

全画面大腸内視鏡画像に適した リアルタイム特徴量抽出の FPGA 実装

清水達也^{†1} 小出哲士^{†1} Anh-Tuan Hoang^{†1} 杉幸樹^{†1} 岡本拓巳^{†1} 佐藤光^{†1}
玉木徹^{†2} Bisser Raytchev^{†2} 金田和文^{†2} 吉田成人^{†3,1} 三重野寛^{†3} 田中信治^{†4}

本稿では大腸 NBI 拡大内視鏡観察により得られる大腸粘膜表面の微細血管模様を特徴量とし、病理識別を行う診断支援システムの構築を目指す。最適な診断支援のためには全画面で処理を行う必要があり、高速な処理のためにハードウェアへの実装を行う。本稿では乗算器や除算器をシフトと加算器で代用し、パイプライン化によりストリーミング処理を可能とした、ハードウェア向け D-SIFT 特徴量抽出アーキテクチャを FPGA 実装した結果について述べる。実装の結果、最大動作周波数 170 MHz レイテンシ 61 msec @ 100MHz, スループット 16 fps @ 100MHz を達成し、リアルタイムでの全画面特徴量抽出を実現した。

FPGA Implementation of Real-time Feature Extraction for Full HD Colorectal Endoscopic Images

TATSUYA SHIMIZU^{†1} TETSUSHI KOIDE^{†1} ANH-TUAN HOANG^{†1} KOKI SUGI^{†1} TAKUMI OKAMOTO^{†1}
HIKARU SATOH^{†1} TORU TAMAKI^{†2} BISSER RAYTCHEV^{†2} KAZUHUMI KANEDA^{†2}
SHIGETO YOSHIDA^{†3,1} HIROSHI MIENO^{†3} SHINJI TANAKA^{†4}

This paper shows the implementation for feature extraction of fine vascular patterns of the large intestine mucosa surface taken by NBI magnifying endoscope. It aims to build a diagnostic support system for pathology identification. The implementation in hardware supports for high performance processing of Full HD image. Instead of using complex multipliers, dividers, the implementation uses shifters and adders only. The pipeline implementation is suitable for stream processing. Results of D-SIFT feature extraction module implementation on FPGA shows that the implementation meets requirements of latency and throughput from medical doctors.

1. はじめに

近年大腸ガン罹患患者数は世界的に増加の一途を辿っている。しかし大腸ガンは早期段階での発見、治療によりほぼ完治可能な病気でもある。そのための検査方法として図 1 のような NBI (Narrow Band Imaging) システムを用いた大腸拡大内視鏡検査が行われている [1]。これは大腸壁面の微細血管模様から腫瘍の有無やガンの深達度の診断を行うが、これには医師の経験と知識が必要となり、限られた医師しか診断が行えないという問題がある。そこで症状を定量的に評価することによる医師の診断支援や、経験の少ない若手医師の教育支援を可能とする CAD (Computer-Aided Diagnosis) システムが求められている。これまでに広島大学病院で提唱されている NBI 拡大所見分類 [1]に基づき、非腫瘍性病変、腺腫、浸潤ガンの 3 クラス (Type A, B, C3) に分類する大腸 NBI(Narrow Band Imaging)拡大内視鏡診断支援システム(eCAD システム)がソフトウェア実装により開発されている [2, 3]。

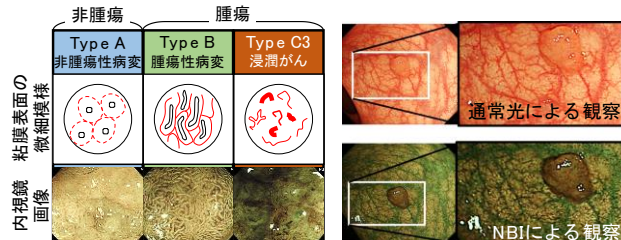


図 1. NBI 拡大所見分類。

これは 120 x 120 pixel の単一領域に対して実用上十分な正率率約 90%、フレームレート 14.7 fps を実現している。しかし病理の正確な診断のためには大腸内視鏡画像全画面で識別を行う必要があり、10 pixel 間隔でラスタスキャンを行った場合、およそ 20 分の処理時間がかかり、リアルタイムでの診断支援が困難である。そこで我々の研究グループでは eCAD システムのハードウェア実装によりリアルタイムな診断支援システムの構築を目指す。医療現場からの要求性能である、(I) フレームレート 1~5 fps 以上、レイテンシ 1 sec 以下、(II) 非腫瘍の識別精度 90 %以上の 2つを実現するため FPGA 実装を行う。

