

ニューラルネットワークを使用した

HD から 4K2K への解像度変換ボケ復元回路の実装

小野 雅晃^{a)}

^{a)} 筑波大学システム情報工学等技術室

〒305-8573 茨城県つくば市天王台 1-1-1

概要

ニューラルネットワークを使用した画像処理研究がシステム情報工学系和田研究室で行われている。画像処理研究のうちの解像度変換ボケ復元システムについて報告する。本システムは HD(High Definition video)解像度画像の HDMI(High Definition Multimedia Interface)信号を入力とし、4K2K 解像度に解像度変換をすると同時にボケ復元を行う。4K2K 画像は 4 つの HD 解像度画像に分割し、4 つのディスプレイに出力されて、4K2K 解像度の画像として表示される。解像度変換とボケ復元は、ハードウェア化したニューラルネットワークを使用する。

キーワード：ニューラルネットワーク、FPGA、4K2K

1. はじめに

筑波大学システム情報系の和田研究室では、ニューラルネットワークを画像処理などに応用している^[1]。その一環として、HD 解像度画像の 4K2K 解像度への解像度変換とボケ復元を同時に実行するニューラルネットワークのハードウェア化を行い、表示システムを構築した。このシステムを解像度変換ボケ復元システムとする。

解像度変換ボケ復元システムに入力されるのは HD 解像度の HDMI 信号である。HD 解像度の HDMI 信号は Red、Green、Blue の 3 つのハードウェア化したニューラルネットワークに入力され、1 ピクセルが 4 ピクセルに拡大される。拡大されたピクセルは集められ、フレーム・バッファの DDR3 SDRAM に書き込まれる。ピクセルは 4 つの独立した画像出力回路により、4 つの HD 解像度画像と

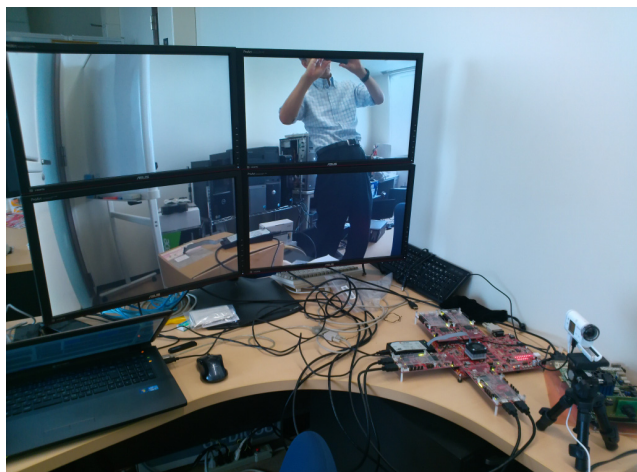


図 1. 解像度変換ボケ復元システム全体の写真

して、4 台のディスプレイに表示される。解像度変換ボケ復元システムの写真を図 1 に示す。

図 1 の右下の白い HD 解像度のカメラで撮影した HDMI 信号が FPGA(Field-Programmable Gate Array)ボードに入力されて、上の 4 台のディスプレイ画像として出力されている。

2. ニューラルネットワーク

ニューラルネットワークは、脳を模式的に実装したものである。脳のニューロンを模倣した実装をコンピュータのソフトウェアとして実現するか、または、ハードウェアの素子として実現することもできる。ニューラルネットワークは脳と同様に学習をさせて能力を獲得することができる。また、ニューラルネットワークの演算はフィードバックがないため演算の並列性が高い。演算の並列性が高いため、ハードウェアつまり、FPGA に実装することが容易である。今回のニューラルネットワークには、3 層フィードフォワード型を使用している。図 2 に今回使用した 3 層フィードフォワード型ニューラルネットワークを示す。

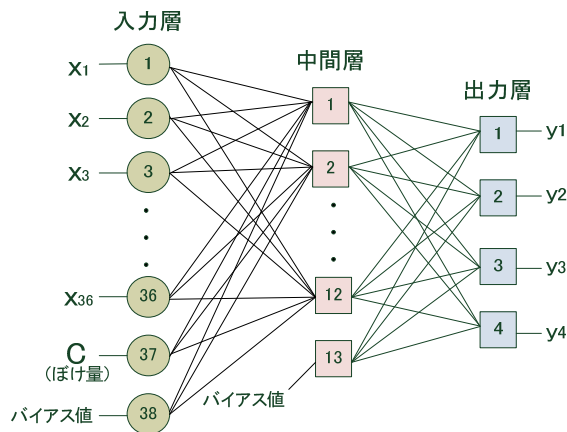


図 2. 解像度変換ボケ復元用ニューラルネットワーク

図 2 に示すように、38 入力、13 中間素子、4 出力のニューラルネットワークを解像度変換ボケ復元用として使用した。また、解像度変換やボケ復元に関して予め学習が終了したニューラルネットワークを使用している。

