

# 1ミリ秒を捉える機械の目で 知能システムを革新

テレビなどの映像では、滑らかに動いて見えるよう画像を連続して提示する。1秒間に提示するのは通常30枚、あるいは60枚程度と人間の目の能力に合わせて設定されている。これをはるかに超える速さで世の中の現象を捉えられたら、どのようなことができるのか。毎秒1000枚の速度で撮影し、処理する機械の目「高速ビジョン」を実現し、その可能性を追究しているのが東京大学大学院情報理工学系研究科の石川正俊教授だ。岸則政JST ACCELプログラムマネージャーとの二人三脚で高速ビジョンプラットフォームを構築し、新たな価値を生み出そうとしている。

いしかわ まさとし

石川 正俊

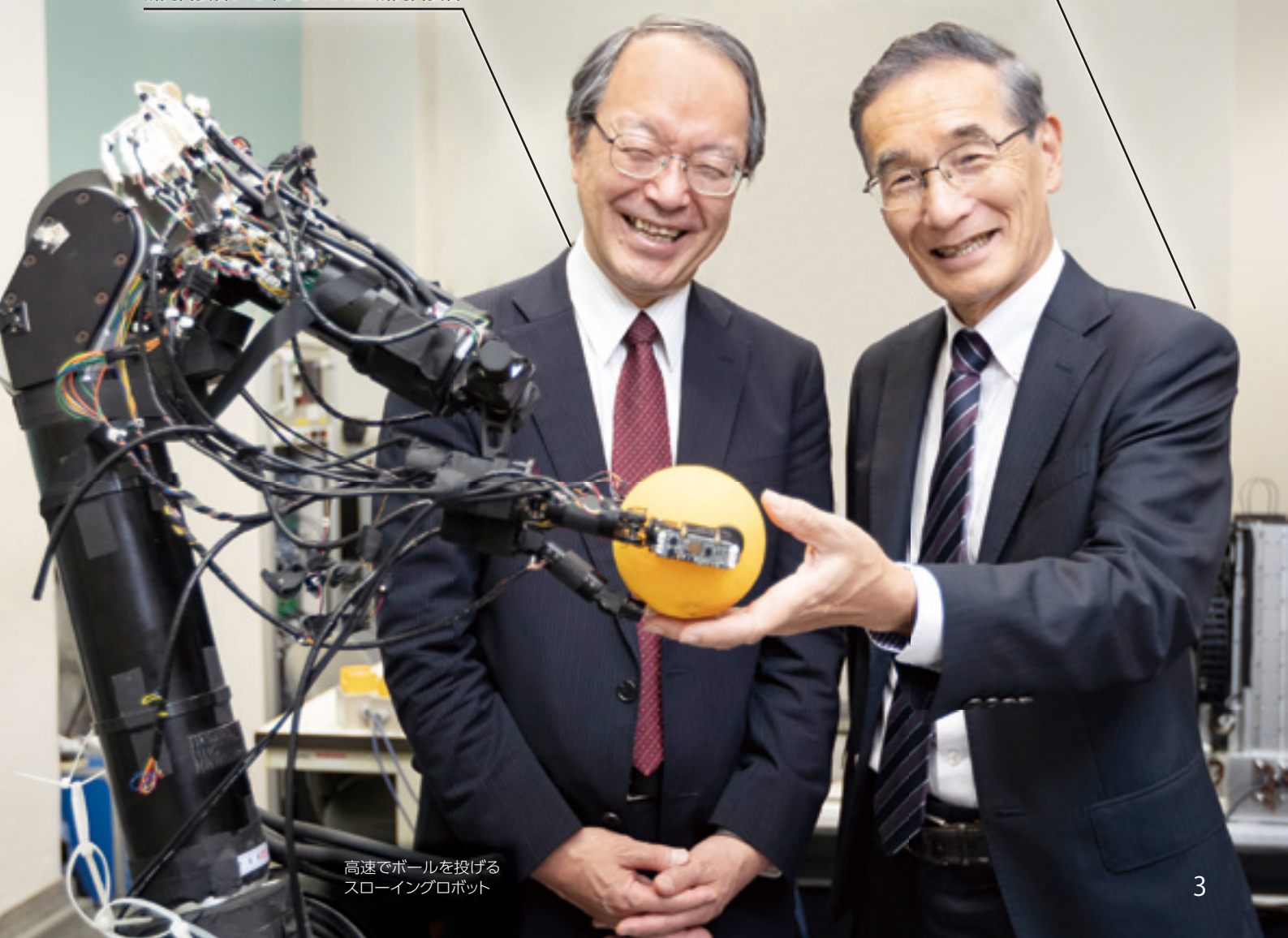
東京大学 大学院情報理工学系研究科  
研究科長・教授

1999-2004年、09～15年 CREST  
研究代表者、16年よりACCEL研究代表者

きし のりまさ

岸 則政

JST ACCELプログラムマネージャー／  
東京大学 大学院情報理工学系研究科  
客員研究員



高速でボールを投げる  
スローイングロボット



## 高速ビジョンで捉える世界 価値の芽を育て波及させる

動画を撮影したり、再生したりする際に、1秒間に何枚の画像を表示するかを表す数値がフレームレート (fps) である。人間の目では30fps程度であればスムーズに動いて見え、高精細画像を実現する8K映像では、30fps、60fpsに加えて120fpsの規格も用意されている。機械の目でこれをはるかに超える1000分の1秒 (1ミリ秒) ごと、1000fpsを可能にするのが「高速ビジョン」だ。