^{†1} 広島大学ナノデバイス・バイオ融合科学研究所
Research Institute for Nanodevice and Bio Systems, Hiroshima University.
^{†2} 広島大学 工学研究院
Graduate School of Engineering, Hiroshima University
^{†3} JR 西日本 広島鉄道病院 消化器内科
Department of Gastroenterology, Hiroshima General Hospital of West Japan

Railway Company
^{†4} 広島大学病院 医歯薬保健学研究所 内視鏡医学
Department of Endoscopy and Medicine, Graduate School of Biomedical and Health Science, Hiroshima University

2. eCAD システム

2.1 NBI (Narrow Band Imaging)

NBI とは消化管内視鏡に用いられる特殊光内視鏡技術で、従来の白色光の前に特殊な光学フィルタを入れることで血管中のヘミグロビンに吸収されやすい2つの狭帯域な波長 (415 nm, 540 nm) を照射することで、大腸壁面の血管や腺構造を強調表示するものである。

2.2 eCAD システム概要

eCAD システムの概要を図2に示す。本システムは Bag-of-Features (BoF) と呼ばれる分類手法を用いる。元々は Bag-of-Words (BoW) と呼ばれる文書分類手法を画像に応用したもので、BoW とは既知の文章から特徴語を抽出し、その特徴語の出現頻度から文書を分類するものである。eCAD の処理は (1) 特徴量抽出、(2) 特徴量変換、(3) タイプ識別の3つの処理で構成されている。オフラインで行う学習フェーズ、オンラインで行う識別フェーズでこの3つの処理を行う。まず学習フェーズで分類 Type が既知である大腸内視鏡画像 (学習画像) を入力する。(1) 特徴量抽出で学習画像の輝度勾配・強度を算出し、その情報を128次元のベクトルとして表すことで、大腸壁面の微細血管模様の特徴量を表現する。特徴量抽出により得られたベクトル情報を(2) 特徴量変換でクラスタリングし、各クラスタの中心を Visual Word として保存する。識別フェーズでは (1)、(2) の処理を、識別する内視鏡画像 (テスト画像) に対して行い、VW の出現頻度を示す Visual Word ヒストグラムを作成する。(3) タイプ識別では学習フェーズで作成したヒストグラムと識別フェーズから得られた Visual Word とを比較することにより、Type A、Type B、Type C3 の3タイプの識別を行う。本システムでは特徴量抽出と Visual Word 作成に VLFeat の Dense Scale Invariant Feature Transform (D-SIFT) [4]、階層的 k-means 法を用い、タイプ識別に LIBSVM の Support Vector Machine (SVM) [5]を使用した。

3. D-SIFT 特徴量抽出アルゴリズム

図3に D-SIFT 特徴量抽出の概要を示す。D-SIFT [4]は、入力画像に対して一定間隔 (grid) で特徴点 (key point) をとる。key point は周囲 20×20 pixel における輝度勾配方

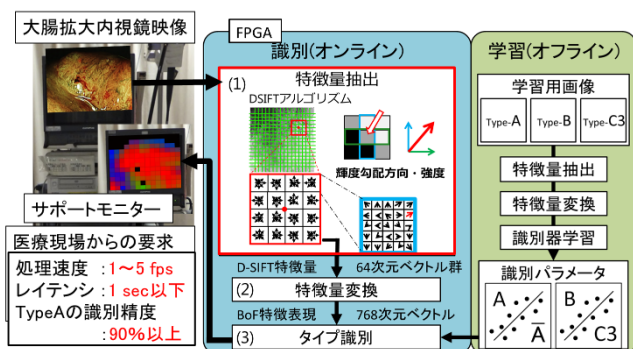


図2. eCAD システム概要.

