

平成22年6月1日現在

研究種目：基盤研究（A）
 研究期間：2006～2009
 課題番号：18200003
 研究課題名（和文） 次世代ウェアラブルコンピュータのための知覚認識モバイルプロセッサの垂直統合研究
 研究課題名（英文） Vertical Integration of VLSI mobile processor for Wearable Perception Device
 研究代表者
 吉本 雅彦 (YOSHIMOTO MASAHIKO)
 神戸大学・大学院工学研究科・教授
 研究者番号：30324099

研究成果の概要（和文）：

次世代ウェアラブルコンピュータの知覚認識能力を向上させるための超低消費電力 VLSI プロセッサの設計技術開発を実施した。視覚認識と音声言語認識を統合した VLSI 向き知覚認識アルゴリズムの開発を行うとともにそれらを高い電力効率で実現できる VLSI アーキテクチャを開発した。HDTV 解像度の動画像に対して、実時間で SIFT (Scale Invariant Feature Transform) 特徴量抽出可能なプロセッサについて、従来比で 98.6%の電力削減に成功した。また、20000 語彙の実時間連続音声認識のための VLSI アーキテクチャの開発を完了した。

研究成果の概要（英文）：

The design technology of ultra-low power VLSI processor to improve the perception recognition ability of the next generation wearable computer has been developed. A SIFT Descriptor Generation Engine which features a VLSI oriented Scale Invariant Feature Transform (SIFT) algorithm was realized and it provides high energy efficiency, 2.79mJ / frame, and processing capability for HDTV resolution video (1920 x 1080 pixels) at 30 frames per second (fps). Moreover, the development of VLSI architecture for the real-time continuous speech recognition was completed about 20000 vocabularies.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2006年度	9,400,000	2,820,000	12,220,000
2007年度	11,500,000	3,450,000	14,950,000
2008年度	8,000,000	2,400,000	10,400,000
2009年度	8,000,000	2,400,000	10,400,000
年度			
総計	36,900,000	11,070,000	47,970,000

研究分野：総合領域

科研費の分科・細目：情報学・計算機システム・ネットワーク

キーワード：計算機アーキテクチャ、ウェアラブル、画像認識、音声認識、VLSI、プロセッサ

1. 研究開始当初の背景

VLSI 微細化技術の進展に支えられたコンピュータのダウンサイジングはとどまるところを知らない。今後、高機能なウェアラブルコンピュータやインプラントブルコンピュータ（体内埋め込み型）へと進化する

とともに、本格的な知能ロボットへ搭載されるモバイルスーパーコンピュータの開発が期待される場所である。しかしながら、その実現には、莫大な処理量を必要とする実時間高精細動画像処理と音声言語処理、ネットワーク通信処理などが必要で、モバイル環境

の電力条件のもとで、従来の汎用プロセッサでそれらを実行するのは不可能である。2008年次に予想されたインテルプロセッサの消費電力とモバイルスーパーコンピュータへの要求低消費電力特性との間には実に3桁以上のギャップがあった。このため革新的な高い電力効率のプロセッサ技術の開発が必要となっていた。

2. 研究の目的

本研究の目的は、視覚・音声・言語処理を統合して知覚認識できる、VLSI知覚認識プロセッサコアを超低消費電力性能で実現することである。つまり、デスクトッププロセッサで実現する際に必要な消費電力を大幅に低減し、ウェアラブル・パーセプション・デバイス(Wearable Perception Device:WPD)を実用化するためのプロセッサ技術を開発することである。ここでWPDとは、たとえば“めがね”タイプのウェアラブルコンピュータで、高精細ディスプレイとして働くだけでなく、ユーザや周囲の人や環境を知覚してユーザに適切な情報を提供する機能を有するデバイスを指している。WPDはバーバル(音声・言語)・ノンバーバル(視覚)認識による意図や要求、動機、関心の確実な推定が実施できる。さらにWPD間をネットワークを介して情報を共有させたり、あるいはデータベースアクセスにより学習することで、知覚認識の精度向上をはかることができる。こうしてWPDは「人と環境を理解して人に快適で安全で便利な状況を形成する」機能を有する。

