

かなでたけお

金出武雄

(米国カーネギーメロン大学 ワイタカー記念教授,
独立行政法人 産業技術総合研究所 デジタルヒューマン研究センター・センター長)
自由な視点移動で映像を楽しむ仮想化現実技術の実用化



360 度の自由な視点から映像を楽しむ。これまでの映像放送技術では成し得なかった映画のような世界を、現実の映像から構築する「仮想化現実」の技術を確立。これらの技術をもとに開発された「EyeVision」システムは、2001 年に米国で最高の人気を誇るスポーツ、アメリカンフットボールの大会 SuperBowl XXXV にて、世界で初めてテレビ中継に採用され、映画のような任意視点の中継映像を提供することに成功、視聴者に大きな衝撃を与えた。

カーネギーメロン大学ロボット研究所へ

金出武雄は、1945 年兵庫県に生まれた。中学時代は生徒会長を務めるなど、当時からリーダーシップを発揮し、抜群の記憶力を武器に京都大学に入学、1968 年に京都大学電子工学科を卒業した。1970 年に同大学工学修士号を取得し、1974 年には、世界で最初とされる自動的顔認識システムの開発により博士号を取得。京都大学情報工学科助手を経て、1976 年に同助教授に任官。典型的な日本の大学教官の道を歩んでいた。

助手時代に京都大学の研究室を訪問してきた米国カーネギーメロン大学のアラン・ニューウエル博士との出会いが、金出の渡米のきっかけであった。ニューウエル博士は人工知能研究の権威で、当時の金出の研究にも興味を示した。ニューウエル博士は、研究員として渡米したいと熱心に訴える金出の願いを聞き入れ、人工知能の分野で有名なラジ・レディ教授を紹介、1978 年に金出は 1 年間だけながらラジ・レディ教授のもとで研究に取り組むことになった。米国での整った研究環境を実感した金出は、「日本の三倍の密度で研究ができる」と感

じ、ロボットビジョンや人工知能で先行する米国で、自分がどれだけできるか挑戦したいという思いに駆られた。

1980 年の春、工業用ロボットでは世界一といわれる日本での研究環境を捨てることに対して周囲は難色を示したものの、金出は強い決意から京都大学助教授の職を捨て、高等研究員として5年契約をカーネギーメロン大学と結んだ。「5年だけ」と家族を説得し、5歳の息子と9カ月の娘を連れて、周囲の反対を押し切っての渡米であった。

コンピュータビジョンとロボット

カーネギーメロン大学は、マサチューセッツ工科大学、スタンフォード大学と並んでいわゆる「コンピュータ・サイエンス御三家」といわれる、コンピュータ・サイエンスの世界で有名な大学である。世界中から我こそはと思う優秀な頭脳が狭き門をくぐり集まっている。そのような刺激的な環境の中で、金出は1981年に、動画像を解析し、その中での物体の動きを高速にかつ正確に追跡する方法を研究した。

図1のように、複数の対象物体を移動するカメラで撮影すると、撮影された各フレームの中では、それぞれの物体はその位置関係によって異なる方向に異なる速度で動くことになる。この物体ごとの見かけ上の動きを検出することにより、物体の移動方向や速度、さらには奥行きに関する情報などまで得ることができる。このような動画像中の見かけ上の速度を検出することは、コンピュータに人間の目と同じ働きをさせようとするコンピュータビジョンの分野において大変重要な処理である。画像中で動く物体の追跡では、その物体のエッジや模様など特徴的な点を抽出し、次のフレームでその特徴点がどこにあるかを探索する必要がある。

