vorliegen, wird dies durch einen Puls des Signales **ready** signalisiert, wodurch der Core seine Berechnungen startet. Nach einigen Zyklen des Clocksignals liegt dann die Iterationszahl (in diesem Fall 3) am Ausgang **iter** an.

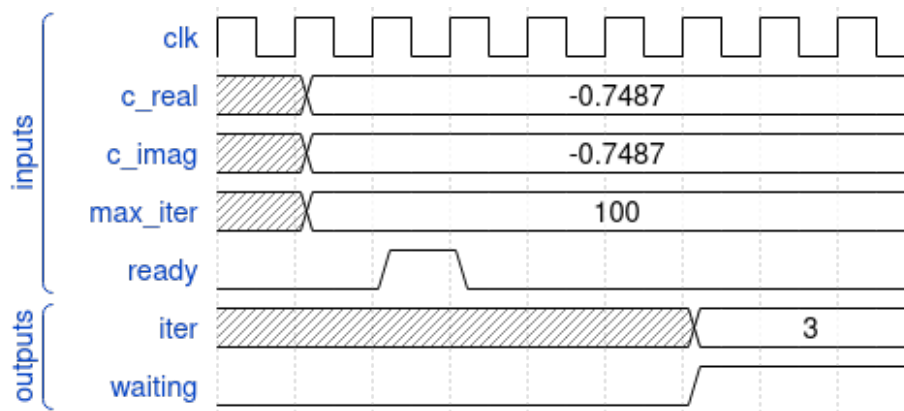


Abbildung 11: Funktionsweise Mandelbrot-Core, Impulsdiagramm

Vor diesem Zeitpunkt ist der Zustand von **iter** undefiniert, erst nach der steigenden Taktflanke von **waiting** liegt das richtige Ergebnis an. Ab diesem Zeitpunkt ist der Core auch wieder bereit, neue Werte von c zu verarbeiten.

Da der Mandelbrot-Core über einen den Eingang **max_iter** verfügt und diesen auch bei jeder Berechnung neu ausliest, kann die maximale Iterationszahl im laufenden Betrieb jederzeit geändert werden,² was nützlich ist um bei steigender Vergrößerungsstufe den Detailgrad der Berechnung erhöhen zu können.

Die Mandelbrot-Cores werden mit einer eigenen Clock, die unabhängig von z.B. der VGA-Clock läuft, getaktet. Die maximal mögliche Taktrate ist hierbei abhängig von der Anzahl der Rechenschritte, die in jedem Iterationsschritt durchgeführt werden müssen. Eine Iteration besteht aus folgenden Bestandteilen:

- Überprüfen, ob maximale Iterationszahl bereits erreicht ist, wenn dem so ist, ist die Berechnung abgeschlossen
- Wenn nicht, berechne neuen Wert für z , falls dieser größer als 2 ist, ist die Berechnung abgeschlossen
- Zeige den aktuellen Zustand der Berechnung über die Signale **waiting** und **iter** an

Die Abfrage (A) der aktuellen Iterationszahl und der Vergleich mit der maximal zulässigen ist hierbei sehr trivial und nimmt deshalb auch keine nennenswerte Zeit in Anspruch, muss also

²Neue Werte für die Signale **c_real**, **c_imag** und **max_iter** werden erst beim nächsten **ready** Puls übernommen

auch nicht weiter optimiert werden.

Der komplexzahlige Iterationsteil (B) besteht jedoch aus mehreren Multiplikationen, Additionen und Vergleichen, weswegen hier ein großer Teil der pro Iteration entstehenden Verzögerung entsteht. Es gilt, diesen Bestandteil (B) weitestgehend zu optimieren, was durch einfache algebraische Umformungen geschehen kann.