3. 研究の方法

本研究では、視覚認識と音声言語認識を統合したVLSI向き知覚認識アルゴリズムの開発を行う。そしてそれらを高い電力効率で実現できるVLSIアーキテクチャの開発を実施する。電力効率を1~2桁向上させることにより、種々の認識処理を並列に実行でき、かつそれらを統合化することで、より精度の高い、実用的な知覚認識を可能とする。このための開発戦略のキーワードは、並列化、専用化、協調設計の3つである。特に、徹底した垂直統合型協調設計(アルゴリズム階層、アーキテクチャ階層、回路階層)により、電力効率を向上させる。

4. 研究成果

(1) 視覚認識・音声処理アルゴリズム

①視覚認識アルゴリズム

WPDが持つべき視覚認識機能について洗い出しを行い、人に関する認識、物体に関する認識について、研究を行った。

これらの研究成果は、多くのジャーナル、国際会議の予稿集として出版されており、2008年には、物体の切り出しに関して、IEEE ICME 2008 The Best Paper Awardを受賞した。また、2009年には、物体認識に関して、MITA2009 Distinguished Paper Awardを受賞した。

②音声言語処理アルゴリズム

WPDが持つべき音声言語処理機能について洗い出しを行い、音声信号処理、音声認識、音声認識の活用について、研究を行った。これらの研究成果は、多くの国際ジャーナル、国際会議の予稿集として出版されており、2008年には、システムへの指示内容か雑談であるかを判定する手法に関して、情報処理学会から山下記念研究賞を受賞した。また、2009年には、音声の到来方向を推定する手法に関して、日本音響学会よりポスター賞を受賞した。

③視覚認識と音声認識の統合アルゴリズム

上記①と②をどのように統合してウェアラブル・パーセプション・デバイスを構成するかについて研究を行い、階層的統合方法を提案した。また、実世界をWPDで撮影し、撮影された画像から物体を切り出して認識するとともに、音声で指示した検索内容を基に、情報を検索してWPDに提示するアルゴリズムが、実現可能であるかどうかについて研究を行い、実現可能であるとの結論を得た。

(2) VLSIアーキテクチャ

①VLSIメモリアーキテクチャ

大容量メモリは認識プロセッサの高性能化、低消費電力化に極めて重要な要素である。そこで、画像における隣接画素の相関性を利用した2種類の低消費電力SRAMについて研究を行った。一つは多数決論理とReordering処理を用いた2-portSRAMであり、ビット線電力を53%削減することができた。もう一つは読出しビット線に非プリチャージ方式を採用した10トランジスタ構成SRAMであり、読出しビット線電力の8割削減を実現した。

②画像認識プロセッサアーキテクチャ

前述の並列処理アルゴリズムに対応するVLSIアーキテクチャを開発した。HDTV解像度の動画像に対して、実時間、低消費電力で、SIFT(Scale Invariant Feature Transform)特徴量抽出を実行可能なプロセッサアーキテクチャを開発した。VLSI向きSIFTアルゴリズム、3ステージROI(Region of Interest)パイプラインアーキテクチャ、SIFT特徴量抽

出に特化した専用演算回路の実装により、高い電力効率を実現する VLSI を設計し、試作 (65nmCMOS) を完了し (図 1)、従来技術に対して 98.6%の電力削減を達成した (図②)。また画像認識における動き特徴量としてのオプティカルフロープロセッサ、および画像圧縮符号化プロセッサのアーキテクチャを考案し、FPGA実装を行った。