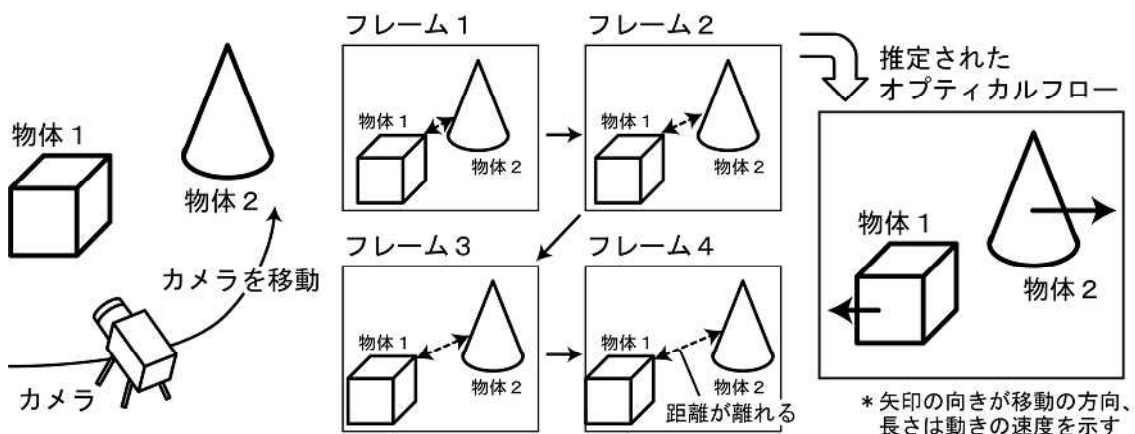


図1. 運動物体の見かけの速度推定

金出らは、この特徴点を追跡する基本的アルゴリズムとして Lucas-Kanade 法を提案した¹。図 2 のように、フレームの中から特徴となるあるサイズの領域 (テンプレート) を切り出す。そして、次のフレームで、切り出したテンプレートの周辺と類似性を評価する。類似性の評価には、2つの画像の差の二乗の合計 (差分二乗和) を用いる。つまり、差分二乗和が小さければ小さいほど類似性が高く、同一の画像であるといえる。単純には次のフレームの画像中のすべての点に対して類似性評価をすることで一致点を探索することができるはずであるが、それには多大の時間を要する。Lucas-Kanade 法では、まず前の画像での特徴点の位置に対して類似性を評価し、その微分変化量を考慮しながら、特徴点の動きを計算していくことで、一致点を高速に導き出す。これは極端に動きが速い場合には、次のフレームで特徴点がフレーム外に飛び出してしまったり、比較評価すべき範囲が広がって評価に要する時間がかかり過ぎるなどの問題が出てくるが、ビデオレートで通常の画像を取得している場合には効果的に適用できる。この方法は、MPEGなどの動画圧縮方式を使った動画画像の解析における基本中の基本的方法となっている。

さらに金出は、Lucas-Kanade 法で差分二乗和により類似性の変化量を求める際に用いる行列式の固有値が、特徴点の運動量の計算のしやすさを示すことを発見した²。つまり、そういう運動量の計算のしやすい固有値の行列式をもつ画像テンプレートは、追跡しやすい、つまり特徴と呼ぶに相応しい個所であると気付いた。それまで、物体の角や模様が特徴点となりやすいことを人は経験的に知っていたが、コンピュータが特徴点に相応しい個所を数学的に認識することは難しかった。この特徴点抽出手法と特徴点追跡手法を合わせた方法は現在、KLT(Kanade-Lucas-Tomasi)-トラッカーと呼ばれ、物体認識・追跡手法のための数あるアルゴリズムの中でも精度が良好かつ高速・ロバストな標準的手法として、現在でも多くの動画画像解析に使われる技術の基礎となっている。

¹ B. D. Lucas and T. Kanade, "An Iterative Image Registration Technique with an Application to Stereo Vision," in Proc. of Int. Joint Conf. on Artificial Intelligence, pp. 674~679, Aug. 1981.

² C. Tomasi and T. Kanade, "Detection and Tracking of Point Features," Carnegie Mellon University Technical Report, CMU-CS-91-132, Apr. 1991.