向・強度情報を128次元のパラメータである特徴量算出したものである。大腸内視鏡画像のように画像全体に重要な特徴量が存在する場合、特徴量の取りこぼしを防ぐことができ、非常に有効である。

処理としては、画像をグレースケールに変換後、Gaussian Filterにより平滑化を行い、各 pixel の x 方向、y 方向の輝度勾配を算出する。次に輝度勾配方向を8方向に分類し、輝度勾配強度を算出する。その後 grid size の矩形領域 bin で key point との距離に応じて方向ごとに畳み込みを行う。最後に得られた Key point 内の $4 \times 4 \times 8$ 方向 = 128 次元の値を正規化することにより、128次元の特徴量ベクトルが算出される。eCAD システムでは grid size を 5 pixel, bin サイズを 5 pixel, 7 pixel に設定している。この手法をソフトウェア手法と言い、以下に示すハードウェア向けに改良したものをハードウェア手法とする。

4. ハードウェア向け D-SIFT 特徴量抽出アルゴリズム

4.1 平滑化処理

入力画像をグレースケールに変換した後、画像の平滑化を行う。これは画素ごとの細かい変化量を少なくし、局所領域内の変化量を明瞭にする目的がある。ソフトウェア手法ではこの平滑化の際にガウシアンフィルタを使用する。Gauss 関数分散 σ 、ウィンドウサイズ S、Scale は図4のように設定することで、良い精度を得られることがわかっている [2]。ハードウェア手法ではこの平滑化処理に、係数を2のべき乗にしてシフトと加算器のみで実現している Simple Gaussian Filter (図4)を使用している [6]。

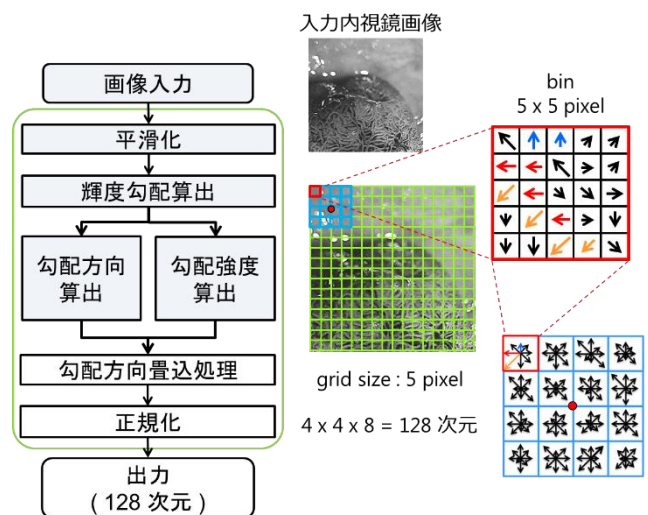


図3. D-SIFT 特徴量抽出概要.

4.2 輝度勾配方向・強度算出処理

輝度勾配方向・強度は各 pixel の x 方向の差分 G_x , y 方向の差分 G_y により求める. ソフトウェア手法では正確な精度で計算を行うため, 輝度勾配方向算出に Arctan を用い, 強度算出にルート計算を用いる. しかし, これは計算が複雑化するため, ハードウェア実装に不向きである. そこでハードウェア手法では図5のように, (1) G_y , G_x の符号ビットの組み合わせにより 4 方向に分類し, (2) さらに G_y , G_x の大小関係により 2 分割することで, 8 方向への分類し, 輝度勾配方向を決定する. これにより計算の複雑化の要因であった Arctan やルートの計算を除くことができ, かつシミュレーションにより, 符号ビットで分類した 4 方向 (64 次元) のみでもソフトウェア手法と同程度の精度があることを確認している [7].

4.3 bin の畳み込み処理省略によるブロック値の共有

ソフトウェア手法では 5×5 pixel サイズの bin 内で輝度勾配強度を方向ごとに足し合わせ, 20×20 pixel サイズの key point で特徴量の畳み込みを行うことで特徴量を算出する. しかしこの場合, 隣り合う Key point が同一のブロックであっても, key point が移動した際に重み付けにより異なるブロック値を持つことになる. 図6に示すハードウェア手法では, この重み付け処理の重み係数を 1 とすることで, ブロック値の共有を可能とし, 約 95 % のメモリの削減が可能となる. また任意のブロックの計算の高速化が可能となる.

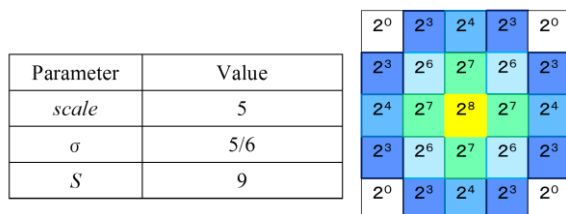


図4. Simple Gaussian Filter 係数.