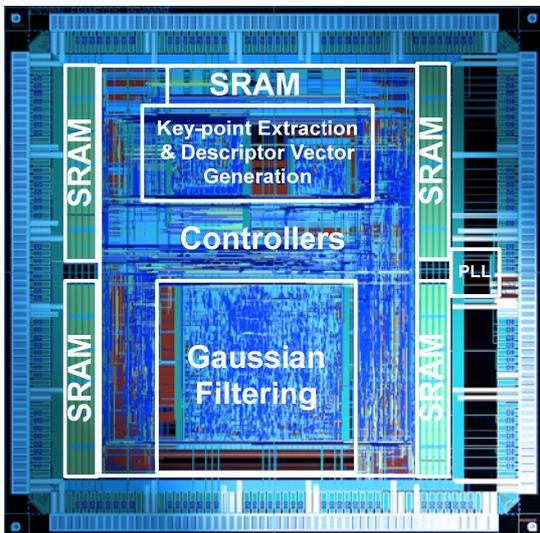


図 1. SIFT 特徴量抽出 LSI のチップ写真

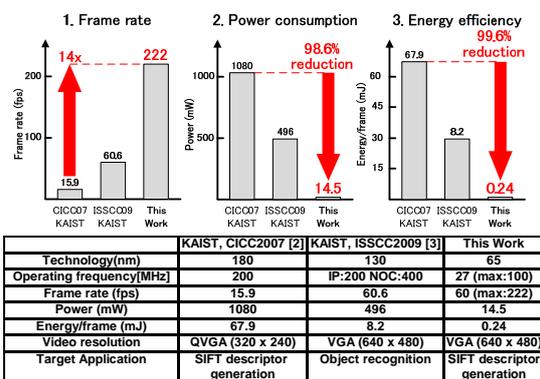


図 2. SIFT 特徴量抽出 LSI の性能改善効果

③音声認識プロセッサアーキテクチャ

20000 語彙の実時間連続音声認識のための VLSI アーキテクチャの開発を完了した。従来 LSI アーキテクチャでは、膨大な演算量とメモリ帯域が問題となり、2 万語彙を超える認識は困難だったが、提案アーキテクチャでは、GMM 演算の高並列化、Viterbi 演算アルゴリズム改良、キャッシュの導入、GMM 演算と Viterbi 演算の 2 ステージパイプラインの導入を行なった。そして提案アーキテクチャを FPGA に実装することで、従来アーキテクチャに比べリアルタイム動作時の必要周波数を、20,000 語で 32%削減し 41.71 [MHz]での動作を確認した。またメモリ帯域は 65%削減し、

218.62 [MB/s]でのリアルタイム認識が可能であった。その際の認識精度は、86.42%であった。

以上により、画像認識と音声認識プロセッサコア技術が開発でき、統合認識プロセッサ VLSI 実現のメドを得ることができた (図 3)。

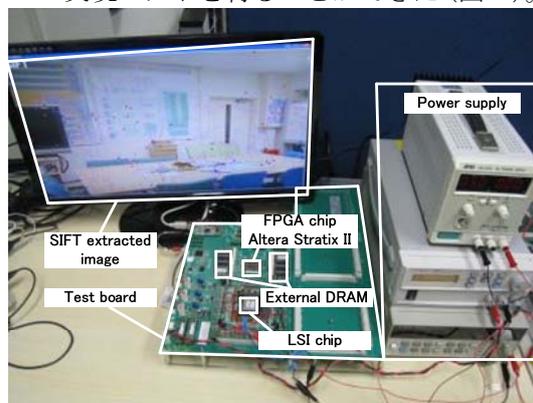


図 3. 画像・音声知覚認識 LSI 評価ボード

(3) 知覚認識システム開発

コンピュータのデスクトップを常時画像処理し、決まったパターンに対してアプリケーション動作を行うルールベースエンジンを構築した。図 4 に作成したウェアラブルアプリケーションの処理の流れを示す。PC の画面やカメラ画像中にある特定の画像を認識し、その認識パターンによりイベントの種類を判別して処理に反映させるという流れを繰り返す。また、視覚以外の知覚情報の活用を考慮し、音声認識を導入して音声もトリガとすることができる。認識手法はテンプレートとなる画像毎に設定することが可能で、主に PC 画面上はテンプレートマッチングを、カメラ画像中の検索は SIFT 特徴量を用いた認識を行う。この部分に関し、上記 LSI と組み合わせで高速動作を実現することで、実用的な実世界アプリケーションの構築が可能になる。またパターン認識については、各テンプレートの時系列における相対的な生起関係を指定し、複数の画像が連動して発生する場合のイベント認識にも対応する。図 5 に作成したウェアラブルアプリケーションの動作画面例を示す。本システムを実際のウェアラブル環境で利用し、ウェアラブルカメラと HMD と組み合わせることで、音声・画像の統合知覚認識処理により、実空間内でのアノテーションや実空間連動型のサービスの構築が容易になる。