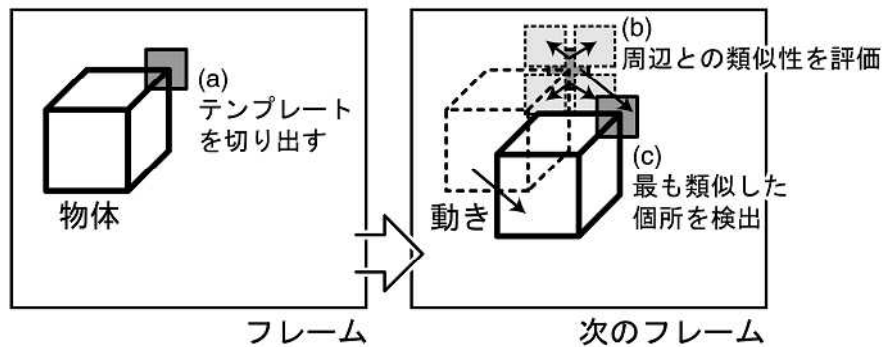


図 2. Lucas-Kanade 法における特徴点抽出と探索

ロボットのハードウェア分野においては、1981 年頃、直接駆動型マニピュレータというギアを使わないロボットの腕の概念と、世界最初のプロトタイプ開発者の一人として携わった。直接駆動型マニピュレータは、現在では一般的に使われる技術となっている。運転手なしの車、自律ヘリコプターなど数多くの自律移動ロボットや応用システムの開発プロジェクトリーダーを務めたが、中でも 1995 年には、「手をはなしてアメリカ横断」という自律走行車のデモンストレーションを行った。米国国防総省の国防高等研究事業局(DARPA)によるプロジェクトとして始まった自動車の自律走行を実現する研究は、初期は大掛かりな改造を施した車両を用いていたものの、デモンストレーションの段階では小型のバンに外付けのモジュールを取り付ける程度で実現できるに至っていた。デモンストレーションでは 2～3 人が乗車し、常に危険がないかを監視しながらではあったが、ピッツバーグから西部海岸のサンディエゴまでの約 5,000km の 98.2%を、自律的に計算機の「目」によって道路の位置を見つけながら走行することに成功した。残り 1.8%の中には、実際に飛び出しなどの危険回避のためにハンドルを握ることもあったが、ほとんどは工事中で道路に目印となるものがない場合や、給油のためガソリンスタンドに立ち寄ったときなどであり、このデモンストレーションによって自律走行技術の初期の成功を世の中に印象付けた。

1995 年以降、画像の特徴抽出や運動物体の動き推定、距離計測、ロボット制御など数多くの研究成果の集大成の一つとして、金出らが「仮想化現実(Virtualized Reality)」と名付けた全く新しい視覚メディアの実現への挑戦が始まった。

「仮想化現実 (Virtualized Reality)」の世界

金出の率いる米国カーネギーメロン大学ロボット研究所は、いわゆる仮想現実 (Virtual Reality)とは全く別のコンセプトに基づいて、「仮想化された現実世界」をユーザへ提供することを目的とした、「仮想化現実 (Virtualized Reality)」の概念を提案した。一般に馴染みの深い「仮想現実 (Virtual Reality)」とは、現実世界に近い状況を、コンピュータグラフィックスを使

って構築し、その中で現実世界では実現の難しい状況をシミュレーションすることである。例えば、航空機パイロットの養成などに使われている飛行シミュレーションなどは仮想現実の典型だろう。ゲームや映像アミューズメントのように、実在しない世界での臨場感溢れる体験を提供するのに用いられている。これはすなわち、「現実に近い仮想世界」といえる。