東京大学大学院情報理工学系研究科の石川正俊教授は「1000fpsで捉えることにより、高速で動く物について、より完全な情報を得ることができます」と話す。例えば、時速150キロメートルのボールを30fpsで撮影すると約140センチメートルごとの姿しか捉えられず、その間のボールの動きは推定するしかない (図1)。しかし、1000fpsであれば約4センチメートルごとに捉えられ、ボールの実際の動きを正確に把握できる。推定が不要なので処理時間も短縮され、これまで不可能だったさまざまなことが実現可能になる。

「高速ビジョンは単に高速で撮影するだけの技術ではありません。撮影から画像処理、そしてロボットの動きなどにフィードバックするまでを、数ミリ秒で行うのです」と石川さんは続ける。カメラをはじめとする感覚系・認識系 (センサー技術)、処理系 (コンピューター技術)、そしてモーターをはじめとする運動系・行動系 (アクチュエーター技術) をそれぞれ1ミリ秒で動かすことで、リアルタイムで実世界の変化に適応、機能するまったく新しい知能システムを実現できるのだ。

高速ビジョン研究は「これまでの科学のアプローチとは異なる」と石川さんは強調する。従来の科学技術は、ある課題を解決するために原因を調べ、理解し、技術を生み出してきた。一方、高速ビジョンでは解決すべき具体的な課題が先にあるわけではない。高速ビ

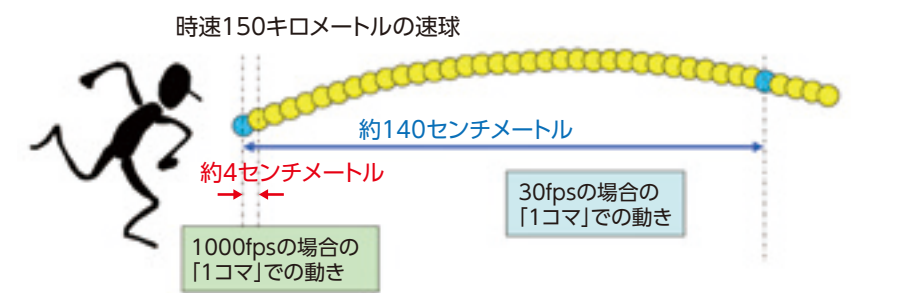
ジョンを出発点に、今まで考えもしなかったような可能性に気付き、新しい価値を生み出すのである。「私たちは高速ビジョンのプラットフォームを世に出していきます。その価値の芽が大きな波となり、多様な分野に波及していく。そうやって新しい世界を築きたいのです」と目を輝かせる。