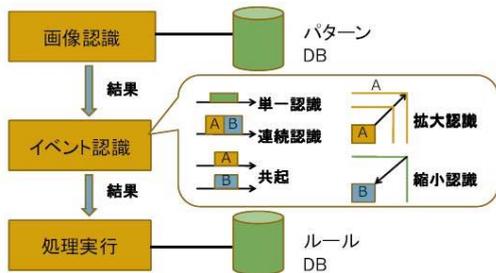


図4. ウェアラブルアプリケーション処理の流れ



図5. 作成したウェアブルアプリケーションの動作画面例

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計11件)

- ① Ryoichi Takashima, Tetsuya Takiguchi, Yasuo Ariki, Monaural sound-source-direction estimation using the acoustic transfer function of a parabolic reflection board”, Journal of the Acoustical Society of America, 査読有, Volume 127, Issue 2, 2010, pp. 902-908
- ② H. Fujiwara, S. Okumura, Y. Iguchi, H. Noguchi, H. Kawaguchi, and M. Yoshimoto, A Dependable SRAM with 7T/14T Memory Cells, IEICE Trans. Electron, 査読有, 2009
- ③ Takuya Tonaru, Tetsuya Takiguchi and Yasuo Ariki, Extraction of Human Activities as Action Sequences Using pLSA and PrefixSpan, International Journal of Hybrid Information Technology, 査読有, Vol. 2, No. 1, 2009, pp. 13-20
- ④ 佐古淳, 滝口哲也, 有木康雄, 音声・状況の同時認識に基づくスポーツ実況中継へのメタ情報付与, 情報処理学会論文誌: 音声ドキュメント処理特集号, 査読有,

Vol. 50 No. 2, 2009, pp. 563-574

- ⑤ H. Fujiwara, S. Okumura, Y. Iguchi, H. Noguchi, H. Kawaguchi, and M. Yoshimoto, A Dependable SRAM with 7T/14T Memory Cells, IEICE Transactions on Electronics, 査読有, vol. E92-C, no. 4, 2009, pp. 423-432
 - ⑥ Y. Murachi, J. Miyakoshi, M. Hamamoto, T. Iinuma, T. Ishihara, F. Yin, J. Lee, H. Kawaguchi, and M. Yoshimoto, A sub 100 mW H.264 MP@L4.1 Integer-pel Motion Estimation Processor Core for MBAFF Encoding with Reconfigurable Ring-connected Systolic Array and Segmentation-free, Rectangle-access Search-window Buffer, IEICE Trans. Electron, 査読有, 2008, pp. 465-478
 - ⑦ Y. Murachi, Y. Fukuyama, R. Yamamoto, J. Miyakoshi, H. Kawaguchi, H. Ishihara, M. Miyama, Y. Matsuda and M. Yoshimoto, A VGA 30-fps Realtime Optical-Flow Processor Core for Moving Picture Recognition, IEICE Trans. Electron., 査読有, 2008, pp. 457-464
 - ⑧ H. Noguchi, Y. Iguchi, H. Fujiwara, S. Okumura, Y. Morita, K. Nii, H. Kawaguchi, and M. Yoshimoto, A 10T Non-Precharge Two-Port SRAM Reducing Readout Power for Video Processing, IEICE Trans. Electron, 査読有, vol. E91-C, no. 4, 2008, pp. 543-552
 - ⑨ H. Fujiwara, K. Nii, H. Noguchi, J. Miyakoshi, Y. Murachi, Y. Morita, H. Kawaguchi, and Masahiko Yoshimoto, Novel Video Memory Reduces 45% of Bitline Power using Majority Logic and Data-Bit Reordering, IEEE Transactions on Very Large Scale Integration (VLSI) Systems, 査読有, vol. 16, no. 6, 2008, pp. 620-627
 - ⑩ 宮原正典, 青木政樹, 滝口哲也, 有木康雄, 顔表情からの関心度推定に基づく映像コンテンツへのタギング, 情報処理学会論文誌, 査読有, Vol. 49, No. 10, 2008, pp. 3694-3702
 - ⑪ 田中克幸, 滝口哲也, 有木康雄, AdaBoost/LogitBoostによるWhyテキストセグメント判定と回答抽出の自動化, 情報処理学会論文誌, 査読有, Vol. 49, No. 6, 2008, pp. 2234-2242
- [学会発表] (計19件)
- ① T. Fujinaga, K. Miura, H. Noguchi, H. Kawaguchi, and M. Yoshimoto, Parallelized Viterbi Processor for 5,000-Word Large-Vocabulary Real-Time Continuous Speech Recognition FPGA System, Proceedings of ISCA Annual

