

身近になった先端デジタルメディア処理技術の動向

Trends in Advanced Digital Media Processing in Familiar Digital Devices

山内 康晋 土井 美和子

■ YAMAUCHI Yasunobu

■ DOI Miwako

パソコン (PC) やテレビ (TV), 携帯電話など身近に存在するデジタル機器の性能は年々向上しており, デジタル機器を通して触れる映像や音声, 知識といったメディア情報に接することで, 豊かで便利な生活を享受することができるようになった。一方で, ブロードバンド インターネット (高速大容量通信インターネット) の普及により, デジタル機器の扱えるデータ量が飛躍的に増加したことで, 真に求めている情報にアクセスできないという問題も発生している。

デジタル機器の性能向上は人間の処理能力を拡張してきた反面, 人間の処理能力を超えたメディア情報の扱いと人間への円滑な情報伝達が課題となっている。機器性能の向上により大量のデジタルメディア情報を容易に取得して処理できるこれからの時代には, 人間の特性を考慮したデジタルメディア処理技術の開発と活用が求められている。

東芝は, 感性に訴える高品位な表現, 自然で直感的なインタラクション (やり取り) の実現, そして時間や興味, 人などの“つながり”に基づく情報フィルタリングといった方向性で, 人にやさしいデジタルメディア処理技術を開発している。

The performance of digital media devices such as PCs, TVs, and mobile phones has rapidly progressed in recent years, and people now depend on these devices for convenience and comfort in life. However, the increasing volume of Internet data, which can now be easily acquired via a broadband network, also creates problems for users in terms of accessing the data that they actually need.

Although people's ability to handle large volumes of media data can be augmented by means of digital devices, it is difficult for users to access most of the data they need due to the gap between the capabilities of human users and digital devices. Human-friendly digital media processing is required to compensate for this gap.

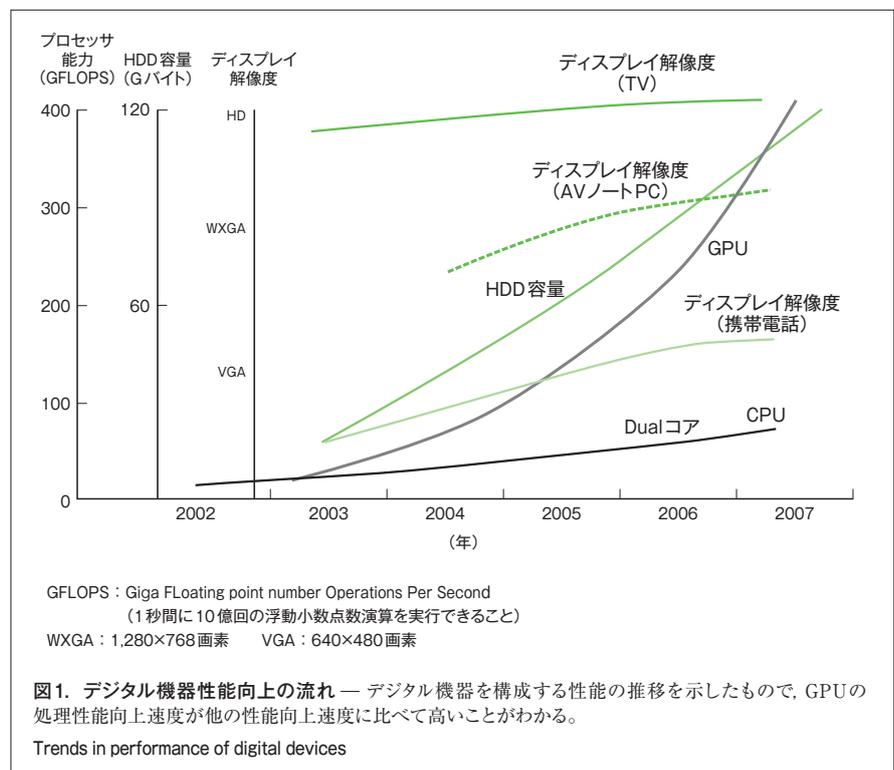
Toshiba has created various advanced media processing methods and digital devices focusing on high-quality media representation that appeals to people's sensibilities, natural and intuitive human-machine interaction, and data filtering using relational information.