## 自動運転から壁面の検査まで いろいろな分野で技術を磨く

自動車への画像処理の応用で長年の経験を持つ岸則政ACCELプログラムマネージャーは、高速ビジョンのわかりやすい用途として自動運転を挙げる。「例えば、自動ブレーキは周囲の物体を捉え、それがどのように動くかを予測して制御するため、突然飛び出し

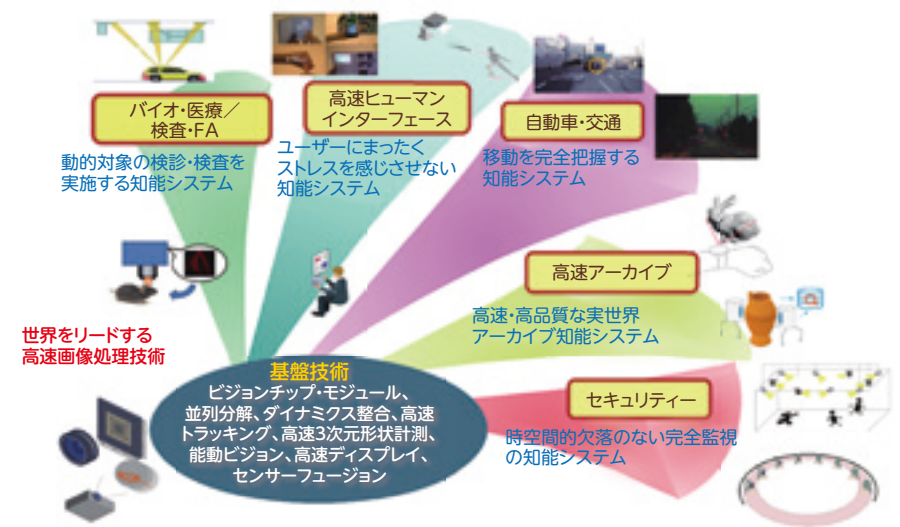
てくる物体に対しては反応が遅れることがあります。一方、高速ビジョンであれば、1ミリ秒ごとに周囲の状況を分析、判断し、瞬時にブレーキを制御できるのです」。プロジェクトではこの他にも、バイオ・医療／検査・工場自動化 (FA)、高速ヒューマンインターフェース、交通、高速アーカイブ、セキュリティなどと実に多岐にわたる応用範囲をターゲットにしている (図2)。

岸さんは「プロジェクトの目的は、新しい技術で気付きを与えることです。そのためには、具体的に何ができるかを示し、多くの人にどうやって使えるかを考えてもらうことが重要です。1つの用途を想定し技術を尖らせるのではなく、いろいろな分野に飛び込んでそこで求められる技術を磨く戦略をとっています」と説明する。この戦略に基づき、ボールなど動く物体を追いか



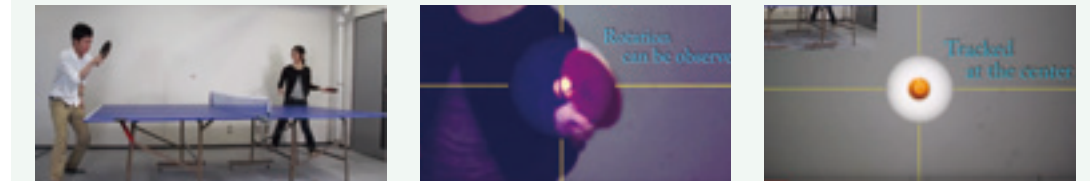
## 高速な動きを確実・正確に認識できる

■図1 時速150キロメートルで投げたボールを30fpsで撮影すると、撮影の間隔は約140センチメートルになる。一方、1000fpsでは約4センチメートル間隔でボールの動きを把握できる。30fpsの場合、約140センチメートルの間に何が起きているかを推定する必要があるが、1000fpsであれば実際の動きを捉えられるので推定のための煩雑な計算は不要になる。



