

メディアストリーミングプロセッサ SpursEngine™とその応用例

SpursEngineSE1000 High-Performance Stream Processor

檜田 和浩 近藤 伸宏

■ HIWADA Kazuhiro

■ KONDOH Nobuhiro

東芝は、CPU及びGPU (Graphics Processing Unit) に続く第3のプロセッサであるメディアストリーミングプロセッサ SpursEngine™を開発した。

SpursEngine™は、Cell Broadband Engine™(注1)の演算コアであるSPE (Synergistic Processing Element) を利用したシステムLSIであり、メディアストリーミング処理において特に高い性能対消費電力を実現するプロセッサである。

また、SpursEngine™の特長の一つである高い画像処理性能を生かしたアプリケーションを開発し、従来はCell Broadband Engine™などの高性能プロセッサを必要とした高精度リアルタイムフェイストラッキングを、市販のパソコン(PC)とSpursEngine™の組合せにより実現した。

Toshiba has developed SpursEngineSE1000, a third processor operating in cooperation with a central processor unit (CPU) and a graphics processing unit (GPU), which utilizes multiple synergistic processor element (SPE) cores derived from the Cell Broadband Engine™.

SpursEngineSE1000 has been designed to achieve a good power-performance ratio, especially in the area of streaming-media processing. Utilizing the advanced image-processing performance of SpursEngineSE1000, we have also developed "FACEMATION", an image manipulation program that realizes high-performance real-time 3D face tracking on a consumer PC in conjunction with SpursEngineSE1000.

1 まえがき

東芝は、Cell Broadband Engine™(1) (以下、Cell/B.E.™(注1)と記す)の演算コアであるSPE (Synergistic Processing Element) を継承したシステムLSI SpursEngine™を開発した(図1)。ここでは、SpursEngine™の概要を述べるとともに、そのアプリケーションの一つであるFACEMATIONの紹介を通してSpursEngine™の特長を述べる。

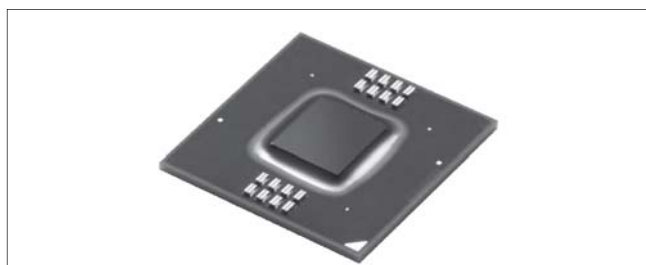


図1. SpursEngine™ — SpursEngine™のチップである。Cell/B.E.™の演算コアであるSPEを継承している。
SpursEngineSE1000

(注1) IBM社、SONYグループ、東芝が共同開発した高性能プロセッサで、Cell Broadband Engine及びCell/B.E.は、(株)ソニー・コンピュータエンタテインメントの商標。

2 SpursEngine™開発の背景

近年、高精細な映像などHD (High Definition) コンテンツの普及が急速に進んでいる。わが国では、2003年12月に地上デジタル放送が開始され、地上波でHDコンテンツが放映されるようになった。それと同期して、液晶テレビ(TV)やプラズマTVに代表される大画面、薄型TVが広まり、家庭でHDコンテンツを楽しむことができるようになった。

その後、HDD (ハードディスク装置) レコーダの普及や大容量記録が可能な光ディスクの登場により、ユーザーがHDコンテンツを記録できるようになった。更に、HD映像が記録できるデジタルビデオカメラの登場で、HDコンテンツの撮影が可能になった。インターネットのプロードバンド(高速大容量通信)化も急速に進み、動画投稿サイトやTV電話など、映像を利用した新しいサービスが開始され人気を集めている。

ユーザーが映像コンテンツを扱う機器の一つにPCが挙げられる。PCは、高性能なCPUの搭載と汎用性の高さから様々な用途に活用され、家庭向けにはTVチューナを内蔵したAV PCというジャンルが確立されている(図2)。AV PCでは、映像や音声などのメディアデータを扱うが、映像コンテンツのHD化とともに、汎用CPUでの処理速度の不足が顕著になってきている。今後、HDコンテンツを扱うPCでは、それを処理するための専用プロセッサが必要になると予想される。