図 2 の 38 個の入力の $x_1 \sim x_{36}$ までの 36 個の入力はピクセルを入力し、残りの 2 個はぼけ量とバイアス値を入力する。38 個の入力にそれぞれの重みを掛けて足しあわせ、13 個の中間素子のそれぞれへ入力する。入力層から中間層への線と中間層から出力層までの線は、重みを掛けて足し合わせることを

表している。中間層ではシグモイド関数を掛けて後段に出力する。シグモイド関数は非線形な飽和関数であり、その数式を式 1 に示す。

$$\text{sigmoid}(z) = 1 / (1 + e^{-z}) \quad (\text{式 1})$$

Z の変域が-10 から 10 までのシグモイド関数のグラフを図 3 に示す。

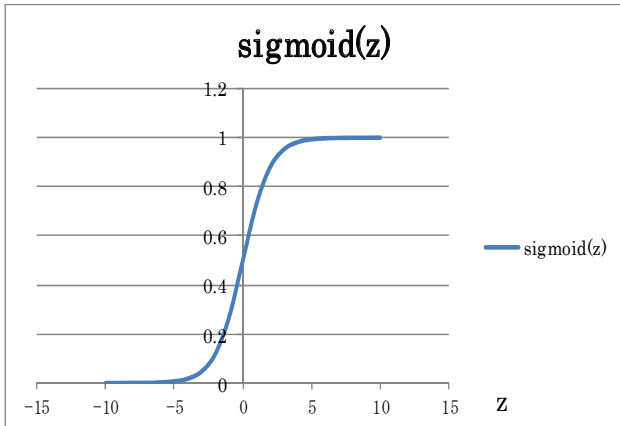


図 3. Z の変域が-10 から 10 までのシグモイド関数のグラフ

13 個の中間素子の出力にそれぞれの重みを掛けて足し合わせ、y1~y4 までの 4 個の出力層に入力する。出力素子は中間層同様にシグモイド関数を掛けて最終的な画像出力とする。

解像度ボケ復元システムでは、ニューラルネットワークを解像度変換とボケ復元に利用している。解像度変換ボケ復元用ニューラルネットワークのピクセルの入出力の様子を図 4 に示す。

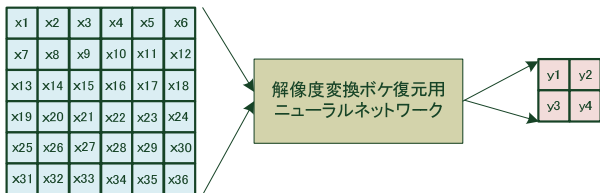


図 4. 解像度変換ボケ復元用ニューラルネットワークの入出力

ニューラルネットワークの 36 個の入力は、HD 画像の内の縦 6 ピクセル、横 6 ピクセルの正方形の画像を入力する。ニューラルネットワークの 4 個の出力は画像の出力となる。

HD 画像に対して 1 ピクセルごとにニューラルネットワークの入力をずらすごとに 4 ピクセルの画像を得る。よって、画像は縦 2 倍、横 2 倍の 4 倍に拡大され、HD の画像が 4K2K の画像に変換される。

3. 解像度変換ボケ復元用 FPGA 回路

解像度変換ボケ復元用 FPGA 回路は、TB-7K-325T-IMG の Xilinx 社の FPGA (XC7K325T-2FFG900) に実装され解像度変換ボケ復元システムの中核をなしている。解像度変換ボケ復元 FPGA 回路は、HDMI Input Module、HDMI Output Module、AXI Interconnect、DDR3 SDRAM Controller、MicroBlaze 回路に分けられる。その様子を図 5 に示す。