- Conference of International Speech Communication Association (Interspeech), 2009/9, London, UK
- ② 栗田雄介, 寺田努, 塚本昌彦, 画像認識を基盤としたアプリケーション制御フレームワークの設計と実装, 情報処理学会マルチメディア, 分散, 協調とモバイルシンポジウム(DICOMO 2009), 2009年7月9日, 大分県別府温泉 杉乃井ホテル
- ③ S. Okumura, A 0.56-V 128kb 10T SRAM Using Column Line Assist (CLA) Scheme, International Symposium on Quality Electronic Design, 2009, 2009年3月17日, San Jose
- ④ H. Fujiwara, A 7T/14T Dependable SRAM and Its Array Structure to Avoid Half Selection, International Conference on VLSI Design 2009, 2009年1月6日, New Delhi, India
- ⑤ K. Miura, A Low Memory Bandwidth Gaussian Mixture Model (GMM) Processor for 20,000-Word Real-Time Speech Recognition FPGA System, International Conference on Field-Programmable Technology 2008, 2008年12月9日, Taipei, TAIWAN
- ⑥ Kosuke Mizuno, An H.264/AVC MP@L4.1 Quarter-Pel Motion Estimation Processor VLSI for Real-Time MBAFF Encoding, IEEE International Conference on Electronics, Circuits, and Systems (ICECS 2008), 2008年9月3日, Malta
- ⑦ Yuichiro Murachi, A Sub 100 mW H.264/AVC MP@L4.1 Integer-Pel Motion Estimation Processor VLSI for MBAFF Encoding, International Symposium on Circuits and Systems (ISCAS), 2008年5月19日, Seattle, Washington, USA
- ⑧ Y. Murachi, A Power-Efficient SRAM Core Architecture with Segmentation-Free and Rectangular Accessibility for Super-Parallel Video Processing, 2008 International Symposium on VLSI Design, Automation & Test (VLSI-DAT), 2008 International Symposium on VLSI Design, Automation & Test (VLSI-DAT), 2008年4月24日, Hsinchu, Taiwan,
- ⑨ Katsunori Onishi, 3D Human Posture Estimation Using the HOG Features from Monocular Image, Proc. of 19th Int'l Conf. on Pattern recognition (ICPR08), 2008, Tampa, Florida USA
- ⑩ Akira Suga, Object Recognition and Segmentation Using SIFT and Graph Cuts, Proc. of 19th Int'l Conf. on Pattern recognition (ICPR08), 2008, Tampa, Florida USA
- ⑪ Keita Fukuda, GRAPH CUTS BY USING LOCAL TEXTURE FEATURES OF WAVELET COEFFICIENT FOR IMAGE SEGMENTATION, ICME 2008 (International Conference on Multimedia and Expo), 2008, Hannover, Germany
- ⑫ T. Ishihara, Y. Murachi, T. Iinuma, F. Yin, T. Kamino, K. Mizuno, H. Kawaguchi, and M. Yoshimoto, A Sub-100mW H.264 MP@L4.1 Integer-Pel Motion Estimation Processor Core for Real-time MBAFF Encoding, International Solid-State Circuits Conference Student Forum, 査読有, Feb. 2008
- ⑬ Atsushi Sako, Language Modeling Using PLSA-Based Topic HMM, IEICE TRANSACTIONS on Information and Systems, 査読有, Vol. E91-D, No. 3, 2008, pp. 522-528
- ⑭ H. Noguchi, A 10T Non-Precharge Two-Port SRAM for 74% Power Reduction in Video Processing, Proceedings of IEEE Computer Society Annual Symposium on VLSI (ISVLSI), 2007
- ⑮ Y. Sakata, K. Kawakami, H. Kawaguchi, and M. Yoshimoto, An Elastic Pipeline Architecture for Dynamic Voltage Scaling and Its Application to Low-Power Portable H.264/AVC Decoder with Embedded Frame Buffer SRAM, Proceedings of WSEAS European Computing Conference (ECC), 査読有, Sep. 2007
- ⑯ H. Ishihara, M. Miyama, Y. Matsuda, Y. Murachi, Y. Fukuyama, R. Yamamoto, J. Miyakoshi, H. Kawaguchi, M. Yoshimoto, A VGA 30-fps Optical-Flow Processor Core Based on Pyramidal Lucas and Kanade Algorithm, IEEE Asian Solid-State Circuits Conference Proceedings of Technical Papers (A-SSCC), 査読有, Nov. 2007
- ⑰ Tetsuya Takiguchi, PCA-Based Speech Enhancement for Distorted Speech Recognition, Journal of Multimedia, 査読有, Volume 2, Issue 5, 2007, pp. 13-18
- ⑱ Tomoyuki Yamagata, System Request Detection in Conversation Based on Acoustic and Speaker Alternation Features, Interspeech2007, 査読有, 2007, pp. 2789-2792
- ⑲ H. Fujiwara A Two-Port SRAM for Real-Time Video Processor Saving 53% of Bitline Power with Majority Logic and Data-Bit Reordering, Proceedings of ACM/IEEE International Symposium on Low Power Electronics and Design (ISLPED), 2006, pp. 61-66