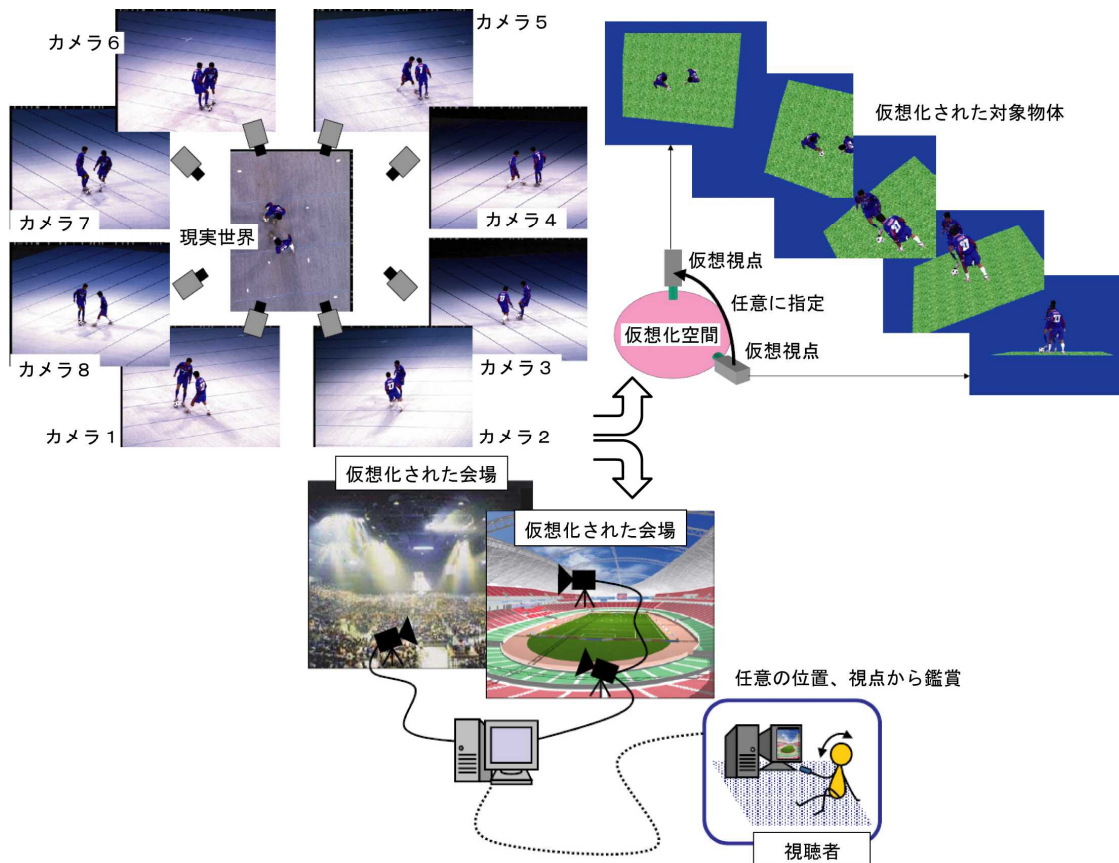


図 3. 仮想化された現実世界: Virtualized Reality³

一方、金出らの提案する「仮想化現実 (Virtualized Reality)」では、現実世界から得られた映像情報を元にして、その実在する現実世界の中を自由に動き回れる「仮想化された現実世界」を創り出すことが目的であった。これは、スポーツや劇場、手術のシーンというような時間的に変化する出来事に対して、何人ものユーザが時間的にも空間的にも全く任意の位置から、あたかもそこにいるがごとく振舞うことのできる、全く新しい視覚メディアであった。図3のように、現実世界を多数のカメラで周りから捉え、デジタル化されたビデオ系列を処理するこ

³ 北原 格, 大田 友一, 斎藤 英雄, 秋道 慎志, 尾野 徹, 金出 武雄, “大規模空間における多視点映像の撮影と自由視点映像の生成,” 映像情報メディア学会学会誌, Vol. 56, No.8, pp. 1328~1333, 2002 年 8 月.

とで、その出来事の「時間」、「空間」、「見え方」という多次元の表現を創り出す試みである。高度なコンピュータグラフィックス(CG)で話題をさらった 1999 年アメリカ公開映画「マトリックス」の中で用いられ、多くの人の知るところとなった任意視点空間の映像表現手段を、現実の動きの中で創り出し提供することができるようになれば、視聴者は自分の好みの位置や角度で現実世界の振る舞いを知ることができるようになる。

メッセージのある研究を

金出の研究に対する取り組みのモットーとして「メッセージのある研究をしろ」というのがある。「『何だか分からないが、難しいことができてしまった』というのは、一見優れた研究のようにも見えるものだが、人の参考にはあまりならないことが多い。研究の成果に対して、それを聞いた人が『それを使えばこんなアイデアが実現できるかもしれない』、『その程度のことだったら自分でもできる』、『それが成功するならば、こんな方法もあるのではないか』といった、驚きや触発を受けるような研究をするべきである。」ということである。こういった考え方は、金出が博士課程で研究に行き詰まった経験から生まれている。

小さな頃からいわゆる優等生であった金出は、そのプレッシャーのようなものから、博士課程に入ってから早く格好のいいことをしなければと考えていた。数学的でちょっと格好の良さそうな研究論文を読むと、すぐに同じようなことができそうだと思って取り組んでみる。しかし、研究は試験と違い、その問題が解いて価値のある問題か、そもそも答えがあるのかも分かっていない。そう簡単にできるものではなく、焦りから次の課題、そしてまた次へと取り組んでみるものの、何もできずに時間が経っていった。