人間の処理能力を拡張する デジタル機器の性能向上

われわれの身近に存在するTV, PC, 及び携帯電話といったデジタル機器は, 臨場感あふれる大量の映像コンテンツの視聴や, インターネットを介したグローバルで多様な情報へのアクセス, コミュニケーションなどの場面で, 家庭の内外を問わず, われわれの日常生活になくてはならないものとなってきた。このようなデジタル機器が日常生活で急速に利用されるようになった背景には, デジタル機器の性能向上が大きく寄与している。

デジタル機器と人間の処理能力

図1は, 代表的なデジタル機器を構成するCPUやGPU (Graphics Processing Unit) といったプロセッサ能力, ハー



ユビキタス化するデジタルメディア処理技術を支えるソフトウェア開発の標準化動向

数百万画素を超える画像センサ搭載の携帯電話や、顔の表情をとらえてシャッタの押せるデジタルカメラが登場するなど、近年の携帯デジタル機器におけるメディア処理技術の進歩は著しい。これらの機能は、ハードウェアである画像センサの高解像度化と、センサで取得した画像に対する顔の検出や表情の同定といったソフトウェア処理との組合せによって実現されている。付加価値の高いデジタル機器を開発するうえで、ソフトウェア処理の重要性は年々増してきている。

一方で、従来のハードウェアに合わせた作り込み主体のソフトウェア開発では、ソフトウェア資産を再利用することが難しいという課題があった。ハードウェアの違いを意識することなくソフトウェア資産を有効活用

メディア	CG (ハイエンド)	CG (組込環境)	ベクトルグラフィクス	映像	音響 (組込環境)
メディア処理API	OpenGL (2.1)	OpenGL/ES (2.0)	OpenVG (1.0)	OpenMAX (1.0)	OpenSL/ES (1.0)
メディア連携API	EGL (1.3)			OpenML (1.0)	
連携データフォーマット	COLLADA (1.4)				
統合開発環境	OpenKODE (1.0)				

* () 内の数値は、現在のバージョンを示す。

できる枠組みとして、Khronos Group⁽¹⁾ による標準化活動が存在する。ハードウェアベンダー、ソフトウェアベンダーを含めて100社強が参加し、オープンで機器に依存しないメディア処理のプログラミング関数 API (Application Programming Interface) の規格策定を目指している。表はKhronos group で策定活動中の規格リストである。

ハードウェア能力を生かしたメディア処理APIや、メディア間の連携と同期処理をサポートするメディア連携APIのほか、ソフトウェア開発全般の統合に必要なデータや開発環境が含まれている。この規格に基づいたシステムを開発することで、ハードウェア及びソフトウェア資産の有効活用や開発期間の短縮などの効果が期待されている。

ドディスク装置 (HDD) の容量、ディスプレイ解像度の推移を表したものである。人が見る対象であるディスプレイ画面の解像度及び、データを蓄積するためのHDD容量の性能向上速度に比べて、画面の描画処理をするGPUによる処理性能の向上速度が高いことがわかる。これは、GPUが演算ユニットを並列化させることで総合的な演算性能を向上させることが容易な構成になっているためである。CPUも演算ユニットのマルチコア化 (複数の演算ユニットを1チップに実装すること) による性能向上が進んでいるほか、CPUとGPUの統合チップ化も進展するなど、今後もデジタル機器におけるメディア情報の処理能力は向上していくことが予想されている。また、デジタルTVで代表的なハイディフィニション解像度 (HD:1,920×1,080画素) は、現在の代表的なディスプレイ解像度であるが、映画などへの利用を想定した大画面プロジェクタでより高い解像度 (4,096×2,160画素) の映像表示技術も開発されるなど、デジタル機器における表示性能についても今後進化していくことが想定される。

視覚や聴覚を通して入ってくる画像や

音声といったメディア情報に対する人間の処理能力は、年代とともに変化するものではない。われわれは身近に利用しているデジタル機器の性能向上によって、メディア情報の処理能力を拡張してきたと言える。一方で、デジタル機器の性能向上は、人間の情報処理能力を超える大量のメディア情報の扱いを可能にしてきた (困み記事参照)。その結果、人間の知覚できる情報量との間に乖離