In jedem Schritt muss überprüft werden, ob z kleiner 2 ist:

$$\text{abs}(z) \leq 2$$

Der Betrag (abs) einer komplexen Zahl mit Realteil z_r und Imaginärteil z_i ist als ihr Abstand vom Ursprung definiert:

$$\text{abs}(z) = \sqrt{z_r^2 + z_i^2}$$

Also muss geprüft werden, ob gilt:

$$\sqrt{z_r^2 + z_i^2} \leq 2$$

Das Ziehen der Wurzel kann in diesem Fall durch das Quadrieren beider Seiten umgangen werden:

$$z_r^2 + z_i^2 \leq 4 \quad (4.1)$$

Falls dies wahr ist, muss der Wert von z aktualisiert werden und die Iterationszahl kann um eins erhöht werden:

$$z_{n+1} = z_n^2 + c$$

$$\text{iter} = \text{iter} + 1$$

Das Quadrat einer komplexen Zahl mit Realteil a und Imaginärteil b wird folgendermaßen berechnet:

$$(a + bi)^2 = a^2 + 2abi + b^2 * i^2 = a^2 + 2abi - b^2$$

Im Fall von z gilt also:

$$z^2 + c = (z_r + z_i i)^2 = z_r^2 - z_i^2 + 2z_r z_i i + c$$

Da die Bestandteile der komplexen Zahlen getrennt gespeichert sind, müssen diese isoliert betrachtet werden (das = Zeichen ist in diesem Kontext als Zuweisungsoperator zu verstehen):

$$z_r = z_r^2 - z_i^2 + c_r \quad (4.2)$$

$$z_i = 2z_r z_i i + c_i = z_r z_i i + z_r z_i i + c_i \quad (4.3)$$

Die Quadrate von z_r und z_i kommen sowohl in Gleichung 4.1 als auch in Gleichung 4.2 vor, müssen aber nur ein mal berechnet werden, was zwei Multiplikationen spart.

Auch das Ersetzen der Multiplikation mit zwei in Gleichung 4.3 durch eine Addition erspart eine Multiplikation. Der nicht optimierte Iterationsschritt (B) enthält eine Wurzel, sechs Multiplikationen und vier Additionen. Durch Wiederverwenden bereits berechneter Werte und algebraische Optimierung sind nun lediglich 3 Multiplikationen und fünf Additionen nötig.

Das optimierte Iterationsverfahren läuft bei einer maximalen Clockfrequenz von 40MHz stabil, was bedeutet dass eine einzelne Iteration maximal 25ns in Anspruch nimmt.

4.3.2 Mandelbrot-Koordinator

Der in Abbildung 12 gezeigte Mandelbrot-Koordinator verwaltet einen Mandelbrot-Core und schreibt dessen Ergebnisse in den in Unterabschnitt 4.3.3 beschriebenen RAM.

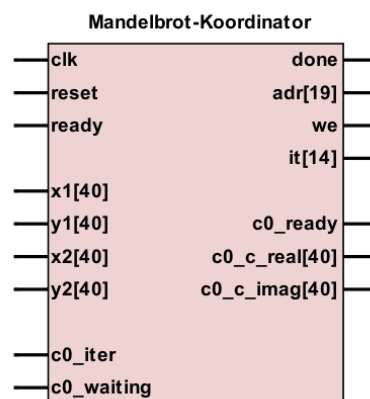


Abbildung 12: Mandelbrot-Koordinator, schematische Darstellung

Berechnet wird die Mandelbrotmenge in dem durch die Punkte (**x1/y1**) und (**x2/y2**) aufgespannten Rechteck (s. Abbildung 13). Liegen diese Eingangssignale stabil an, kann durch das Aktivieren des Eingangs **ready** der Rechengang gestartet werden.