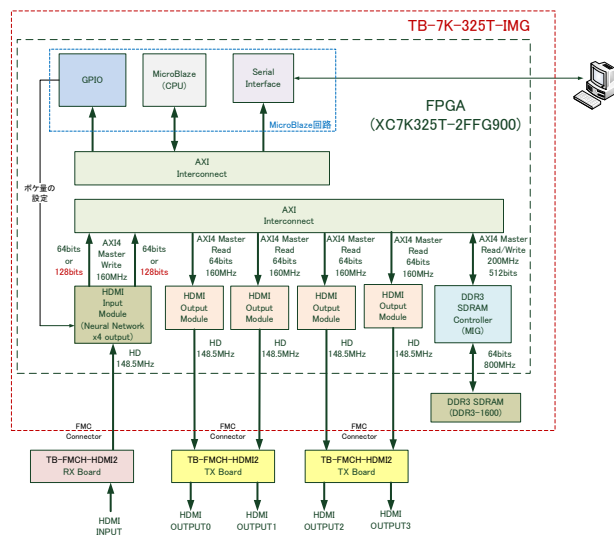


図 5. 解像度変換ボケ復元用 FPGA 回路

図 5 において、赤い破線で示されるのが FPGA ボードの TB-7K-325T-IMG である。FPGA ボード内の DDR3 SDRAM と FPGA を使用している。黒い破線の内側が FPGA 回路である。FPGA 回路内部の青い破線で示される部分が MicroBlaze 回路となる。HDMI Input Module、HDMI Output Module と TB-FMCH-HDMI2 との間は FMC コネクタで接続されている。

各モジュールは、IP (Intellectual Property) 化されている。IP は完全に独立しており、IP を交換することで簡単に機能を変更することが出来る。Xilinx 社の XPS(Xilinx Platform Studio)ツールによって IP 間が接続される。XPS の上位ツールには FPGA の回路を生成するための Xilinx 社の ISE14.7 を使用した。使用言語は VHDL(VHSIC Hardware Description Language) と Verilog HDL(Verilog Hardware Description Language)を混在して使用した。FPGA の LUT(Look Up Table)の使用率は 50 %である。

図 6 に解像度変換ボケ復元用 FPGA 回路の ISE14.7 プロジェクトを示す。

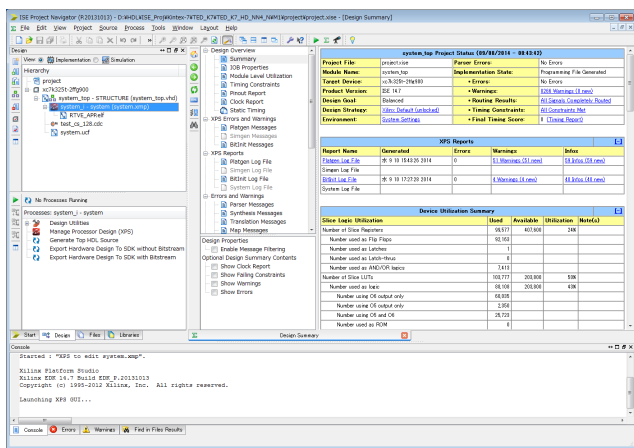


図 6. 解像度変換ボケ復元用 FPGA 回路の ISE14.7 プロジェクト

図 7 に解像度変換ボケ復元回路の XPS プロジェクトを示す。

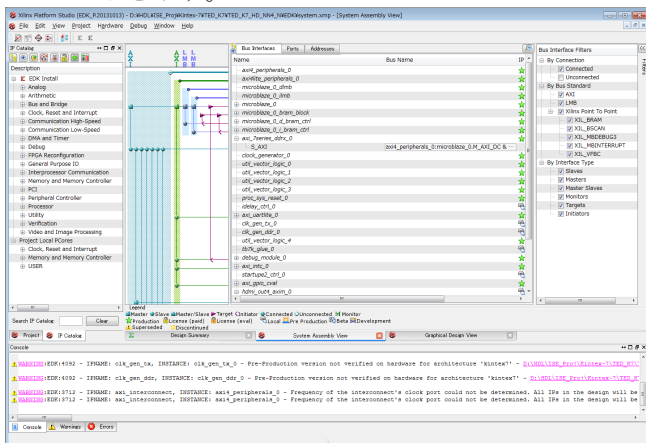


図 7. 解像度変換ボケ復元回路の XPS プロジェクト

3.1 実装に用いた FPGA ボード

解像度変換ボケ復元システムでは、東京エレクトロンデバイス社 inrevium 「Kintex-7 FPGA 搭載 ACDC 1.0 ベースボード」TB-7K-325T-IMG^[2] をベースボードとして、FMC コネクタで、HDMI1.4a インタフェースカード TB-FMCH-HDMI2^[3] を 2 組接続している。図 8 に解像度変換ボケ復元システムに使用している FPGA ボードの写真を示す。

TB-FMCH-HDMI2 は、TX ボードと RX ボードのセットで 1 組となっていて、1 枚のボードに HDMI の出力または入力コネクタが 2 個ずつ搭載されている。RX ボードは HDMI 信号を標準的なビデオの信号に変換し、TX ボードは標準的なビデオ信号を HDMI 信号に変換する。