■図2 高速ビジョンが目指す適用分野。高速ビジョンの基盤技術（プラットフォーム）は多彩な分野への応用が可能だ。

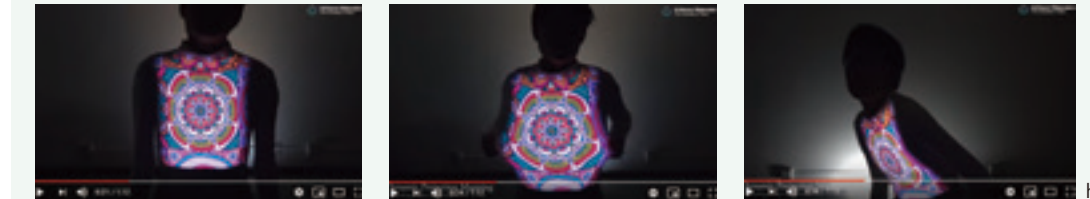
■図3 高速トラッキング技術。高速で動く卓球の球の動きを、自動的に追いかけて撮影でき、球の回転まではっきりと観察できる。画像を解析し、ボールが常に画像の中央にくるようにミラーを高速で動かしている。



[https://youtu.be/9Q\\_IcFZOgVo](https://youtu.be/9Q_IcFZOgVo)



■図4 ダイナミックプロジェクションマッピング。着ている人の動きに合わせて、Tシャツの正面に画像が投影される。生地伸び縮みも検出し、画像を瞬時に調整する。



<https://youtu.be/QDppJ9NWtaE>



け常に画面の中心に捉えて撮影する技術や、動いたり伸び縮みしたりする物体に映像を映し出すダイナミックプロジェクションマッピングなど多彩な技術が開発されている (図3、4)。

応用が期待される技術の1つが、トンネル内走行中の壁面検査だ (図5)。中日本高速道路との共同で、導入に向けた実験を進めている。従来、トンネルの安全検査は道路を封鎖した上で行われていた。しかし、高速ビジョンを利用すれば、カメラを搭載した検査車両で通常速度で走行しながら撮影し、トンネル内の亀裂の有無などを高精度で検出できる。時速100キロメートルで走る車からトンネルの壁面を撮影すると、1000fpsでも1コマで約3センチメートルも進んでしまう。そこで、ミラーを高速で動かして通り過ぎた箇所を画面の中央に捉え続けることで動きを打ち消し、ぶれののない画像を撮影する。

検査のための画期的な技術で、交通規制による渋滞なども避けられ、コストも下がるため、検査の頻度を増やすことが期待される。この実現のためには高速画像処理だけではなく、車に搭載するための小型化や、トンネル内で他車の走行を妨げない低照度照明での撮影なども必要だった。そういった周辺技術も含めた開発により、実用化が着実に近付いている。

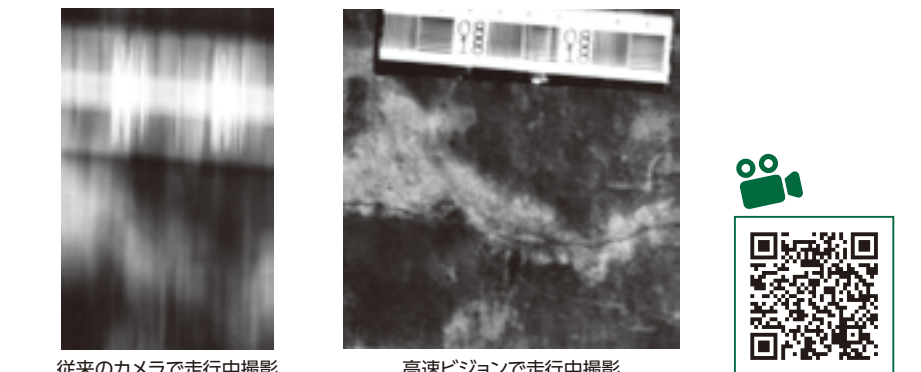
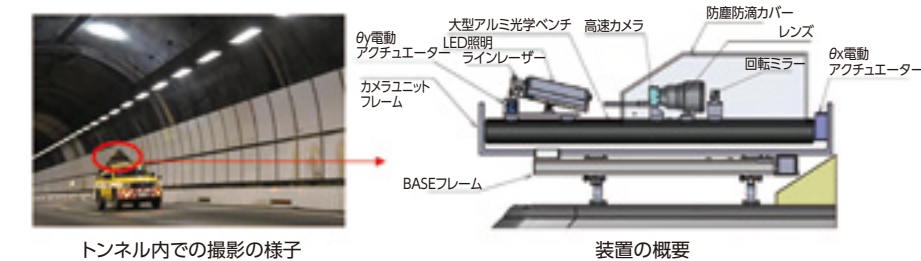
## 高速化を支える基盤技術 価格や消費電力も大幅に低減