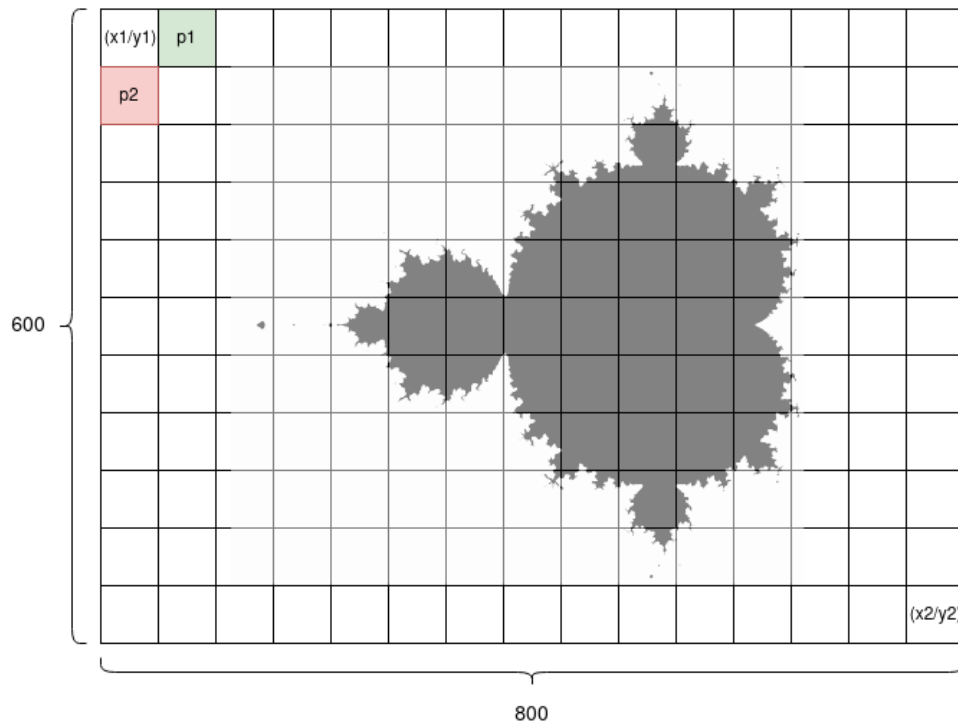


Abbildung 13: Mandelbrot-Koordinator, schematische Darstellung

Der obere linke Pixel repräsentiert den Punkt $(\mathbf{x1/y1})$. Um zu einer gegebenen Pixelkoordinate (p_x/p_y) den zugehörigen komplexzahligen Punkt (c_r/c_i) zu finden wird folgende Formel verwendet:

$$c_r = \mathbf{x1} + step_x * p_x$$

$$c_i = \mathbf{y1} + step_y * p_y$$

mit:

$$step_x = \frac{x2 - x1}{800}$$

$$step_y = \frac{y2 - y1}{600}$$

Damit das dargestellte Bild nicht gestaucht ist, sollte das Seitenverhältnis von 800x600 Pixeln eingehalten werden:

$$\frac{x2 - x1}{y2 - y1} = \frac{4}{3}$$

Wenn die Mandelbrotmenge beispielsweise für das Rechteck mit den Eckpunkten $(-2/-2)$ und $(2/1)$ dargestellt werden soll gilt:

$$step_x = \frac{2 - (-2)}{800} = 0,005$$

$$step_y = \frac{1 - (-2)}{600} = 0,005$$

Dem Pixel (250/100) würde folgender Punkt $c : -0,75 - 1,5i$ zugeordnet werden:

$$c_r = -2 + 0,005 * 250 = -0,75$$

$$c_i = -2 + 0,005 * 100 = -1,5$$

Diese Werte werden an die zugehörigen Outputsignale **c0_c_real** und **c0_c_imag** angelegt und anschließen durch Aktivieren des Ausgangs **c0_ready** bestätigt.

Im Anschluss wartet der Mandelbrot-Koordinator bis das Eingangssignal **c0_waiting** aktiv wird und schreibt den an **c0_iter** anliegenden Wert in den RAM.

Hierfür werden die Signale **adr,we** und **it** verwendet, wobei deren genaue Funktionsweise im Kapitel näher erläutert wird. Wenn jeder Pixel abgearbeitet wurde, wird dies durch das Signal **done** signalisiert und der Koordinator ist bereit, neue Eingangssignale aufzunehmen.