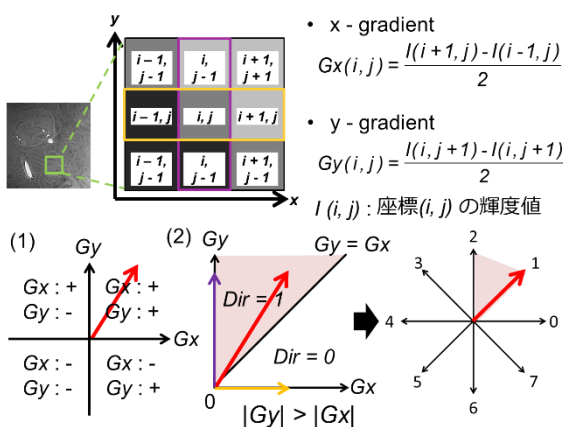


図5. 輝度勾配方向・強度算出ハードウェア手法.

4.4 正規化処理の代替案

内視鏡画像には様々な明るさがあり, 明るさによって出力される特徴量が異なる. ソフトウェア手法では, 各特徴量を算出した後, 画像の明暗変化に対応するため, 特徴量の正規化処理を行う. しかし正規化処理には除算, ルート計算, 2 乗計算, 128 次元の入力が必要となり, 計算が複雑化してしまう. そこでハードウェア手法では Gaussian Filter による平滑化の後にしきい値処理を行うことで輝度値による変化量を一定にし, 正規化を省略した. これにより, 計算の簡略化をし, かつピーク部分を残したまま各特徴量のレンジを制御した.

4.5 ハードウェア向け D-SIFT の有効性検証

ソフトウェア手法とハードウェア手法の比較結果について示す[7]. Visual Word 数を 768 個とし, タイプ識別にはハードウェア向け SVM 手法 [8,9]を用いる. データセットには広島大学病院内視鏡診療科で撮影された, NBI 拡大所見分類の各 Type の構造が観測される領域をトリミングした大腸 NBI 拡大内視鏡画像 (約 100×300 pixel ~ 900×800 pixel) 1260 枚を使用する. 評価指標には表1をパラメータとし, 各タイプの識別の正確さを判断する True Positive を使用する. 図7には 10 fold-Cross Validation を 10 回行った平均値を示す. ソフトウェア手法を SW, ハードウェア手法を HW とし, 輝度勾配方向で 8 方向算出したものを 128 次元, 4 方向のみ使用したものを 64 次元としている. 結果から, ソフトウェア手法 128 次元と比較して, ハードウェア手法は同程度かそれ以上の精度があることが確認できる. また輝度勾配方向算出の際に 4 方向の分類のみ行ったハードウェア 64 次元手法でも同程度の精度が確認している.

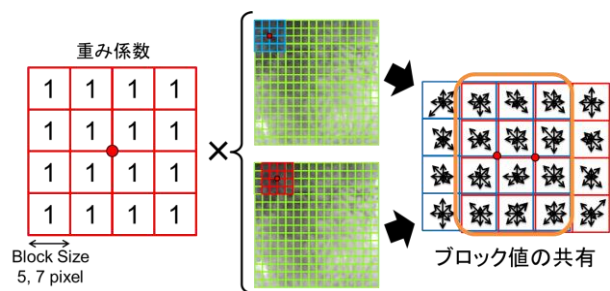


図6. 畳み込み省略によるブロック値の共有.

表1. 評価方法に用いるパラメータ.

i	Type A, Type B, Type C3, All
$Img_Num(i)$	i の全画像データ数
$Pos_Num(i)$	正しく i と識別された数

$$True\ Positive(i) = \frac{Pos_Num(i)}{Img_Num(i)} \times 100[\%] \quad (1)$$

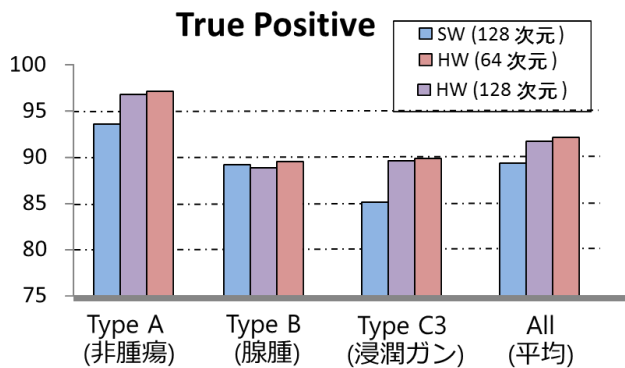


図 7. ハードウェア向け D-SIFT の有効性 [7].