TB-7K-325T-IMG は、Xilinx 社の FPGA (XC7K325T-2FFG900) が搭載されている基板である。DDR3 SDRAM は 4 個搭載されて、合計 64 ビットバス幅の DDR3-1600 として使用されているので、12.8Gbytes/sec の帯域がある。高性能な FMC(FPGA Mezzanine Card)コネクタが 4 個搭載され、そこに 2 組の TB-FMCH-HDMI2 が挿入されている。2 組の TB-FMCH-HDMI2 RX ボードは、合計 4 ポートある内の 1 ポートのみ HD 解像度の HDMI 信号を入力する。入力された HDMI 信号

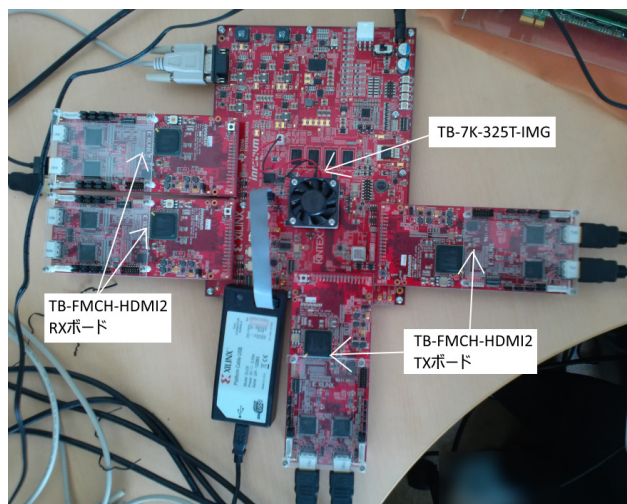


図 8. 解像度変換ボケ復元システムに使用している FPGA ボードの写真

は RX ボード内の FPGA で処理され、標準的なビデオ信号に変換されて FMC コネクタに出力される。FMC コネクタを通して接続された TB-7K-325T-IMG は、標準的なビデオ信号を受け取り、FPGA 内に実装された解像度変換ボケ復元ニューラルネットワークに画像を入力する。ニューラルネットワークから出力された 4 倍画像は、一旦、DDR3 SDRAM に Write される。DDR3 SDRAM から Read されたピクセルは、画像の各象限に分けられ、FMC コネクタ経由で 2 組の TB-FMCH-HDMI2 TX ボードに送られる。2 組の TB-FMCH-HDMI2 TX ボードから 4 台のディスプレイに送られて、全体で 4K2K 画像として表示される。

3.2 HDMI Input Module

HDMI Input Module は、HD 解像度のビデオ信号を 4K2K 解像度に変換し、その 4K2K のピクセル・データを DDR3 SDRAM に Write するモジュールである。HDMI Input Module は、1 個の HD 解像度のビデオ信号入力と 2 個の 128 ビット・データ幅または 64 ビット・データ幅の AXI4 Master インターフェイス(Write のみ)を持っている。現在は 128 ビット・データ幅を使用している。この理由については、“5. AXI Interconnect と MIG の仕様による不具合とその解消手法”で詳しく述べる。HDMI Input Module は、ビデオ信号入力から得られたピクセルを Red, Green, Blue を担当する 3 つの解像度変換ボケ復元ニューラルネットワークに入力する。なお、図 5 に示すようにビデオ信号は TB-FMCH-HDMI2 RX ボードから送られる。ニューラルネットワークの出力は、図 4 に示すように y1 ~ y4 の 4 つの出力となる。y1 と y2 は画像のフレーム・バッファ上で連続したアドレスになるが、y3, y4 は y1, y2 と連続したアドレスにならない。例えば、4K2K 解像度画面の左上を 0 番目とした時に y1 をピクセルの 0 番目とすると、y2 は 1 番目となる。y3 は下の行になるので、y0 の番目に 4K2K 解像度分の横のピクセルが足され、3840 番目となる。同様に y4 は 3841 番目となる。HD 解像度画面のピクセルの様子を図 9 に示す。



図 9. HD 解像度画面のピクセルの番地

スループットを向上させるためには、メモリのアクセスにおいて連続したデータ転送を行う必要がある。y1, y2 と y3, y4 は別々にデータ転送すると AXI4 バスのスループットが向上するので、それぞれ別々の AXI4 Master インターフェイス回路にデータ転送を任せている。従って、HDMI Input Module は 2 つの AXI4 Master Write インターフェイス(axi4_master_inf)を持つことが図 5 に示されている。それぞれの axi4_master_inf は、128 ビット・データ幅の 32 バーストを 30 回行うことで、3840 ピクセルの画像の 1 行分を DDR3 SDRAM に Write することができる。