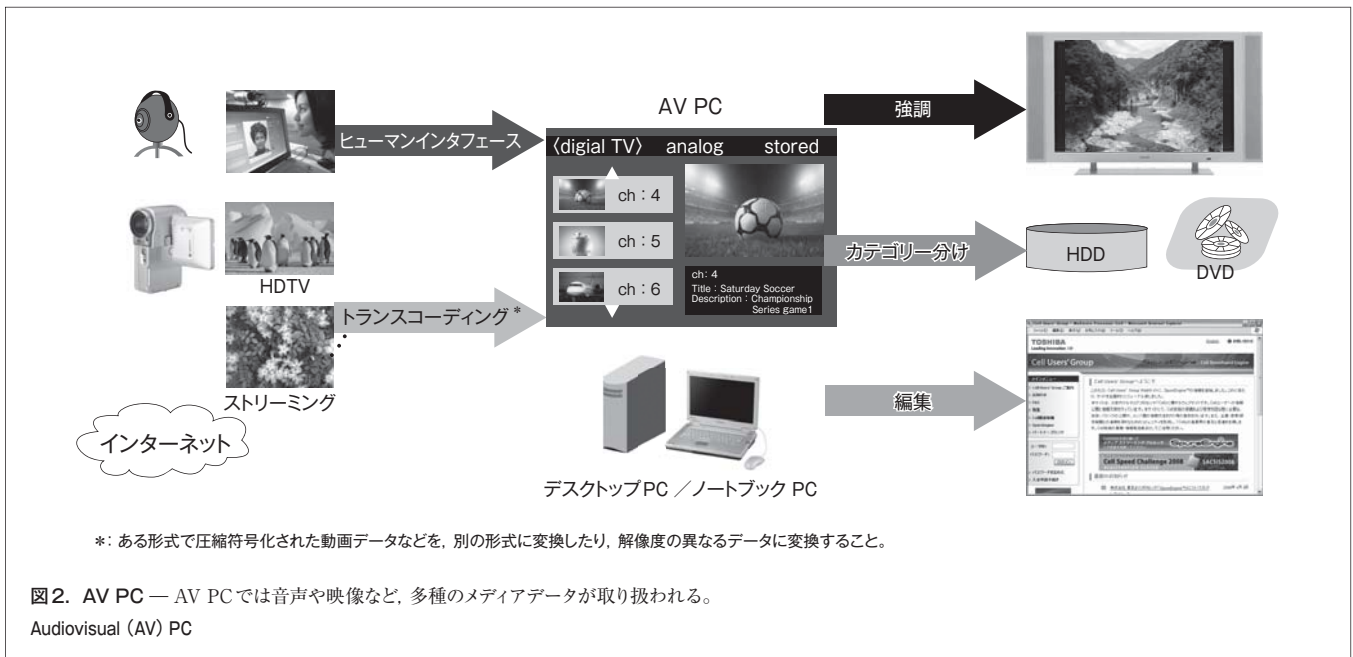


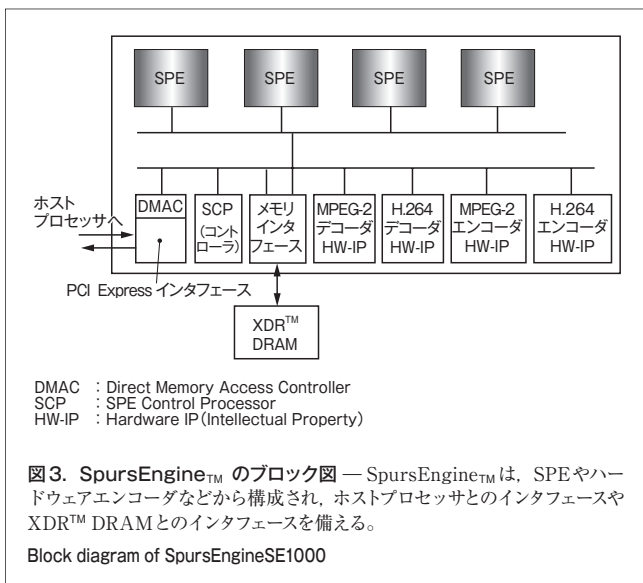
図2. AV PC — AV PCでは音声や映像など、多種のメディアデータが取り扱われる。
Audiovisual (AV) PC