(かいり)が発生し、デジタル機器との円滑な情報伝達を阻害する要因として近年問題になってきている。人間の特性を考慮したメディア情報の管理と処理技術が今まで以上に重要である。

メディア処理技術の発展の方向性

ディスプレイの表示解像度とデジタル機器の処理能力との関係を図2に示す。ディスプレイの画面サイズにも依存す

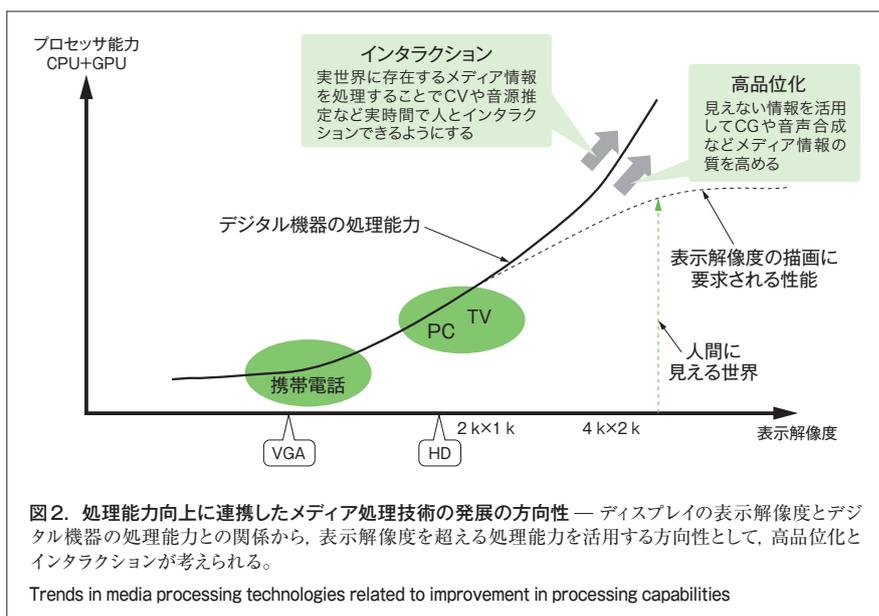


図2. 処理能力向上に連携したメディア処理技術の発展の方向性 — ディスプレイの表示解像度とデジタル機器の処理能力との関係から、表示解像度を超える処理能力を活用する方向性として、高品位化とインタラクションが考えられる。

Trends in media processing technologies related to improvement in processing capabilities

るが、ディスプレイの解像度は人間の目に見える世界の限界に近づきつつある。一方で、デジタル機器のメディア処理能力は目に見えるものの処理、つまり表示解像度分の描画処理に要求される性能を大幅に超えるようになってきた。余った処理能力を活用する方向性には大きく分けて、提示するメディア情報そのものを高品位化する方向性と、実世界に存在するメディア情報を処理することで人とのインタラクションを実時間で実現する方向性の二つが考えられる。

以下では、大量のメディア情報の処理技術が発展する方向性として高品位化とインタラクションについて、更に、情報を見つけやすくする“つながり”を基にしたフィルタリングについて、技術の動向を述べる。

感性に訴える高品位な表現へ

高品位化を実現するためのメディア処理技術としては、従来、見えない、あるいは聞こえないものとして扱われてこなかったメディア情報を活用することが考えられる。

高次元テクスチャ技術（この特集のp.22-25）は、環境条件（視点位置や光源位置）の異なる状況で取得した大量の実写テクスチャ（物体表面の“きめ”など）画像をデータベース化しておくことで、見る位置に応じて変化する質感の高いコンピュータグラフィックス（CG）映像を生成できるようにしている。

ベクトルデータという、画素ではなく目には見えない構造を活用することで、高品位な文字を解像度に依存せずに描画できるようにする技術が高速3次元ベクトル表示技術（p.26-29）である。この技術は、高精細化を求められている携帯電話などで、GPU搭載を背景に画像インタフェースとしての活用が期待されている。

“文字”とそれ以外の“図”などで構成されるドキュメントの構造に着目して高品位化と高圧縮率を両立する技術が、

高圧縮PDF（Portable Document Format）技術（p.30-33）である。

高音質音声合成技術（p.34-37）は、多数の単位音声データを大量の実音声から生成しデータベース化しておくことで、高品位で自然な抑揚のある音声を合成しており、カーナビインタフェースなど幅広く展開されている。