そんな中、当時の長尾真助教授(前京都大学学長)に「金出君、もう少し具体的なことをやったら」と言われ、1970 年の大阪世界万博の際に集められた 1000 人の顔がデジタル化されたデータベースの存在と、このデータベースの画像全てをうまく処理し認識できたらいいのでは、と教えられた。金出はもっと理論的な研究がしたいという気持ちを抑え、この研究に取り組み始めた。すると、さまざまな課題に行き詰まりはしたものの、目標が具体的であることも手伝って、最後には人の顔を画像の入力から抽出し、さらに判別まで自動的に行うことが可能になった。そして、それは世界で最初の自動顔認識システムとして知られる研究となった。これがきっかけとなって、金出は研究や開発というものは、具体的な目標を必ず持ち、そして多くの人にメッセージを与えられるものであるべきという考え方を持つようになった。

「仮想化現実」の概念では、バスケットボールやブロードウェイの熱狂的なファンが、試合や舞台を自分たちの好きな座席から見ることができ、それがリアルタイムに処理されれば、選手や俳優の動きに合わせて座席を変えることすら可能となる。金出は「仮想化現実の世界」の

研究を、「Let's watch the NBA on the court.」(NBA のゲームをコートの中で見よう)⁴というメッセージを投げかけることで考えた。

なぜ人の眼は2眼なのか

カーネギーメロン大学での金出の取り組みの一つとして、人の眼と同じように複数のカメラを使って立体的な映像を取得するステレオ視の問題があった。ステレオ視は、私たち人間が左右の眼の見え方のずれから奥行きを知ることができるように、それぞれのカメラによって得られた画像の微妙なずれから距離推定を行う方法である。金出らは、特にビデオレート・ステレオマシンという、動く物体の距離情報をビデオレートの高速でしかも信頼性良く取得するカメラの実現に取り組んでいた。それまで、3次元視覚というものは、人間の目のように2台のカメラを使ってやるというのが常識的な考えであった。しかし金出らは、ビデオレート・ステレオマシンに取り組むうちに、どうやら人の眼と同じ2眼での距離推定では信頼性に限界があるということに気付く。

2眼ステレオ視では、図4(a)のようにある物体を2つのカメラで撮影したときの見え方の違い(視差)から、距離を推定することができる。しかし、2つのカメラが水平に設置されている場合、水平線に対しては図4(b)のように、その端点付近以外は左右で同じ画像となり、距離の推定ができない。人は水平線の端を見て、その視差の違いからそこまでの距離を知り、水平線の中の部分も同じか滑らかにつながる距離にあるだろうと予想する。しかし、計算機でこのような予想をするのは容易ではないし、また、もし端が見えていなければ、本質的に距離を推定することすらできない。また、図4(c)のように、柵のような周期性をもった物体に対しても、右目画像のどの柵と左目画像のどの柵とが対応するかはあいまいである。金出はそう考えていくにつれ、人の眼は水平方向の2眼で十分であったのかもしれないが、コンピュータの視覚としては2眼のシステムにこだわるべきでない、あるいはそれは不適切なのではないかという結論に至った。そこで金出らは、物体の距離をあいまい性なく正しく知るには2眼ではできないということを数学的に示し、3つ以上の眼を用いて複数の視差から距離情報を得る複数視差ステレオ理論(Multi-baseline stereo 理論)を提唱した。

ステレオ視のカメラの数を2つに拘る必要はないという、当時の常識から脱却した発想は、その後の仮想化現実の世界を実現する第一歩であった。「みんなの言うことの反対をすればいい」という発想が、独創的なアイデアを生む秘訣であると金出は半分冗談半分本気で言う。