4.3.3 Speicher

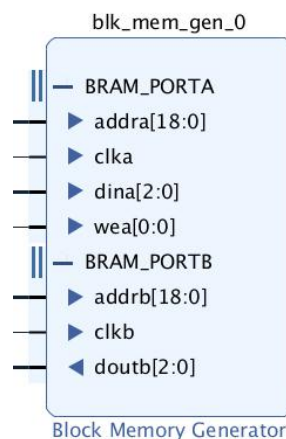


Abbildung 14: Block Memory Generator, IP-Bauteil

Der Block Memory Generator ist ein mit Vivado mitgeliefertes IP-Bauteil, welches ermöglicht den auf FPGA-Boards verfügbaren Block-RAM zu verwenden.

Verwendet wird der Block-RAM im Rahmen dieses Projekts, um die Iterationswerte, die jedem Pixel zugeordnet werden (s. Unterabschnitt 4.3.1) zu speichern.

Der Block-RAM des Zybo-Board setzt sich aus 60 32768-Bit RAMB18E1-Blöcken zusammen, verfügt also über 1966080 Bit (240 KB) Speicher [6, S. 14] [14]. Da in diesem Speicher ein 800x600 Pixel Bild gespeichert (und ausgelesen) werden soll, ergeben sich 4 Bit Speicher pro Pixel:

$$\left\lfloor \frac{1966080 \text{ Bit}}{800 * 600 \text{ Pixel}} \right\rfloor = 4 \frac{\text{Bit}}{\text{Pixel}}$$

Ein Datenwort in diesem Speicher kann also 16 (2^4) Zustände annehmen, wodurch 16 verschiedene Farbwerte gespeichert werden können.

Um alle Datenworte adressieren zu können wird ein 19-Bit Adresswort benötigt:

$$\lceil \log_2(800 * 600) \rceil = 19$$

Da die Mandelbrot-Koordinatoren einen 14-Bit Iterationswert liefern, muss dieser zur Einspeicherung auf eine Breite von 4 Bit gemappt werden. Diese Problemstellung wird in Unterabschnitt 4.3.4 behandelt.

Im Betrieb wird der Speicher permanent vom VGA-Modul ausgelesen, während die Mandelbrot-Koordinatoren neue Werte im Speicher ablegen. Da die Synchronisation dieser beiden Vorgänge enorm aufwändig wäre und die Anzahl der Lese-/Schreibzugriffe pro Sekunde eingeschränkt würde, muss der Speicher im sog. *Dual-Port*-Modus betrieben werden.

Ein Speicher im Dual-Port-Modus verfügt über zwei Ports, welche völlig unabhängig voneinander agieren (**BRAM_PORTA**, **BRAM_PORTB**; fortan **Port A** bzw. **Port B** genannt). Jeder dieser Ports verfügt über einen eigenen Takt (**clka**, **clkb**) und einen eigenen Adressbus. **Port A** wird rein zum Schreibzugriff verwendet, während **Port B** rein zum Lesezugriff konfiguriert ist. Dies äußert sich darin, dass **Port A** einen *Write Enable*-Eingang (Schreibzugriffsaktivierung) **wea** hat, jedoch keinen Datenausgang. **Port B** verfügt wiederum über einen Datenausgang **doutb**, jedoch über

4.3.4 Lookup Tables

Lookup Tables (kurz LUTs) sind Bausteine, welche für einen Index i einen bestimmten Tabelleneintrag als Output liefern können. Die Tabelleneinträge sind hierbei vordefiniert, was einige Berechnungen, die sonst in Hardware durchgeführt werden müssten, erspart.

Im folgenden werden die im Rahmen dieses Projekts verwendeten LUTs vorgestellt.

Da in dem in Unterabschnitt 4.3.3 beschriebenen Speicher nur 4-Bit Datenworte gespeichert werden können, ist es nicht möglich, für jeden einzelnen Pixel RGB-Werte zu hinterlegen, da diese auf dem Zybo-Board 16 Bit breit sind [14]. Es muss also eine Farbtabelle existieren, welche von 4-Bit Iterations-Werten auf 16-Bit RGB-Werte übersetzt (**LUT1**).

Da der Mandelbrot-Koordinator 14-Bit Iterationswerte liefert, der Speicher jedoch nur 4-Bit Datenworte speichern kann, muss auch hier eine Übersetzung stattfinden (**LUT2**).