HDMI Input Module のブロック図を図 10 に示す。

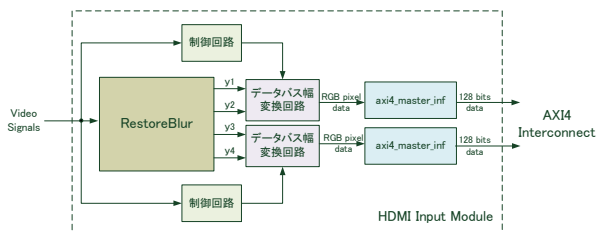


図 10. HDMI Input Module のブロック図

図 10 に示すように HDMI Input Module は、RestoreBlur、2 個の制御回路、2 個のデータバス幅変換回路、2 個の axi4_master_inf により構成されている。

RestoreBlur は、RGB(Red Green Blue) 3 つのニューラルネットワークを搭載し、解像度変換ボケ復元を行うシステムの要の回路である。

データバス幅変換回路は、ニューラルネットワークの出力の内の y1, y2 を担当するデータバス幅変換回路と y3, y4 を担当するデータバス幅変換回路の 2 個が搭載されている。データ幅変換回路は、それぞれのピクセル・データのデータ幅を 128 ビットまたは 64 ビットに変換する。データバス幅変換回路を制御する制御回路も、2 個搭載されている。

axi4_master_inf は、AXI4 Master インターフェイス用の汎用モジュールである。AXI4 バスの Write 及び Read トランザクションを行うインターフェイス回路であるが、HDMI Input Module では Write の機能のみを使用している。なお、axi4_master_inf は AXI4 バスの利点を最大限利用するため、アドレス転送とデータ転送、応答転送を完全にオーバーラップできるように設計されている。

HDMI Input Module は AXI4 バスのデータバス幅を 64 ビット幅と 128 ビット幅の 2 つに変更することができる。当初は 64 ビット幅固定になっていたのだが、64 ビットデータバス幅では画面が乱れて、正常に 4K2K 解像度で表示することができなかった。そこで、AXI4 バスのデータバス幅を 128 ビットにしたところ正常に表示することができた。

DDR3 SDRAM のデータ帯域幅には余裕があったのだが、AIX Interconnect と MIG の仕様によって、HDMI Output Module の Read トランザクションが HDMI Input Module の Write トランザクションの間に入れなかったのが原因だった。MIG のデータ帯域幅には、まだ余裕があったので、HDMI Input Module の AXI4 バスのデータ幅を 128 ビットとした。これで、Read トランザクションが無理なく Write トランザクションの間に入ることができて正常に画像を表示することができた。

3.3 RestoreBlur

HDMI Input Module 内で解像度変換ボケ復元用ニューラルネットワークを 3 つ持つモジュールが RestoreBlur である。

RestoreBlur のブロック図を図 11 に示す。

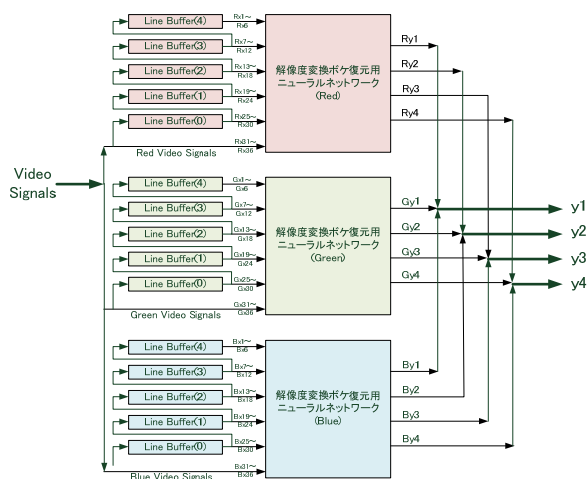


図 11. RestoreBlur ブロック図

図 11 に示すように、RestoreBlur は RGB それぞれの 5 つのライン・バッファと解像度変換ボケ復元ニューラルネットワークで構成される。

図 4 に示すニューラルネットワークの入力の x36 は現在のビデオ信号のピクセルとなる。x35 は 1 クロックの遅延素子で x36 よりも 1 クロック遅れたピクセルとなる。また、x34 は同様に 1 クロックの遅延素子で x35 よりも 1 クロック遅れたピクセルとなる。x31 までは同様に 1 クロック遅れたピクセルをニューラルネットワークに入力する。

x25~x30 までのピクセルは、1 ライン前のピクセルとなる。同様に x25~x30 までは、それぞれ数字の若いほうが 1 クロック前のピクセルとなる。例を挙げると x25 は x26 よりも 1 クロック前のピクセルとなる。それ以降も同様である。