さまざまな可能性を秘めた高速ビジョンを支える基盤技術について「知能システムに関わる、センシング、処理、そして運動系や表示系の制御の全てが高速であることが重要です。これをダイナミクス整合と呼びますが、どれか1つが欠けても1000fpsを実現することはできません」と石川さんは語る (図6)。岸さんも「処理速度だけを上げても、センサーの能力が低かったり、ロボットの制御に時間がかかったり

すればシステムとしては遅くなります。モーターなど周辺技術が遅れの原因になる場合、新たに開発し高速化を図ってきました」と続ける。

例えば、動く対象を追いかける高速トラッキング技術では、複数のミラーを動かして対象の姿を追跡するが、この動きがずれてしまったらそこで失敗だ。ミラーをできるだけ速く、精密に動かすためにはミラーの大きさなど細部まで作り込む必要がある。

また、水などの流動物や高速で動く対象に映像を映し出すダイナミックプロジェクションマッピングでは、映像を

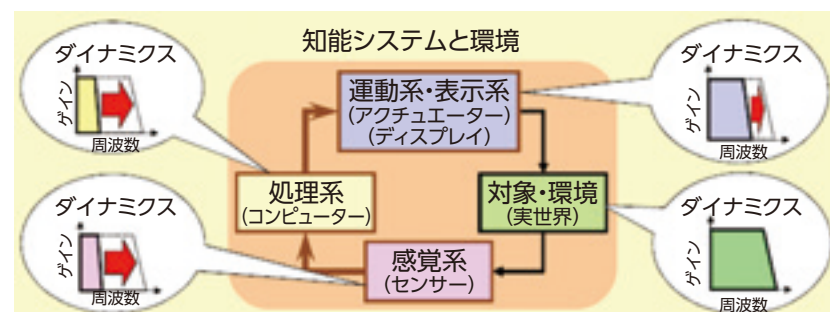
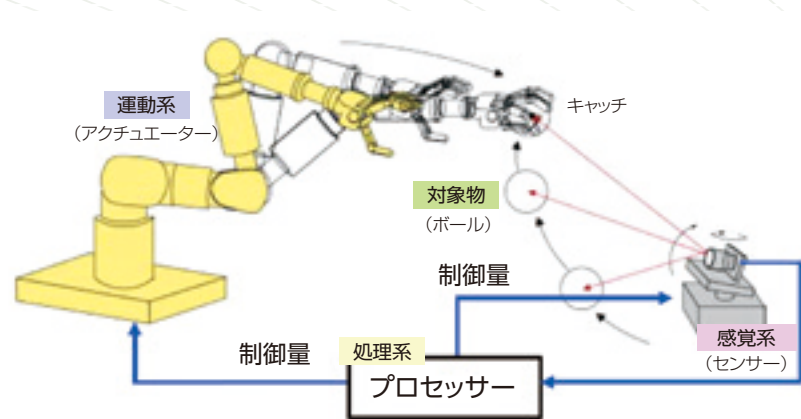