5. ハードウェア向け D-SIFT アーキテクチャ

図 8 にハードウェア向け D-SIFT アーキテクチャのブロック図を示す。図 9 に平滑化処理回路 (SGF : Simple Gaussian Filter), 図 10 に輝度勾配方向・強度算出回路 (GC : Gradient Calculator), 図 11 にブロック特徴量算出回路 (BFC : Block Feature Calculator)を示し、ハードウェア向け D-SIFT はこれら 3つの回路により構成される。これら 3つは、Line Buffer を使用することで pixel 毎のパイプライン処理を実現している。本稿での実装結果については、SGF の 5 列目、GC の 3 列目のレジスタの値を後段の Line Buffer に入力するチェーン構造となっている。また、リソースと計算時間の削減が可能のため、ハードウェア 64 次元手法の実装を行った。

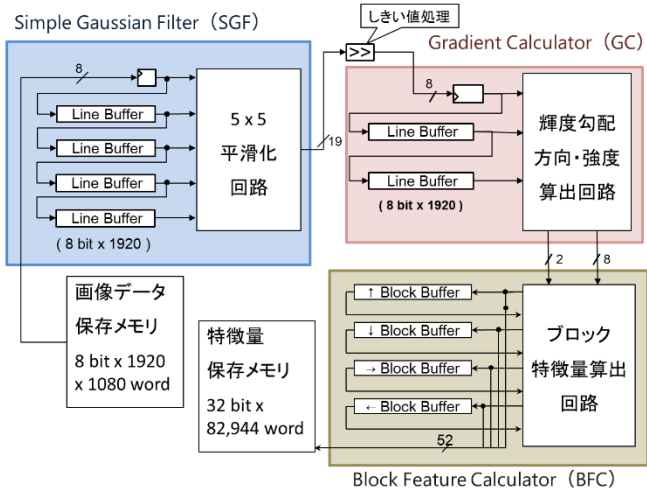


図 8. D-SIFT 特徴量抽出アーキテクチャ。

6. ハードウェア向け D-SIFT 動作検証

ハードウェア向け D-SIFT の動作概要について説明する。1つ目の bin (5x5 pixel) が出力されてから 1つ目の Scan Window (60x60 pixel) の特徴量抽出処理が完了するまでのラスタスキャンの様子を図 12 に示す。入力のピクセルは Scan Window ごとではなく、1行ずつラスタスキャンを行う。そのため、1つ目の key point が出力されるのは、図

12 の key point valid #1 の bin が算出されたタイミングとなる。2つ目からの key point の出力は次の bin ごとに算出される。key point valid #381 が出力された後に次の行の bin の処理を開始し、key point valid #382 で 2 行目の最初の key point を出力する。key point valid #4,115 (Scan Window valid #1) が出力されると、後段の特徴量変換、タイプ識別で処理の最小単位となる Scan Window が生成され、特徴量変換の処理が開始される。

ハードウェア向け D-SIFT の動作検証を Altera 社の ModelSim-Altera を使用して行った。検証結果より作成したタイミングチャートを図 13 に示す。SGF では、入力画像 1920x1080 pixel の外周 2 pixel に対してフィルタ処理を行う。そのために、領域の外 2 pixel にダミーのデータとして 0 を入力画像外周に付与し、 $(1920+2) \times (1080+2)$ pixel のデータとして処理を行う。このことを踏まえて、SGF の 1つ目の出力のタイミングは図9の上から2つの Line Buffer に 2 行分のデータが入り、3 行 3 列目のデータがレジスタに入力された時であり、 $(1920+2) \times 2 \text{ row} + 3 \text{ column} = 3,847$