3.4 解像度変換ボケ復元ニューラルネットワークのハードウェア実装

図 11 の 3 個の解像度変換ボケ復元ニューラルネットワークは図 2 に示すように、3 層フィードフォワード型ニューラルネットワークである。ニューラルネットワークは、当初はソフトウェアの浮動小数点演算で実装されていた。ソフトウェアからハードウェアに実装するにあたって固定小数点演算へ変換を行った。最初に重みの変域をプロファイリングし、重みのビット長を確定した。このフォーマットは符号部 1 ビット、整数部 7 ビット、小数部 4 ビットの計 12 ビットで表現した^[1]。重みの演算も可能な限り演算ビット長を削減するように考慮して、ニューラルネットワークの回路が小さくなるように工夫した。図 2 にビット長を書き加えたニューラルネットワークを図 12 に示す^[4]。

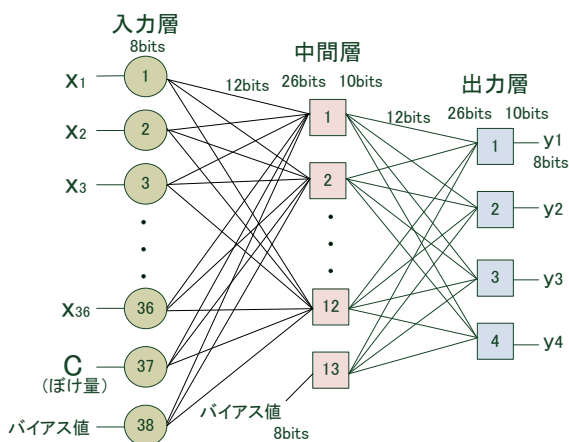


図 12. ニューラルネットワークのビット長

図 12 において、入力は 8 ビット長の RGB のいずれかの値となる。入力に 12 ビット長の重みを掛けて、それぞれの値を足し合わせる部分は 26 ビット長の演算器を持てばオーバーフローすることはない。中間層のシグモイド関数は 10 ビット入力 10 ビット出力のシグモイド関数テーブルを使用した。これはリソース削減のために FPGA の 1 個の BlockRAM に収まるテーブルを使用したためである。シグモイド関数テーブルに入力する前には、適切なビット位置で飽和演算を行い 10 ビットに丸めた。中間層の出力はシグモイド関数テーブル出力の 10 ビットとなる。中間層出力と 12 ビット長の重みを掛けて足し合わせる部分は、同様に 26 ビット長の演算器を持つ。適切なビット位置で飽和演算を行い 10 ビットに丸め、出力層のシグモイド関数テーブルに入力する。出力層のシグモイド関数テーブル出力 10 ビットを 8 ビットに丸めて y1~y4 出力とした。

このニューラルネットワークは、VHDL で書かれているが、この VHDL は重みのテーブルからオブジェクト指向言語の Ruby で書かれたスクリプトで自動生成されている。変換する際には、VHDL の掛け算記号の*を使うのではなく、重みが固定値であることを利用して、最適化を行った。入力値をシフトし、加算していくことで掛け算を実行している。論理合成ツールに最適化を丸投げするのではなく、あ