<https://youtu.be/AkrORWW7NO4>

■図5 時速100キロメートルで走行しながら、幅0.2ミリメートルのトンネル内のひび割れを検出できる。ミラーを高速で動かして対象を捉え続けることで、動きによるぶれを抑えることが可能だ。渋滞の解消、点検頻度の向上、作業の効率化などが期待される。







■図6 上：高速ビジョンを実現するダイナミクス整合。対象物を動きに合わせて高速で検出し、それを高速で処理し、高速でアクチュエーターやディスプレイを制御する。下：このループにおいて、それぞれの系に対象物の動きを十分にカバーできる速度と安定性を持たせることで、タイミングの遅れが生じないシステムを実現している。

映し出す面の3次元形状とその変化を瞬時に計算し、それに応じてプロジェクターで映し出す映像を高速で制御しなくてはならない。計算速度はもちろんのこと、データを転送する速度の遅れも問題になる。速さにこだわって研究開発を進め、必要な要素技術を全て磨いて確立されたのが高速ビジョンに基づく知能システムなのだ。

システムを支える基盤技術の中には、すでに商品化されているものもある。その1つがソニーと共同開発したビジョンチップだ。撮像素子と取得画像から特徴量などを抽出する演算器が1平方センチメートル以下のチップに収まっている。「研究を始めた時には高速画像処理を実現するために1メートル四方のシステムが必要でした。1000fpsという速さはそのまま1万分の1以下(面積比)まで小型化したのです。これにより、社会で使える技術になりました」と石川さんは胸を張る。

岸さんも「価格は1000万円から10万円に、消費電力は100ワットから0.3

ワットになりました。空冷や水冷も不要になり、今までは使えなかった用途で使えるようになります。世界の変わり目といえるくらい、大きな進化です」と普及への自信をのぞかせる。

しかし、2人ともここで満足しているわけではない。「現在は、システムの中で一番遅い運動系の上限がシステムの限界となっています。運動系の制御を1ミリ秒より高速化できれば、さらに性能を向上できます。機械の限界を目指したいですね」と石川さんが語るように、さらに先を見据えている。

## わかる科学からつくる科学へ 動画を利用し企業を巻き込む

開発された技術の多くについて、デモンストレーション動画を制作し、研究室のホームページや動画共有サービス「YouTube」上で公開している。その理由を石川さんは熱く語る。「従来の科学を、物事を理解する『わかる科学』とするなら、私たちが目指すのは新たな価値を『つくる科学』です。何に使い、どのような価値を生み出すかはアイデア次第です。それを考えてもらうために、私たちの技術で何ができるのかをわかりやすく紹介しているのです」。

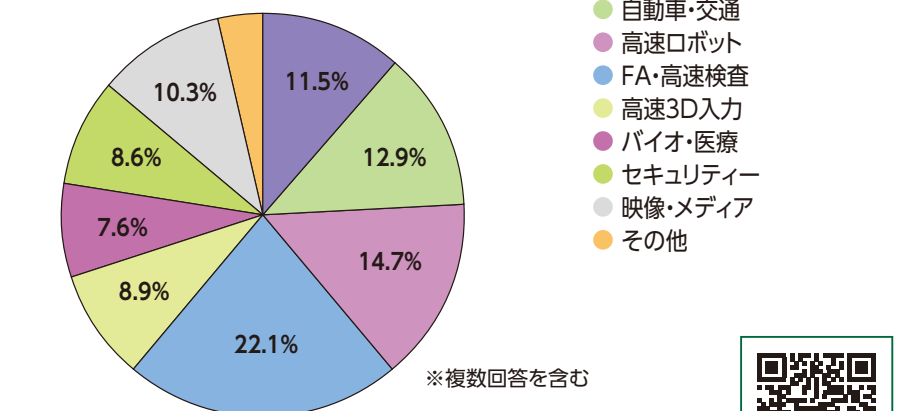
紹介されている動画の中で、最も多く視聴されているのがじゃんけんロボットだ(図7)。人間が出した手の形を瞬時に解析し、ロボットハンドに指令を出すため、同時にじゃんけんするにもかかわらず絶対に負けることがない。この動画でいかに速く画像を処理し、ロボットを制御できるのかを伝えることで、「ロボットと人間が共に働く工場などで、生産性を圧倒的に向上させる