Für die Umsetzung von **LUT2** wird folgende Formel verwendet:

$$I_F = \left\lceil \frac{i}{i_{max}} * 15 \right\rceil$$

Hierbei sind i und i_{max} 14-Bit Zahlen, wobei I_F eine 4 Bit Zahl ist, wodurch I_F in den Speicher geschrieben werden kann.

Beim Auslesen wird wiederum **LUT1** verwendet, um die per VGA dargestellten RGB-Werte zu erhalten. Die verwendeten RGB-Werte sind in einer frei konfigurierbaren Farbtabelle hinterlegt. Sei die maximale Iterationszahl $i_{max} = 100$. Zur Ermittlung der Farbe eines Pixels muss zunächst der Farbindex I_F errechnet werden. Wenn etwa 33 Iterationen benötigt wurden sieht ergibt sich:

$$I_F = \lceil \frac{33}{100} * 15 \rceil = \lceil 4,95 \rceil = 5$$

Nun kann der zugehörige RGB-Wert aus einer Farbtabelle (z.B. Tabelle 3) entnommen werden ³:

Index I_F	Farbe(RGB)	Index I_F	Farbe(RGB)
0	#FF0000	8	#778800
1	#EE1100	9	#669900
2	#DD2200	10	#55AA00
3	#CC3300	11	#44BB00
4	#BB4400	12	#33CC00
5	#AA5500	13	#22DD00
6	#996600	14	#11EE00
7	#887700	15	#00FF00

Tabelle 3: Farbtabelle, Übergang von Rot nach Grün

4.4 Peripherie und funktionale Beschreibung

³Die hier dargestellten RGB-Werte sind 24 Bit breit, da dies die übliche RGB-Notation ist. Die RGB-Werte auf dem Zybo-Board sind nur 16 Bit breit.

5 Umsetzung des Systems

6 Ausblick & Fazit

Literaturverzeichnis

- [1] FPGA Architecture for the Challenge. http://www.eecg.toronto.edu/~vaughn/challenge/fpga_arch.html. Zugriff am: 12.06.2019.
- [2] Go Board - VGA Introduction. <https://www.nandland.com/goboard/vga-introduction-test-patterns.html>. Zugriff am: 12.06.2019.
- [3] Stackoverflow Antwort: Drawing a Fractal Tree in Python, User sheldonzy. <https://stackoverflow.com/a/46754459>. Zugriff am: 12.06.2019.
- [4] Gordon E Moore et al. Cramming more components onto integrated circuits, 1965.
- [5] Xilinx, Inc. *7 Series DSP48E1 Slice*, März 2018. v1.10.
- [6] Xilinx, Inc. *7 Series FPGAs Memory Resources*, Februar 2019. v1.13.
- [7] Chronology of IBM Personal Computers. <https://web.archive.org/web/20150221071923/http://pctimeline.info/ibmpc/ibm1987.htm>. Zugriff am: 12.06.2019.
- [8] Benoît Mandelbrot. *Die fraktale Geometrie der Natur*. Springer-Verlag, 2013.
- [9] The First Completely Computer-Generated (CGI) Cinematic Image Sequence in a Feature Film. <http://www.historyofinformation.com/detail.php?entryid=3584>. Zugriff am: 12.06.2019.
- [10] Vol Libre: The First Fractal CGI Movie. <http://www.historyofinformation.com/detail.php?entryid=3690>. Zugriff am: 12.06.2019.
- [11] user147263. Mandelbrot sets and radius of convergence. Mathematics Stack Exchange. URL:<https://math.stackexchange.com/q/890448> (version: 2014-12-19).
- [12] NUMERIC_STD arithmetic package for synthesis. https://www.csee.umbc.edu/portal/help/VHDL/packages/numeric_std.vhd. Zugriff am: 12.06.2019.
- [13] K.I. Martin. *Superfractalthing maths*, 2013.
- [14] Zybo Reference Manual. <https://reference.digilentinc.com/reference/programmable-logic/zybo/reference-manual>. Zugriff am: 12.06.2019.