る程度、自分で最適化を行うことにより、より最適化され、FPGA の使用リソースが小さくなる。

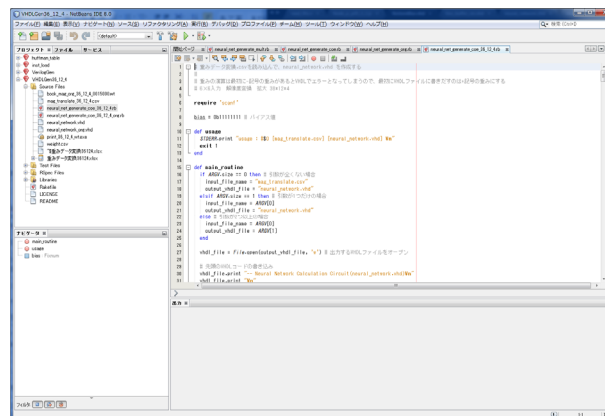


図 13. 重みテーブルから VHDL に変換する Ruby スクリプトの一部

図 13 に重みテーブルから VHDL に変換する Ruby スクリプトの一部を示す。

なお、Ruby スクリプトは、NetBeans 8.0 IDE を使用して作製した。

図 14 に Ruby スクリプトで生成されたニューラルネットワークの VHDL コードの一部を示す。

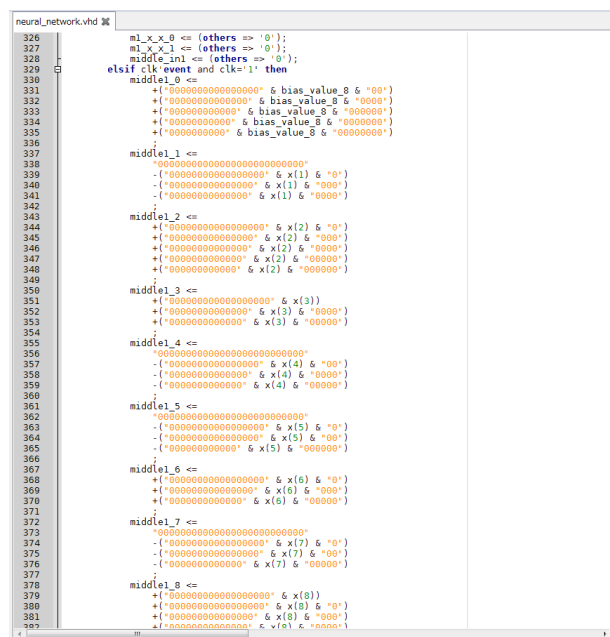


図 14. Ruby スクリプトで生成されたニューラルネットワークの VHDL コードの一部

3.5 HDMI Output Module

HDMI Output Module は、HDMI Input Module によって DDR3 SDRAM に書かれた 4K2K のピクセル・データを Read してディスプレイに表示させるモジュールである。4K2K 解像度画面は 4 個の HD 解像度画面として、4 個の HD 解像度ディスプレイに表示する。4 個の HD 解像度ディスプレイに対応した 4 個の HDMI Output Module を搭載してある。4K2K 解像度画面の分割方法を図 15 に示す。

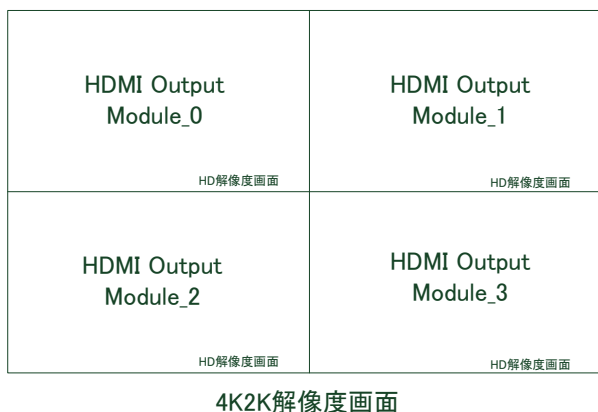


図 15. 4K2K 解像度画面の分割方法

HDMI Output Module のブロック図を図 16 に示す。

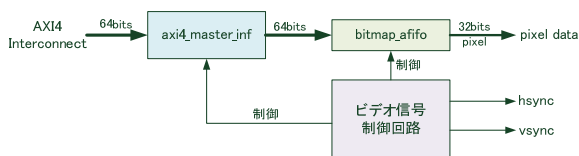


図 16. HDMI Output Module のブロック図

HDMI Output Module は、図 16 に示すように、axi4_master_inf、bitmap_afifo、ビデオ信号制御回路の 3 個のブロックから構成される。axi4_master_inf は前述したように、AXI4 Master インターフェース用の汎用モジュールである。ここでは、AXI4 Master Read 機能のみを使用している。bitmap_afifo は入力が 64 ビット幅 512 深度、出力が 32 ビット 1024 深度の非同期 FIFO である。64 ビット幅の AXI4 バスを 32 ビット幅に変換する機能と AXI4 バスの動作周波数とビデオの動作周波数の変換機能がある。ビデオ信号制御回路は、各モジュールの制御とビデオ制御信号の hsync と vsync を出力する。

HDMI Output Module の動作を説明する。

1. 画面を表示する前にビデオ信号制御回路は、axi4_master_inf に Read の要求を送る。
2. ピクセルは DDR3 SDRAM から Read され、bitmap_afifo に送られる。
3. bitmap_afifo にピクセルが貯まる。
4. 画面を表示する時には、ピクセルの RGB の値が bitmap_afifo から TB-FMCH-HDMI2 TX ボードに送られる。
5. ビデオ信号制御回路からは、hsync と vsync も同様に TB-FMCH-HDMI2 TX ボードに送られる。
6. 画面を表示していると bitmap_afifo に貯められたピクセルは消費されて減ってくる。bitmap_afifo に貯められたピクセルが bitmap_afifo の容量の半分を下回ると、ビデオ信号制御回路は、bitmap_afifo の半分の容量のピクセルを Read するために axi4_master_inf に要求を送る。
7. axi4_master_inf は、AXI4 バス越しに DDR3 SDRAM からピクセル値を Read し、bitmap_afifo に供給する。