設備に使える」といった用途を視聴者に想起させることが狙いだ。動画は500万回以上再生され、米国の三大ネットワークや英国、カナダの放送局でも取り上げられた。「こういった動画や展示会でのデモを見て関心を示す企業は確実に増えています。まさに百聞は一見にしかずです」と岸さんも手応えを感じている。

一方、日本の企業からの反響は、海外の企業と比べると少ない。産業分野へのさらなる普及を目指し、新エネルギー・産業技術総合開発機構の支援も受けて立ち上げたのが、「WINDSネットワーク(Network for World Initiative of Novel Devices and Systems)」だ。「高速ビジョンの素晴らしさは、動画やデモンストレーションを見てもらえば一目瞭然なのですが、どうやって知ってもらうのが課題でした。そこでコンソーシアムを立ち上げ、情報を発信し、フィードバックを受ける体制を整えたのです」と石川さんは振り返る。現在は182法人から200の組織が参画しているが、その興味は実にさまざまだ(図8)。研究段階では特定の用途に限られてきた高速画像処理技術を多様な分野に展開し、産業の活性化や新しい産業の創出を目指している。

学術分野についても、日本バーチャルリアリティ学会、計測自動制御学会、日本ロボット学会、映像情報メディア学会、日本印刷学会、応用物理学会、日本機械学会、電子情報通信学会といった幅広い領域で研究が高く評価され、論文賞などを受賞している。「1つの分野に閉じこもらずに、どんどん外に飛び出していった結果です。手を広げた先できちんと成果を出してきたことが、自信につながっています」と石川さ

## 会員の技術応用想定分野の分布



■図8 WINDSネットワークに参画する会員の技術応用想定分野は実に多岐にわたる。  
<https://www.winds-network.org/about.php>



んは笑顔を見せる。眼球の微細な動きのせいで強力な照明が必要だった眼底検査を高速ビジョンにより簡便化したり、顕微鏡下で動き回る細胞を自動で追いかける技術を開発したりと医療や生物学など応用の範囲はますます拡大している。

## 未来の真実をつくる科学 柔軟な発想で分野を開拓

新しい価値を生み出すための高速ビジョンのプラットフォームは、特定の課題を解決するための技術ではないと石川さんは繰り返す。「教科書に書いてあることを勉強する基礎力は大切ですが、そこにあるのは現在までの真実です。これからは問題も答えも書いていない未来の真実を探す、つくる科学が重要になるでしょう。そこで研究者に求められるのは柔軟な発想力です」と説明し、それが欧米諸国や中国などとの開発競争に不可欠だと語る。

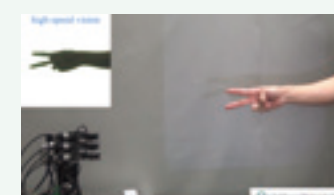
岸さんも「高速ビジョンは世界最高クラスの技術だと自負しています。さらなる普及には固定観念に縛られず、



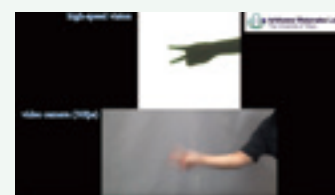
150以上の動画を公開中。  
高速ビジョンのデモンストレーション動画はこちら  
<http://www.k2.t.u-tokyo.ac.jp/movies/index-j.html>



■図7 高速視覚フィードバック。左：相手の動きを瞬時に捉えて分析し、ロボットハンドを制御することで、絶対に負けないじゃんけんロボットが実現できる。右：30fps(下)では動きの影響で画像はぶれているが、1000fps(上)ではしっかりと「チョキ」を出していることがわかる。



ロボットハンド



<https://youtu.be/Qb5UIPeFCIM>