SpursEngine™は、このような背景の下で、ユーザーがHDコンテンツを快適に楽しむことができるような機器の実現を目指して開発された。すなわち、ユーザーが記録したHDコンテンツを快適に編集し、インターネット経由で流通させたり、携帯端末で閲覧したりできるような環境の提供である。

3 SpursEngine™のアーキテクチャ

SpursEngine™のブロック図を図3に示す。

SpursEngine™は、MPEG-2 (Moving Picture Experts Group-Phase 2) 及びH.264^(注2)に準拠したHD対応のエンコーダ (符号化を行う回路)とデコーダ (復号を行う回路)、及



びSPEと呼ばれる演算コアを4個搭載している。SPEは、Cell/B.E.™の演算コアと共通であり、メディア処理性能が高く、性能対消費電力が高いことを特長としている。SpursEngine™では性能と消費電力のトレードオフに対し、エンコード、デコードなどの定型処理はハードウェアで処理し、それ以外の非定型処理はSPEによるソフトウェアで処理するという解決法を採用した。

ホストシステムと接続するための外部インタフェースとしては、PCI Express^(注3)をサポートしている。また、ワーキングメモリとしてXDR™^(注4) DRAMを採用しており、メモリバンド幅を必要とするアプリケーションの実行が可能である。

SpursEngine™はHD対応のエンコーダ、デコーダ、及び高性能マイクロプロセッサが1チップに搭載されており、これらを密に連携させることで、より豊かな体験をユーザーに提供するアプリケーションを実現できる。典型的な使い方として、ホストシステムから入力された映像データをデコーダでワーキングメモリに展開し、SPEで画像認識処理を施し、その結果とともに、ユーザーが必要とするエンコードを行った映像データをホストシステムに出力することを想定している。このような使い方をするすることで、ホストシステム側に負荷をかけることなくHDコンテンツをリアルタイムに処理しながら、高度な画像認識処理によって新しいアプリケーションが創造されることが期待される。

(注2) ITU (国際電気通信連合電気通信標準化部門)とISO (国際標準化機構)が2003年に勧告した動画圧縮符号化の国際標準規格。

(注3) 2002年にPCI SIG (Peripheral Component Interconnect Special Interest Group)によって策定されたI/O (Input/Output) シリアルインタフェース。

(注4) XDR DRAMは、米国及びその他の国におけるRambus社の登録商標。

4 SpursEngine™のアプリケーション

当社が開発したSpursEngine™の適用によるデモアプリケーション FACEMATION について述べる。FACEMATION は、ユーザーの顔を映した入力映像に対し、仮想的に髪型を変更したり化粧を施したりするアプリケーションである(図4)。このアプリケーションでは、フェイストラッキングという画像処理を行っているが、高精度のフェイストラッキングをリアルタイムで実行するには高い処理性能が必要である。当社は、2005年に高性能プロセッサであるCell/B.E.™を用いることでリアルタイム化し、今回、市販のPCとSpursEngine™を組み合わせたシステム上にこのアプリケーションを実装した。

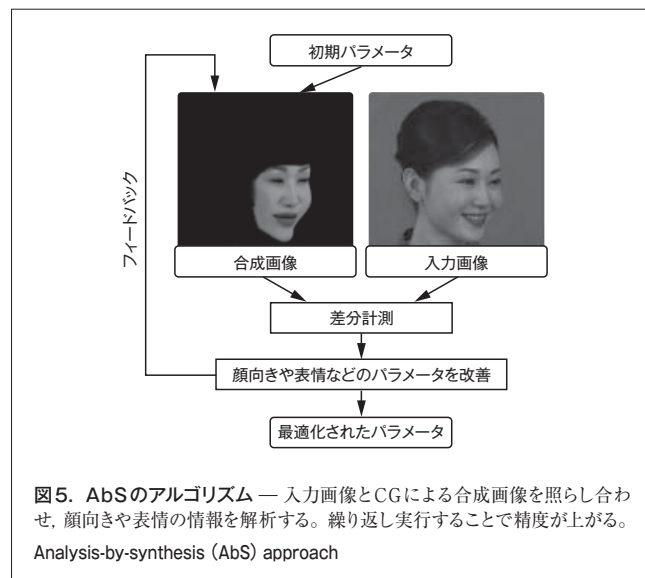


4.1 フェイストラッキング アルゴリズム

フェイストラッキングは、コンピュータビジョン(コンピュータによる視覚情報処理)分野において広く研究されている。当社はその中でも、実写のCG(Computer Graphics)の合成という高精度を必要とする場面に有効なAnalysis-by-Synthesis(以下、AbSと略記)という手法をベースに、リアルタイム向けに改良を加えた手法を用いている⁽²⁾。