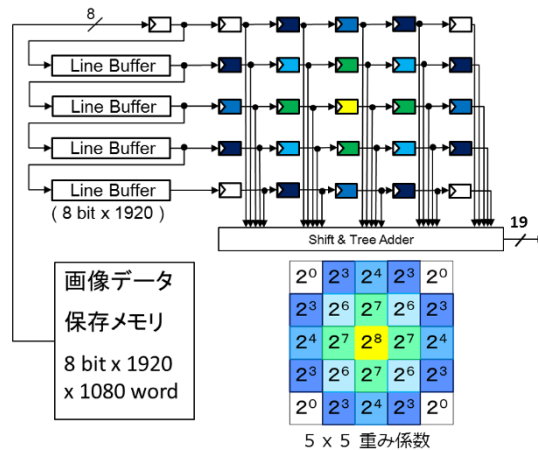


図 9. 平滑化回路 (Simple Gaussian Filter)。

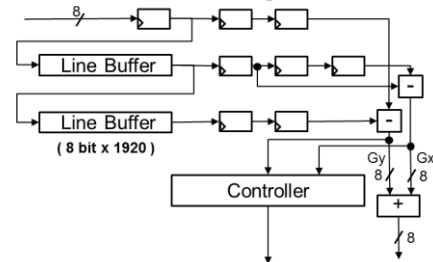


図 10. 輝度勾配方向・強度算出アーキテクチャ。

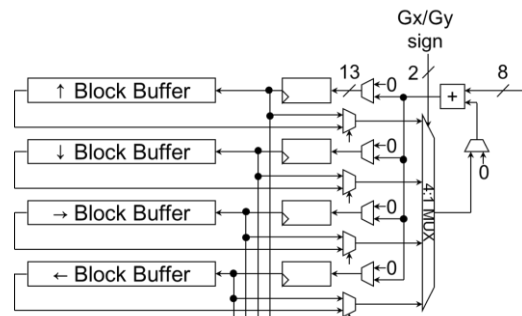


図 11. ブロック特徴量算出アーキテクチャ。

clock で出力されることになる。 Gradient Calculator についても同様で、領域の外 1 pixel に 0 を入力する。そのため図 10 の上段の Line Buffer を満たし、 2 行 2 列目のレジスタに入力されるタイミングで出力されるので、 Simple Gaussian Filter の出力から 1 つ目の Gradient Calculator の出力までに 1925 clock かかる。 Block Feature Calculator では領域外へのダミーデータ入力を行わない。そのため 4 行分のデータを入力し、図 11 の 5 行 5 列目のレジスタにデータが格納されたタイミングで出力されるので、 Gradient Calculator から 7,697 clock かけて出力される。これら 3 つの回路の動作について図 13 のタイミングチャートで確認することができる。図 13 の key point (20×20 pixel), Scan Window (60×60 pixel)の動作については、3 つの回路の動作検証により得られた結果から算出したものとなる。処理速度については、動作周波数 100 MHz で動作させた場合、レイテンシ 61 msec、スループット約 16 fps で動作可能であることを確認した。