8. 画面の表示期間中、6 と 7 を繰り返すことにより、ピクセル値を枯渇させることなくディスプレイに供給することができる。

3.6 AXI Interconnect

AXI Interconnect は、Xilinx 社の IP で、AXI Master モジュールと AXI Slave モジュールを接続する。周波数変換やバス幅変換の機能があり、動作周波数やバス幅の異なるモジュールを手軽に接続することができる。図 5 に示すように、MicroBlaze 関連と HDMI Input/Output 関連の 2 個の AXI Interconnect が実装されている。

3.7 DDR3 SDRAM Controller

DDR3 SDRAM Controller は Xilinx 社の MIG(Memory Interface Generator)と呼ばれる IP を使用している。64 ビット・データバス幅、DDR3-1600 で動作しているもので、最大データ帯域は 12.8GBytes/sec である。AXI4 バスのインターフェース幅は 512 ビット、200MHz の動作周波数であり、DDR3 SDRAM の最大データ帯域に見合ったバス・スピードとなっている。

3.8 MicroBlaze 回路

MicroBlaze 回路は、Serial Interface と MicroBlaze、GPIO(General Purpose Input/Output)で構成される。各モジュール間は AXI Interconnect で接続されている。MicroBlaze 回路の IP は全て Xilinx 社の IP を使用している。MicroBlaze は FPGA 上にソフトコアで実装したプロセッサである。BlockRAM に実装したソフトウェアによりパーソナルコンピュータとシリアル接続を介して通信をしながら、ボケ量の設定

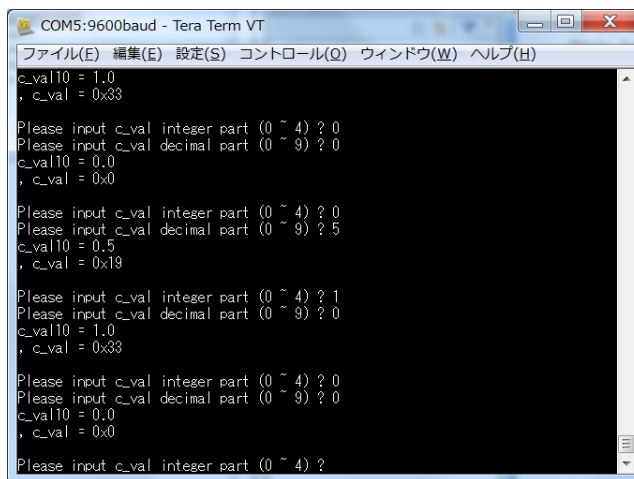


図 17. Tera Term を使用したボケ量設定の様子

を行う。Serial Interface のビットレートは 9600bps(bits per second)である。Tera Term を使用したボケ量設定の様子を図 17 に示す。

4. まとめ

RGB 用の 3 個の解像度変換ボケ復元用ニューラルネットワークを FPGA に搭載し、解像度変換ボケ復元システムを構築した。まだ多少の不具合が

あつて開発途上ではあるが、HD 解像度の画像を4K2K に変換することができた。

今回の事案では、初めて機能モジュールを IP として作製し、IP 間を XPS で接続することで FPGA 内の回路を構築していく新しい FPGA の構築方法を採用した。これによって、HDL(Hardware Description Language)を書く手間が省け、効率化することができた。

謝辞

解像度変換ボケ復元システムを制作する機会を与えて頂いた、筑波大学システム情報系の和田耕一教授に深く感謝いたします。

参考文献

- [1] 伊藤 翼, FPGA を用いた画像の高品質復元ハードウェアの研究, 筑波大学大学院博士課程システム情報工学研究科修士論文, 2014 年 3 月.
- [2] TB-7K-325T-IMG ハードユーザマニュアル, Rev.1.05, 2013/06/05,
http://ppg.teldevice.co.jp/m_board/inrevium/TB-7K-325T-IMG_HWUserManual_1.05.pdf
- [3] TB-FMCH-HDMI2
ハードユーザマニュアル 2 入出力版(工場出荷版), Rev.1.03, 2014/08/13,
http://ppg.teldevice.co.jp/m_board/inrevium/TB_FMCH_HDMI2_HWUserManual_1.03.pdf
- [4] 小野 雅晃, ボケ復元ニューラルネットワークのハードウェア化, FPGA エクストリーム・コンピューティング 第 4 回 発表, 2013/11/10,
<http://connpass.com/event/3666/>,
<http://www.slideshare.net/marsee101/ss-28113255>

Development of 4K2K Resolution Conversion and Blur Restoration Circuit based on Neural Network

Masaaki Ono

Technical Service Office for Systems and Information Engineering, University of Tsukuba,
1-1-1 Tennodai, Tsukuba, Ibaraki, 305-8573 Japan

Keywords: neural network, FPGA, 4K2K