AbS手法は、入力画像を模した画像を合成しつつ、その合成画像と入力画像を照らし合わせることで、詳細な解析を行う手法である(図5)。AbS手法では、まず顔モデルを用いて利用者の画像を合成する。次にこの合成画像と入力画像を比較し、それらの差分を計測する。この差分を最小化するように顔の姿勢や表情などのパラメータを最適化することで、より入力画像に近い利用者の画像を合成する。より入力画像に近い合成画像を用いることで、より詳細な最適化が可能になるため、合成、比較、及び最適化の手順を繰り返すことにより、高精度の解析ができる。

AbSにおける主な処理は、合成画像の生成、画像間の差分



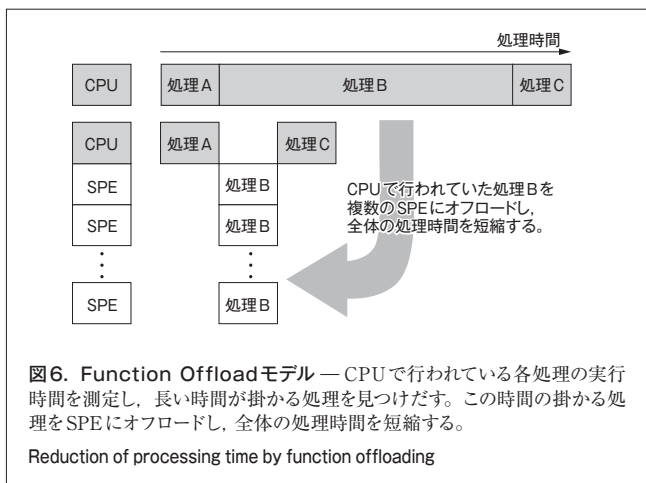
計測、及びパラメータ最適化の三つである。この中でもっとも時間を要する処理が画像間の差分計測処理であり、この処理をいかに高速に行うかが、リアルタイム処理を実現するうえで重要である。

4.2 SpursEngine™による高速化

AbS手法は精度が高い反面、処理は非常に重く、SpursEngine™などの高性能プロセッサの能力を十分に引き出すことが、リアルタイム処理を実現するうえで必須である。当社は、市販のPCにAbS手法用いたFACEMATIONを実装し、結果として10フレーム/sという結果を得た。しかし、自然な動画画像を生成するには30フレーム/s以上処理できることが望ましく、市販のPCに比べ3倍程度の高速化が必要であった。

SpursEngine™などのプログラマブルなプロセッサを用いてプログラムを高速化する手法は多種あるが、当社は、SpursEngine™やCell/B.E.™などの異種プロセッサ間の並列処理による高速化に有効な手法であるFunction Offloadという開発モデルを用いた⁽³⁾(図6)。Function Offloadに基づく開発では、まずメインプロセッサですべての処理を実装したうえで各処理に要する時間を測定し、処理時間が長い処理を見つけだす。この時間の掛かる処理を複数のSPEに担わせる(オフロードする)ことで、全体の処理を高速化する。SPEがオフロードされた処理を実行している間、メインプロセッサはほかの処理を実行できるため、更にシステム全体の性能向上に寄与できる。

当社は、Function Offloadモデルを用い、AbS手法におけるもっとも重い処理である画像間の差分計測を、SpursEngine™の持つ四つのSPEにオフロードした。画像間の差分計測処理には、テンプレートマッチングというアルゴリズムを用いた。テンプレートマッチングは、合成画像上の特徴点の周りの部分画像に着目し、その部分画像にもっともよく似た入力画像上の部



分画像を見つけるという処理である。実装では部分画像は8画素×8画素とし、部分画像が似ていることの指標にはZSSD (Zero-Mean Sum of Squared Difference) 関数を用いた。

