*立石賞特別賞の受賞記念講演概要*

# 超高速ビジョンの開発と高速知能システムへの応用

東京大学大学院 情報理工学研究科 教授　石　川　正　俊

従来の画像処理が主に空間分解能の向上に主眼があり，時間軸としてはビデオレート，すなわち 30 fps を用いていたのに対して，画像の時間分解能にこだわり，機械システムのダイナミクスをカバーする高速領域の画像処理システムを様々な形で実現し，今まで画像の遅さに起因して高速性が実現で来ていなかった応用分野を開拓し，高速ビジョンの世界を構築してきた。

この際，画像処理だけを高速化しても，その高速性を利用するシステムに低速の要素が混じっているとその性能に起因して全体の性能に限界を生じるため，システム全体を高速化することが求められる。このため，高速画像処理を基盤として，処理系や運動系等とともに新しい機能を生み出すシステムを総称して高速ビジョンと呼んでいる。

従って高速ビジョンは，システムのダイナミクスを把握・制御し，新たな知能システムを生み出すことを意味している。歴史的には，1993 年に筆者らが発表した SPE-4k が高速画像処理の第一歩であったが，その後，種々の高速画像処理ハードウェアを開発するとともに，様々な応用展開の中で高速ビジョンが生み出す新しい世界を作り上げてきた。

## 高速ビジョンの基本構造

高速ビジョンを実現するためには，そのシステムアーキテクチャも高速性を実現するものでなくてはならない。一般に知能システムは階層的並列分散構造を取ることになるが，その際，与えられたミッションに対してタスク分解の構造が重要となる。筆者が直列分解と名付けている方法は，センサの情報を認識モジュールに送り，高次の判断を経て運動系の制御信号が運動系に送られるというアーキテクチャであるが，この場合，時間がかかるモジュールが感覚系と運動系の間に存在するため，高速性を実現できない。これに対して，低次のセンサフィードバックループを並列に動作させる並列分解は，高速性を実現できるモデルであり，さらに数学的に厳密な直交分解が理想的なタスク分解方法ととして提案されている。

さらに，筆者らは，システムの設計思想として，「ダイナミクス整合」を提案している。これは，感覚系としてのセンサ，処理系としてのコンピュータ，運動系・表示系としてのアクチュエータやディスプレイのすべてのダイナミクスを対象や環境のダイナミクスをカバーする帯域で設計することを意味している。ある意味で自明の考え方であるが，従来のシステムでは，実際に実現できるデバイスを優先して，対象のダイナミクスをカバーできない要素を用いたシステムで高速性を実現しようとしている点に無理がある。

このようなシステムでは，対象や環境の情報に欠落を生じ，それを補うために，学習，推測，補間等，情報の欠落を補う複雑なアルゴリズムが必要であった。これに対して，ダイナミクス整合のもとでは情報の欠落はなく，簡単なアルゴリズムで大きな効果が実現できる。

従来の知能ロボットは，繰り返し精度を評価指標としたプレイバック制御が主流であったが，

非プレイバック制御時には，ダイナミクスの不確定性やギアのバックラッシュによって，極端な精度の低下を生じるが，従来のロボットの手先に高速ビジュアルフィードバック装置を付加することにより，ロボットのダイナミクスや不確定性等も含めて高精度な制御を実現することが可能となる，筆者らはこのような制御方法をダイナミック補償と呼び，高速高精度な産業用ロボットの実現のための構成方法・制御方法として提案している。

また，カメラの光軸を高速のアクチュエータで駆動・制御することにより，固定視野のカメラでは得られない画角や位置精度を得る方法をアクティブビジョンと呼び，高速画像処理による対象の認識と位置の計測から得られる時間密度の高い情報を用いて，カメラの光軸を対象物に合わせたトラッキング撮像を可能とする方法として，新しい応用が開拓されている。