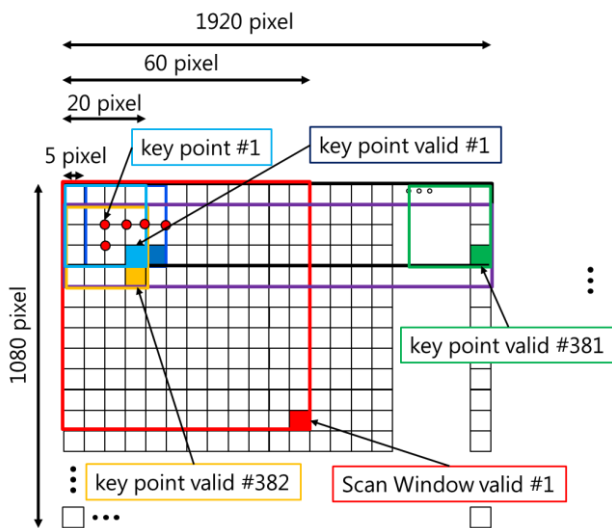


図 12. 各 key point 生成時のブロック数。

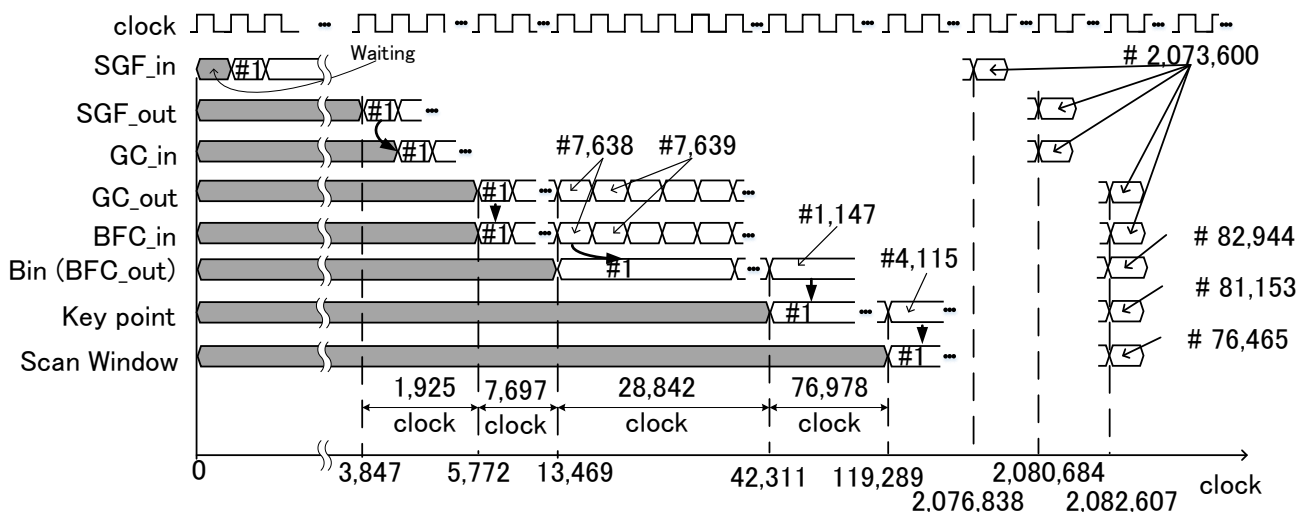


図 13. ハードウェア向け D-SIFT タイミングチャート。

7. ターゲットプラットフォーム

ハードウェア向け D-SIFT アーキテクチャの実装のため GiDEL 社製の PROCe IV 530-A の FPGA ボードを使用し、Altera 社の Stratix IV へ FPGA 実装を行った。プラットフォームの詳細について図 14 に示す。FPGA ボードと Host PC は PCI Express x4 で接続し、インタフェースには Proc Wizard 9.0 を使用した。各モジュールのリソースの使用率と、全体のリソース使用率については表に示す。各モジュールの特徴としては各リソースの使用率 1 % 以下を実現していることである。また後段の特徴量変換部、タイプ識別部で DSP ブロックを約 600 個使用し、高速化のためにはさらに必要となる。そのため DSP を使用しない特徴量抽出の構成は本システムに非常に有効であると言える。

プラットフォーム

使用ボード	PROCe IV 530-A (GiDEL 社)
Memory	DDRII オンボード 512MB DDRISODIMM 1GB x 2
使用FPGA	Altera Stratix IV EP4SE530F43C2
FPGA個数	1個
Host Interface	PCIe x4



図 14. ターゲットプラットフォーム。

表 2. リソース使用率。

	SGF	GC	BFC	全体	リソース
ALUT	555	319	360	1,234	424,960
レジスタ	392	392	3,305,191	1,023	424,960
トータルRAM ビット数	65,536	32,768	26,624	124,928	21,233,664
DSP block	0	0	0	0	1,024

8. 入出力インタフェース

実装した D-SIFT の動作確認のため、入出力のインタフェースを提案する。ProcWizard の IP である MegaFIFO を使用しデータの入出力インタフェースを作成した。インタフェース全体図を図 15 に示す [12]。ホスト PC と FPGA ボードの間は PCI Express x4 を使用し、オンボードのメモリを経由してホスト PC と User Interface 間のデータの送受信を行う。図に状態遷移図を記載する。Initial の state で転送するデータをホスト PC で生成し、生成が完了次第 FIFO to IF_mem の state に遷移する。FIFO to IF_memory では、ホスト PC から DMA 転送を行い、PCI Express x4 Bridge, on-chip bus を経由して on-board memory に書き込む。再び on-chip bus を経由して on-chip memory に形成した IF_memory にデータを書き込む。書き込みが完了次第、D-SIFT の state に遷移し、前述のハードウェア向け D-SIFT の処理を行う。D-SIFT の出力である key point の特徴量を 1 つずつ on-chip bus を経由して on-board memory に書き込む。Full HD 画像から抽出される key point の最大数を出力次第、Complete の state に遷移する。Complete ではホスト PC にて全特徴量ブロックを受け取ったことを確認次第、Initial に戻る。入出力インタフェースを含んだ、ハードウェア向け D-SIFT の実装は未実装のため、発表時に実装結果の報告を行う。

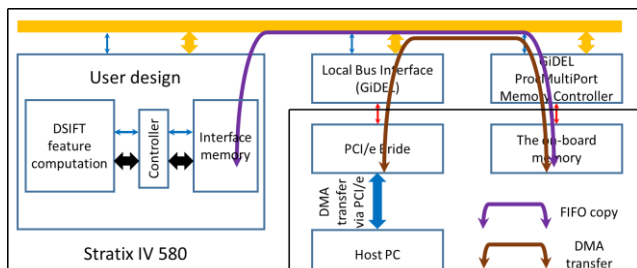


図 15. インタフェース概要.