⁴金出 武雄, 「素人のように考え、玄人として実行する」 PHP 研究所, 2003 年 6 月。

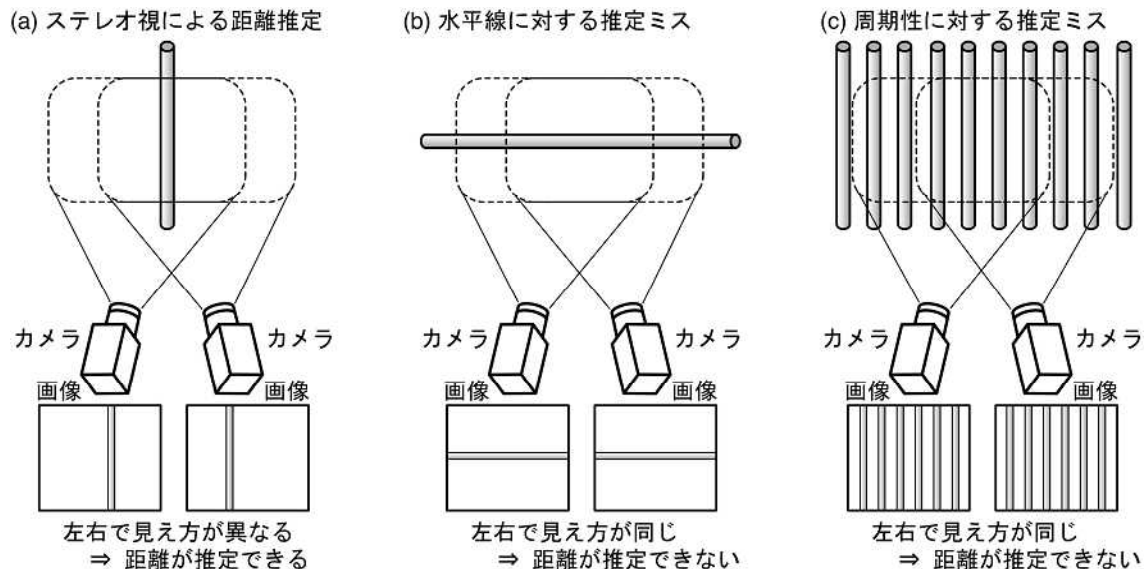
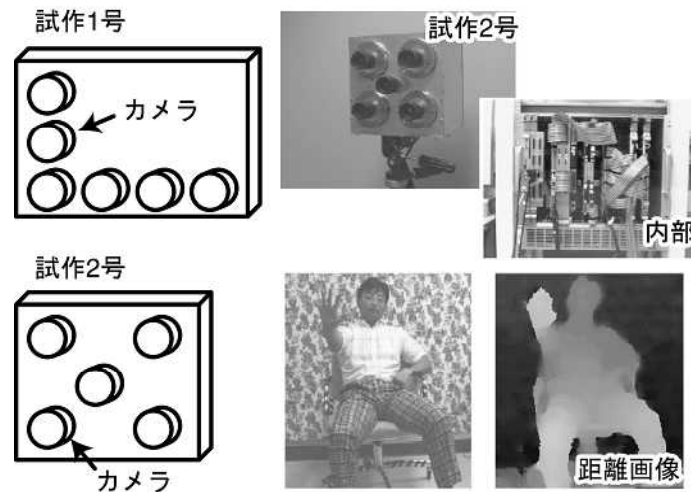


図 4. 2つのカメラで構成されたステレオ視による距離推定の失敗例

金出は1992年から1995年にかけて、DARPAから2億円以上の予算を得て、複数視差ステレオ理論による実時間多眼ステレオカメラの実現に取り組んだ。試作1号は、図5に示すような6つのカメラを縦に3つ、横に4つ並べたものであった。縦3横4の構成には理論的な根拠がなかったが、当時の金出は、部屋をぐるりと見渡したら縦方向のものと横方向のものが4対3の割合であったと冗談まじりに主張していた。これは動きのあるシーンに対して精度の高い3次元画像をリアルタイムで生成できる、世界で最初の完全ビデオレート・ステレオマシンの開発であった。

理論が確立するにつれ、カメラを斜め方向に配置することで縦横両方に対して全てのカメラが寄与できることや、周期的な物体に対して誤検出の発生しにくいカメラ配置などが明らかとなり、5つのカメラを用いた試作2号を作り上げた。図5のように、シーン上の物体の距離を濃淡で表現する距離画像をビデオレートで得られる装置であった。

図 5. 複数視差ステレオ理論による距離計測カメラ⁵

現実世界の中を自由に動き回れる仮想化現実「3D Room」

複数視差ステレオ理論による実時間多眼ステレオカメラの成功は、金出に新しい可能性を感じさせた。金出は、カメラ本来の機能であるシーン画像の取得と、実時間多眼ステレオカメラによって得られた物体の形状を意味する距離画像とを用いれば、いながらにして多くの人にその物体を異なる自由な方向から観察させることができるのではないかと考えるようになった。距離情報から物体の形状をモデル化してしまえば、コンピュータグラフィックを用いて異なる方向からの画像を合成することが可能であることはあきらかである。しかし、カメラが設置してある位置と異なる視点からの物体の画像を合成しようとする、当然、もとのカメラから見えていない部分は欠落してしまう。金出は、自分たちが提案した複数視差ステレオ理論は、本来カメラの台数を限りなく増やすことのできる理論であり、それら多くのカメラから得られた画像や距離情報を扱えることに注目した。つまり、数台のカメラで構成していたビデオレートステレオマシンを、数10台といったシステムに拡張し、物体をさまざまな方向から観測すれば、自由な視点から物体を隅々まで観察することが可能になり、そして、ユーザにまるでその空間に実際にいるかのような感覚を味わわせることができるはずだと考えた。これが、現実世界から得られた映像情報をもとにして、その現実世界の中を自由に動き回れる仮想化された現実世界、「仮想化現実」のアイデアのきっかけであった。