## 高速画像処理のハードウェア

高速画像処理の実現に，通常の CPU や GPU を用いると，走査回路に速度限界が存在することから，十分な速度を得ることは難しい。そこで，撮像素子とボトルネックのないアーキテクチャが必要となる。筆者らは，イメージャと並列処理を結合する方法として，画素ごとにフォトディテクタとプロセッシングエレメントを直結した完全並列タイプ，行側の走査回路を残した列並列タイプ，さらには小数のスタンダードな転送回路を用いたタイプ等を開発してきた。

昨年の ISSCC で，われわれの研究室とソニーが共同で，積層型のビジョンセンサーを発表した。これは，1 層目が高速高感度の CMOS イメージャで，2 層目が列並列の A/D 変換とそれに繋がる並列 PE を実装したものとなっている。現在，商品版が開発・市販され，㈱エクスビジョンから SDK が開発・市販されている。

高速画像処理のアルゴリズムは，一般の画像処理の高速化ではない。処理対象の帯域をカバーする時間密度でのサンプリングを前提とすることにより，従来の画像処理とは違ったアルゴリズムが効果を発揮する。特にトラッキングでは，フレーム間で，対象の変化がほとんどないことから，極めてシンプルなアルゴリズムでも優れた性能を実現できる。筆者らの研究室では，この考え方に基づき，様々な高速画像処理用のアルゴリズムを開発している。

また，知能化されたイメージャという観点から，高速ビジョンネットワークへの展開が期待されている。その際，高速化でカメラ間の同期の問題が顕著となるため，ネットワークカメラの時間管理方法の開発が必要となる。

## 高速ビジョンの応用展開

これらの設計思想に沿って，ハードウェアを開発し，様々な形で高速ビジョン並びにその応用システムが開発されている。

ヒューマンインターフェイスの分野では，速度が 100 km/h を超える人間の動作を計測し，遅延なく人間に提示する高速低遅延ジェスチャー認識システムが開発されている。このシステムでは，高速画像処理と高速提示デバイスにより，人間が入力と提示のずれを感じることのない快適なインターフェイスが実現されており，パソコンのタッチ操作をエミュレートするものや没入感の高いゲームインターフェイス等が実現されている。

これを利用して，無拘束空中像や 3 次元像のジェスチューでの操作や自動車のヘッドアップディスプレイ（HUD）の 3 次元化（コニカミノルタとの共同研究）が実現されている。

また，ターゲットトラッキングを用いたトラッキング撮像装置では，卓球のボールのトラッキング撮像等のスポーツサイエンスへの応用や高速飛翔体の周りの衝撃波の観測（科学警察研究所との共同研究）等の科学計測で活用されるとともに，トラッキングされている光軸を利用することにより，運動する対象物にプロジェクションマッピングを実現するダイナミッ

クプロジェクションマッピングやレーザードップラー装置を用いて運動する対象の振動計測等も実現されている。

ダイナミックプロジェクションマッピングとしては，対象が布や紙のように変化する物体である場合には，新たに東京エレクトロンデバイスと共同研究により開発した 1,000 fps の高速プロジェクターを用いることにより，変形する T シャツや紙，さらには人間の顔等にプロジェクションマッピングを実現するシステムが開発されており，メディアアートとしても作品が発表されている。

ミラーによるトラッキング撮像を用いると，動いているカメラからブラーのない撮像を実現することも可能である。Nexco 中日本との共同研究により，トンネル壁面の 0.2 mm のクラックを時速 100 km の車両から観測するシステムも開発されており，既に実証実験が行われている。このシステムは工場で長尺ものを扱う際の検査装置にも適用可能であり，様々な展開が期待されている。

高速マルチターゲットトラッキングでは，1,600 個程度までの複数対象物の高速同時トラッキングが可能であり，例えば 32×32 のスポットのパターン照明を用いることにより，1,024 点の 3 次元情報を 1 ms ごと得るシステムが開発されている。これもエクスビジョンにより，実用化されている。

この高速 3 次元形状計測機能を用いて，本の電子スキャンシステムも開発されている。1 分間に 250 ページの電子化が可能である。

回転対象の検査では回転光学系を用いて回転方向のトラッキング撮像を実現し，顕微鏡像に対しては，XY ステージに高速トラッキング情報をフィードバックして，運動する対象物の安定かつ長時間の観察を可能にしている。