[図書] (計0件)

〔産業財産権〕

○出願状況（計7件）

名称：半導体メモリおよびプログラム
発明者：吉本雅彦，川口博，藤原英弘，
奥村俊介

権利者：神戸大学

種類：特許

番号：特願 2008-000357

出願年月日：2008年1月7日

国内外の別：国内

名称：水平、垂直 1/2 画素間引きが可能な任意位置任意サイズブロックアクセスメモリ

発明者：吉本雅彦，川口博，宮越純一，
村地勇一郎

権利者：神戸大学

種類：特許

番号：特願 2007-298743

出願年月日：2007年11月18日

国内外の別：国内

名称：画像処理装置及び方法
発明者：吉本雅彦，村地勇一郎，川口博，
福山祐貴，山本亮，松田吉雄，
深山正幸

権利者：半導体理工学研究センター

種類：特許

番号：特願 2007-290754

出願年月日：2007年11月15日

国内外の別：国内

名称：IMAGE ENCODING OF MOVING PICTURES
発明者：Junichi Miyakoshi, Masayuki Miyama,
Masahiko Yoshimoto, Hideo Hashimoto,
and Imamura Kousuke

権利者：半導体理工学研究センター

種類：特許

番号：US Patent No. :US7, 236, 634 B2, Date
of Patent

出願年月日：Jun. 26, 2007

国内外の別：国外

名称：半導体装置，
発明者：藤原英弘，宮越純一，村地勇一郎，
川口博，吉本雅彦，新居浩二

権利者：神戸大学

種類：特許

番号：特願 2006-237637

出願年月日：2006年9月1日

国内外の別：国内

名称：画像処理用メモリ
発明者：宮越純一，吉本雅彦，村地勇一郎，
濱本真生，松野哲郎

権利者：神戸大学

種類：特許

番号：特願 2006-203542

出願年月日：2006年7月26日

国内外の別：国内

名称：画像処理プロセッサ

発明者：宮越純一，吉本雅彦，村地勇一郎，
松野哲郎，濱本真生

権利者：メガチップス

種類：特許

番号：特願 2006-203541

出願年月日：2006年7月26日

国内外の別：国内

○取得状況（計0件）

〔その他〕

ホームページ等

6. 研究組織

(1) 研究代表者

吉本 雅彦 (YOSHIMOTO MASAHIKO)
神戸大学・大学院工学研究科・教授
研究者番号：30324099

(2) 研究分担者

塚本 昌彦 (TUKAMOTO MASAHIKO)
神戸大学・大学院工学研究科・教授
研究者番号：60273588

有木 康雄 (ARIKI YASUO)
神戸大学・都市安全研究センター
研究者番号：10135519

滝口 哲也 (TAKIGUCHI TETUYA)
神戸大学・大学院工学研究科・講師
研究者番号：40397815

川口 博 (KAWAGUCHI HIROSI)
神戸大学・大学院工学研究科・准教授
研究者番号：00361642

(3) 連携研究者

なし。