今後も、多次元で大量のメディアデータや隠された構造を抽出して活用することで、解像度や画面サイズといったデジタル機器の特性に依存することなく、高品位で感性に訴えかけられるメディア情報の表現が探求されていくと思われる。

自然で直感的な インタラクションの実現へ

デジタル機器の処理能力の向上は、実世界に存在する音声や映像といったメディア情報を容易にかつ大量に取得できるようにしたほか、実時間でメディア情報を処理できるようにしてきた。メディア情報は、映像に国境がない一方で、音声には多言語が存在するなど、グローバルな視点から研究・開発を行う必要がある。

東芝は、グローバルな研究開発拠点を欧州と中国に設け、日本を含めた3極体制で連携しながらメディア処理技術の研究・開発を進めている。東芝欧州研究所のケンブリッジ研究所では、当社研究開発センターとも連携し、コンピュータビジョン（CV）（p.3-17）と音声情報処理分野の研究を主に推進している。今後は人間であるわれわれが、ストレスを受けることなく、デジタル機器と実時間でインタラクションできることが求められている。そのための要素技術として、複数方向から取得した実写映像をもとに3次元モデルを獲得する技術や、ジェスチャを利用したインタフェース技術の開発を行っている。

人間が自然に行う行為をデジタル機器に解釈させることも重要である。

顔認識技術（p.38-41）は、複数台の

カメラで取得した大量の画像列を解析することで顔の識別ができるようにした技術であり、歩行中という自然な行動形態での個人識別を可能にしている。ユーザーに負担を掛けない個人同定技術としてセキュリティシステムに組み込まれ、安全な社会の実現に活用されている。

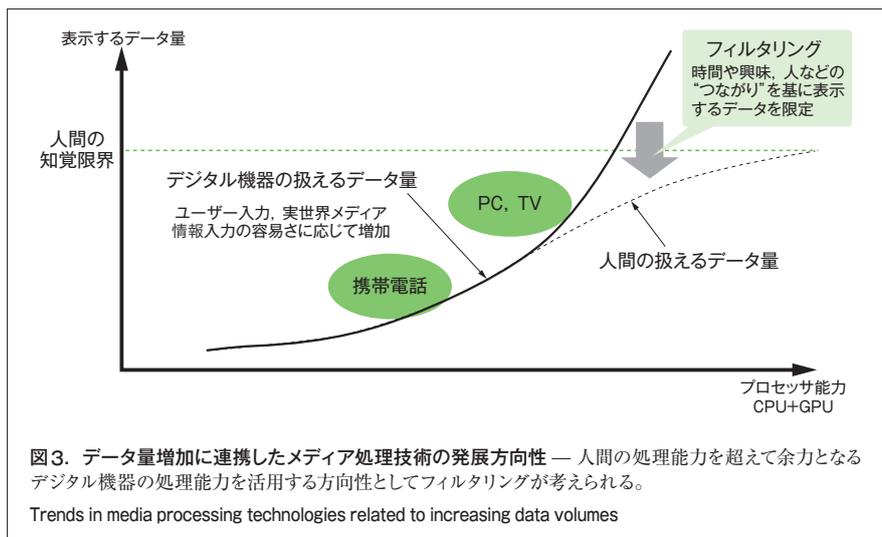
重ね書き文字認識技術（p.42-45）は、画面の狭い領域に手書きの要領で重ね書きすることで、連続して文字を入力できるようにする技術である。同じ場所に書き続ける形式のため、文字の入力に自然に専念でき、携帯電話などへの活用が期待されている。

全方位聴覚技術（p.46-49）は、複数方向から入力される音声データを解析することで音源方位の特定を行っている。当社が開発しているロボットApriAlpha™に搭載することで、人間どうしの対話と同様に、その人に向かって話すというインタラクションができる。更に、ApriAlpha™がネットワークにつながることで、時間と空間を越えた人とのコミュニケーションを実現したものがブログインタフェース技術（p.50-53）である。いずれも、家庭などでロボットが人間社会に溶け込み、自然なコミュニケーションを行うために必要な技術である。