Z 方向に対しては，高速可変焦点レンズを用いてトラッキングが可能であり，また複数の奥行き方向高速スキャン像を用いた全焦点画像の取得も 1,000 fps で可能となる。

一般的なロボットは，プレイバックあるいはビデオレートのビジュアルフィードバックを用いているので，高速であれば対象の変化に対応できず，ビジュアルフィードバックで適応性を出そうとすれば速度が落ちてしまうという問題を抱えている。そこで，手先に高速ビジュアルフィードバックを導入することにより，人間が投げたボールを 100% 打ち返すバッティングロボットや逆に動作範囲内に入ってきた人間等の対象物を回避するロボットシステムが開発されている。

また，通常の産業用ロボットをラフな教示，例えば簡易型のビジョンによる教示に基づく動作に対して，ダイナミクスを含めて手先に取り付けた高速ビジュアルフィードバックシステムでエラーを補償することにより，高精度・高速な動作を実現するロボットシステムも提案されている。このような方法をダイナミック補償と呼んでいる。

産業用ロボットの代わりに，人間の動作に対しても，この方法を適用することが可能で，人間の動作の高速化・高精度化を実現することが可能となる。このようなシステムをアクティブアシスタントと呼び，次世代の人間機械協調システムの一つのソリューションを提供するものと考えている。

また，高速のロボットハンドを開発し，高速ビジョンと組み合わせることにより，落下する生卵をキャッチするハンド，ツールの持ち替えを投げ上げで実現するリグラスピング，さらには高速指ドリブル，ピンセットキャッチ，柔軟体ハンドリング，人間に 100% 勝利するじゃんけんロボット等が実現されている。

さらに，このハンドで開発された高速アクチュエータを用いて，ビジュアルフィードバックにより，不安定な姿勢でも，多少の悪路でも走行可能な高速二足走行ロボットも開発されている。

この他，高速ビジョンは，自動車や飛行体のビジョンによる制御やビジョンチップの演算機

図１　高速ビジョンの開発

能をネットワークにアップする情報の高機能化や情報圧縮に使おうという試みも行われており，数多くのビジョンチップを用いたビジョンネットワークによるセキュリティシステムが考えられている。

これらの研究成果は，研究室の成果集[1]としてまとめられ，動画は動画サイト[2]にアップされている。それらの普及のため，WINDS ネットワークが２年前に設立され，様々な分野に対して応用展開が検討されている[3]。

## 知能システムの未来

知能システムあるいは AI に関しては，現在，様々なコンセプトやプロジェクトが提案・実施されている。それらの多くは，既存の膨大なデータに対して，学習により，高度な機能を実現したものが多く，その高機能性はデータの大きさに起因するものであり，その要因以外については言及されることが少ない。そもそものデータの構造をその取得から考えることが少なく，センシングの構造を目的の機能の実現のために議論することは少ない。特に，サンプリングが不十分な時間密度と不十分な空間密度で得られたデータに対しては，限界が存在する。

これを解決するには，知能システム全体の時間密度と空間密度をセンシング，プロセシング，アクチュエーションあるいはディスプレイの各要素並びにシステム全体の観点から，ダイナミクス整合のデザインコンセプトに基づき，並列分解に基づくタスク分解を設計する必要がある。

このことは，リアルタイム実世界に対応する知能が，非リアルタイムの世界で動作する知能とは一線を画すものであることを意味し，工場の製造ラインの設計や自動車の制御系の設計に通じるものである。つまり，欠落のある情報に対する非リアルタイムの知能を起点とすることではなく，時間密度と空間密度が十分で欠落のない情報を用いたリアルタイム系の知能に起点を置き，そこに高次の知能を埋め込む方向性が重要となる。

### 文　献


1）　石川妹尾研究室 研究成果集
http : //www. k2. t. u-tokyo. ac. jp/Booklet/index-j. html

2）　YouTube Ishikawa Senoo Laboratory Channel
https : //www.youtube.com/user/IshikawaLab

3）　WINDS ネットワーク
http : //www.winds-network.org/