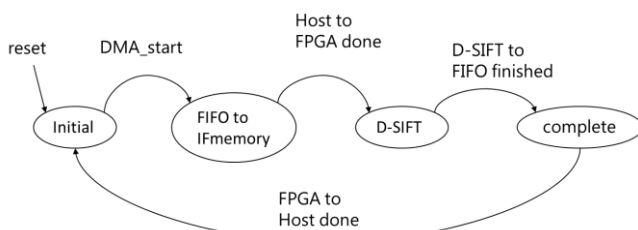


図 22. 状態遷移図.

9. まとめ

本稿ではリアルタイム大腸内視鏡診断支援システムのためのハードウェア向け特徴量抽出アルゴリズムを示し、アーキテクチャの FPGA 実装を行った。Full HD 全画面認識におけるレイテンシ 61 msec @ 100MHz, スループット 16 fps @ 100MHz を達成した。

今後の課題としては、図に示すように本研究グループが開発している特徴量変換ハードウェア [10]とタイプ識

別ハードウェア [9,11]を統合して 1 つの FPGA 上に実装し、システム全体の性能評価を行うことが挙げられる。

謝辞 本研究の一部は、JSPS 科研費基盤研究(C)2459102 と基盤研究(B)26280015 の助成を受けたものです。

参考文献

- 1) H. Kanao, et al., "Narrow-band imaging magnification predicts the histology and invasion depth of colorectal tumors, Journal of Gastrointestinal Endoscopy, vol. 69, no.3, pp. 631-636, 2009.
- 2) 吉牟田 淳基, 他, "大腸拡大内視鏡画像の NBI 拡大所見分類に基づく特徴量と識別器の設計," 電子情報通信学会技術研究報告, Vol. 111, no. 47, pp. 13-18, 2011.
- 3) Toru Tamaki, et al., "Computer-aided colorectal tumor classification in NBI endoscopy using local features," Medical Image Analysis, Vol. 17, No. 1, pp. 78-100, 2013.
- 4) Andrea Vedaldi, Brian Fulkerson, "Vlfeat: an open and portable library of computer vision algorithms," <http://www.vlfeat.org/>, 2012/09/20 accessed.
- 5) Chin-Chung Chang, Chin-Jen Lin, "LIVSVM - a library for support vector machines," <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>, 2012/08/10 accessed.
- 6) 三島 翼, 他, "大腸 NBI 拡大内視鏡画像診断支援システムにおける特徴量抽出部のハードウェア設計," 信学技報, Vol. 112, No. 237, CPSY2012-33, pp. 13-18, 2012 年 10 月 12 日.
- 7) 三島 翼, 他, "大腸 NBI 拡大内視鏡画像診断支援のためのリアルタイム特徴量抽出アーキテクチャ," 電子情報通信学会コンピュータシステム研究会 (CPSY 2013), 2013 年 11 月 8 日.
- 8) S. Shigemi, et al., "An FPGA implementation of support vector machine identifier for colorectal endoscopic images with NBI magnification," Proc. of the 28th International Conference on Circuits / Systems, Computers and Communications (ITC- CSCC2013), pp. 571-572, 2013.
- 9) S. Shigemi, et al., "Customizable hardware architecture of support vector machine in CAD system for colorectal endoscopic images with NBI magnification," Proc. of the 18th Workshop on Synthesis And System Integration of Mixed Information Technologies (SASIMI 2013), R5-2, pp. 298-303, 2013.
- 10) 杉幸樹, 他, "大腸内視鏡診断支援のための高速 Visual Word 特徴量変換の FPGA 実装," 情報処理学会 DA シンポジウム 2015.
- 11) 岡本拓巳, 他, "大腸内視鏡画像のタイプ識別に適した SVM の FPGA 実装," 情報処理学会 DA シンポジウム 2015.
- 12) M. Omori, et al., "HW/SW Co-design of Region Growing Image Segmentation", Proceedings of the 26th International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC2011), Gyeongju, Korea, pp. 322-325, (2011.6).