SPEは、SIMD (Single Instruction Multiple Data) 演算器を持ち16バイト単位のデータを一度に処理できる。また、画像処理に向けた命令を多数実行できるため、テンプレート

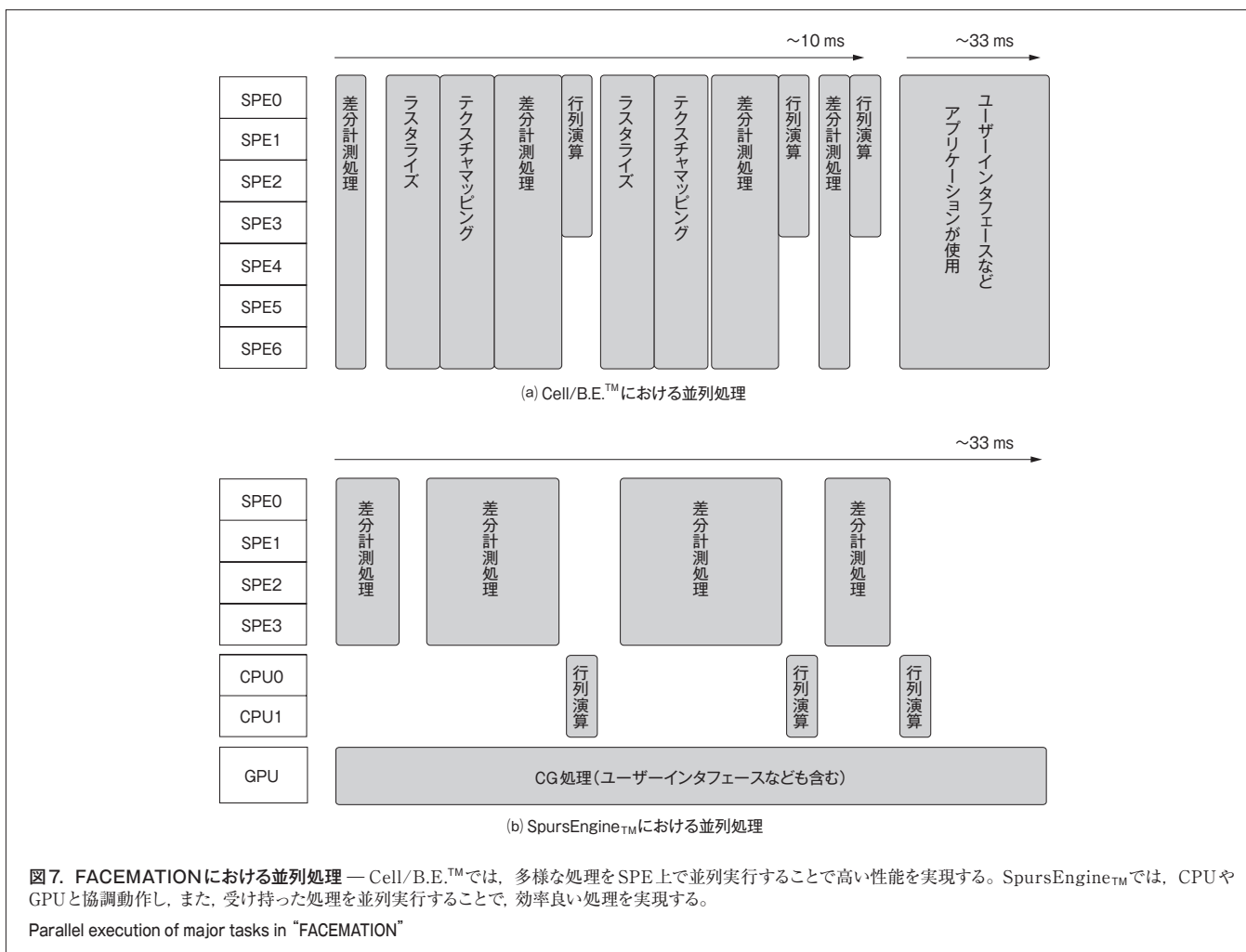
マッチングを通常の4倍から16倍という高い効率で実装することができた。更に、多数の特徴点についてテンプレートマッチングを行う必要があるため、特徴点を複数のSPEに割り分けることで効率良い並列化を行えた。