従来は人間が機械に合わせる形で発展してきたデジタル機器であるが、今後は機械が人間に合わせる時代へと移行し、人間主体のインタラクションを実現することが求められている。そのためには、人間の身体性や行動の理解、及び知見の発見を通して、受動的で存在を感じさせないインタフェースの実現が重要になってくる。

“つながり”に基づく情報 フィルタリングの実現へ

デジタル機器がブロードバンドに代表される広帯域のネットワークを経由してインターネットとつながることで、大量の知識の取得が容易になった反面、人が



求める情報にたどりつけない状態が社会的な問題として提起されている。画面に表示するデータ量とデジタル機器の処理能力との関係を図3に示す。人間が一度に知覚できるデータ量には限界があるため、大量にかつ容易に取得できるネットワーク上の情報を、人間の知覚可能なデータ量に変換処理して見せる必要がある。一方で、探している情報がわかっている場合にはその言葉を用いて検索すればよいが、今後は、探している情報が明確でない、あいまいな情報を頼りに探索を支援する技術が求められる。ここに、フィルタリングの技術が考えられる。

関連性に基づいた情報の絞込みを行う連鎖検索インタフェース (p.54 - 57) は、画面の文字情報を頼りに連鎖的に関連情報へアクセスできるようにしている。新たな知見につながったり、興味のある情報が効率的に収集できるといった応用が考えられる。

ホットワードリンク™技術 (p.58 - 61) は、時間を限定することでネットワーク上の最新動向の把握とアクセスを容易にする技術である。当社のAVノートPC Qosmioシリーズに搭載され、AV視聴中に興味を持った句 (しゅん) の情報を知るなど、情報に敏感なユーザーへのサービスとして提供されている。ユーザーの多様な興味や関心への対応も考

えていく必要がある。

同じ興味のあるDVDコンテンツに関する評価やコメントをネットワーク上で共有するシステムとして開発されたものがコミュニティ™ (p.62 - 65) である。ユーザー間の“つながり”を支援するサービスとしての活用が考えられている。

また、ユーザーの状況や要求に応じた情報アクセスも重要である。特に高い専門性を求められる医療診断支援システム (p.66 - 69) では、患者の状況に応じた診断支援や情報提供が可能になっている。遠隔地など空間的に離れた医者と患者を結び付け、安全で安心な社会を支える医療システムの社会インフラ化が求められている。

今後も、大量に存在するデータの中から“つながり”に基づく情報のフィルタリングと画面への提示をデジタル機器が支援していくことになる。多様な価値観を持つユーザーがデジタル機器を通して時間と空間を越えてつながり、新しい価値ある情報が生み出される場面で、メディア処理技術は活用されていくものと思われる。

人にやさしく、人と人をつなぐ メディア処理を目指して

デジタル機器の性能向上が、デジタル機器とその背後にある大量の情報を

人間の能力に応じた形で伝えるためのデジタルメディア処理技術として活用されていることを述べた。

今後もメディア処理技術は、われわれに感動をもたらす高品位なメディア情報による表現、自然なインタラクションを可能にするメディア処理、そして多様な“つながり”に応じたメディア情報のフィルタリング、といった方向性で開拓されていくものと思われる。そして、身近に存在するデジタル機器がこれらのメディア処理技術を搭載し、時間と空間の制約を越えて人と人の距離を縮め、多様なつながりを支援するプラットフォームとして発展していくことを目指している。

今後も人の理解、人や環境に対する新たな知見の獲得、更に、新しいデバイスやアルゴリズムの開発を通して、人にやさしい機器、インタフェース、及びサービスの創出に取り組んでいく。

文献

- (1) Khronos Group. Open Standards, Royalty Free, Dynamic Media Technologies. <<http://www.khronos.org/>>, (accessed 2007-11-02).



山内 康晋
YAMAUCHI Yasunobu

研究開発センター ヒューマンセントリックラボラトリー 研究主務。コンピュータグラフィックス及びヒューマンインタフェースの研究・開発に従事。情報処理学会、ACM 会員。
Humancentric Lab.



土井 美和子
DOI Miwako, Ph.D.

研究開発センター技監、工博。ヒューマンインタフェースの研究に従事。情報処理学会副会長、電子情報通信学会評議員、映像メディア学会評議員、HI 学会副会長、IEEE、ACM 会員。
Corporate Research & Development Center