4.3 第3のプロセッサとしてのSpursEngine™

当社は以前、Cell/B.E.™にフェイストラッキングを実装した際にもFunction Offload手法を用い、SPEにより高速化した。Cell/B.E.™とSpursEngine™それぞれの並列動作のようすを図7に示す。

Cell/B.E.™への実装時には、画像間の差分計測だけでなく合成画像の生成やパラメータ最適化も含め、すべての処理をCell/B.E.™上でやっている。Cell/B.E.™は非常に高性能なプロセッサであるため、このような多様な処理を行いながらもリアルタイム動作に成功した。

一方、SpursEngine™は性能対消費電力を高めたプロセッサであり、Cell/B.E.™に比べ消費電力が大幅に少ないが、SPEのピーク性能はCell/B.E.™に及ばない。しかし、CPUやGPUなどの異種プロセッサ間で処理を適切に分割することで、システム全体としてはCell/B.E.™と同等のリアルタ



ム動作を実現できた。

CPUは汎用プロセッサであり、多様な処理を効率良く実現できるよう進化しているが、GPUなどの特定用途LSIには性能や消費電力の面で及ばない。GPUはグラフィクス処理に特化することで、高い処理性能を実現するとともに、性能対消費電力の面でも汎用プロセッサに比べ優れており、結果としてシステム全体の低消費電力化に寄与している。

一方、SpursEngine™はCPUやGPUとは異なる特徴を持つプロセッサであり、動画のエンコードやデコード、及びSPEの得意とする画像処理などのメディア処理において、高い性能対消費電力を実現する。このように、異種プロセッサがそれぞれの得意とする処理を分担し実行することで、システム全体として高い性能を、限られた消費電力の範囲内で実現することができる。

SpursEngine™を用いた場合と用いなかった場合の、FACEMATIONの性能比較を表1に示す。SpursEngine™を用いることで、処理速度が3倍に向上するとともに、Function Offloadを用いた結果、CPU負荷率は90%から40%に下がっている。システムの性能が向上することでリアルタイム化が実現でき、CPUの負荷が低下することで余ったパワーを用いてほかの処理を行うこともできる。

表1. SpursEngine™によるFACEMATIONの性能向上
Reduction of "FACEMATION" processing time by SpursEngineSE1000

項目	性能	
	フレームレート (フレーム/s)	CPU負荷率 (%)
CPU (2コア) + GPU	10	90
CPU (2コア) + GPU + SpursEngine™	30	40

5 あとがき

当社は、CPU及びGPUに続く第3のメディアストリーミングプロセッサであるSpursEngine™を開発した。ここではSpursEngine™の概要を述べるとともに、そのアプリケーションの一つであるFACEMATIONを例に、高速化の事例や性能、及びSpursEngine™の特長を説明した。

SpursEngine™はCell/B.E.™に由来するSPEを持っており、Cell/B.E.™の持つソフトウェア資産を活用できるという特長がある一方で、LSIとしての位置づけはCell/B.E.™とは異なる。すなわち、CPU及びGPUと協調し、効率良く処理を分担することでシステム全体の効率を高め、ひいてはAV PCなどでの快適な環境を提供する。

今後も、デジタル家電向けの高性能半導体には、性能向上と低消費電力化というトレードオフの解決が求められる。当社は、より高い性能と低消費電力を目指し、後継製品の開発を進めていく予定である。

文献

- (1) 齊藤智隆, ほか. "特集: Cellからの始まり". 東芝レビュー. 61, 6, 2006, p.1-59.
- (2) Hiwada, K., et al. "Mimicking Video: Real-Time Morphable 3D Model Fitting". Proc 10th ACM Symposium on Virtual Reality Software and Technology. Osaka, 2003-10, ACM, New York, ACM Press, p.132-139.
- (3) Kahle, J.A., et al. "Introduction to the cell multiprocessor". IBM Journal of Research and Development. 49, 4/5, 2005, p.589-604.



檜田 和浩 HIWADA Kazuhiro

セミコンダクター社 システムLSI事業部 先端SoC応用技術部 主務。並列処理システム、画像処理システムの研究・開発に従事。

System LSI Div.



近藤 伸宏 KONDOH Nobuhiro

セミコンダクター社 システムLSI事業部 先端SoC応用技術部 主務。並列処理システム、画像処理システムの研究・開発に従事。

System LSI Div.