1995年から、「仮想化現実」の世界を実際に創り上げることを目的として、金出らは51台のカメラを持つ「3D Room」(図 6)という施設をカーネギーメロン大学内に構築し、予備的な実験

⁵ T. Kanade, A. Yoshida, K. Oda, H. Kano and M. Tanaka, "A Stereo Machine for Video-rate Dense Depth Mapping and Its New Applications," in Proc. of Computer Vision and Pattern Recognition Conference (CVPR), Jun. 1996.

を重ね始めた。この「3D Room」の中に人などの対象物体が入ると、それぞれのカメラが様々な方向から映像を取得し、その映像をもとに位置や形状を推定できる。10年近くにわたって、年間5,000万円前後の予算と、3～4人の学生とスタッフを常に割り当て、東芝や松下、ソニー、インテル、DEC など数多くの企業から支援を受けながら続けられた「仮想化現実」プロジェクトのスタートであった。

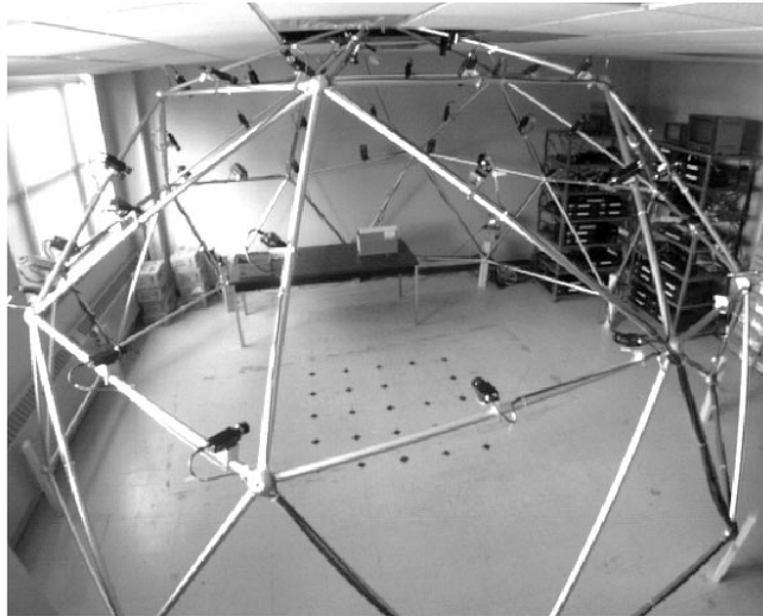


図 6. 51 台のカメラを周囲に配置した 3D Room⁶

⁶ P. Rander, P. J. Narayanan, and T. Kanade, “Virtualized Reality: Constructing Time-Varying Virtual Worlds From Real World Events,” in Proc. of IEEE Visualization, pp. 277 – 283, Oct. 1997.

図 7. 複数の方向から撮影された対象⁷

図 7 のように、「3D Room」内のある対象を複数のカメラ（この場合は 6 台）で撮影した場合、少しずつ向きの異なる対象の画像が得られることになる。コンピュータ上では、画像の濃淡やエッジ、模様などの特徴を抽出し、それぞれの画像間で対応をとることで、その見え方の違いを扱えるようになる。この見え方の違いから図 8(a)のような奥行き情報を得ることができる。この奥行き画像は、白の領域がカメラに近い部分であることを示し、色が暗くなるほどカメラから遠い部分であることを示している。対象を仮想化するには、計測誤差を抑えるための補正やフィルタ処理などが随時必要となるが、主な処理の流れは図 8 の(a)から(c)に示すとおりとなる。

⁷ T. Kanade, P. J. Narayanan, P. W. Rander, “Virtualized Reality: Concepts and Early Results,” IEEE Workshop on the Representation of Visual Scenes, Jun. 1995.

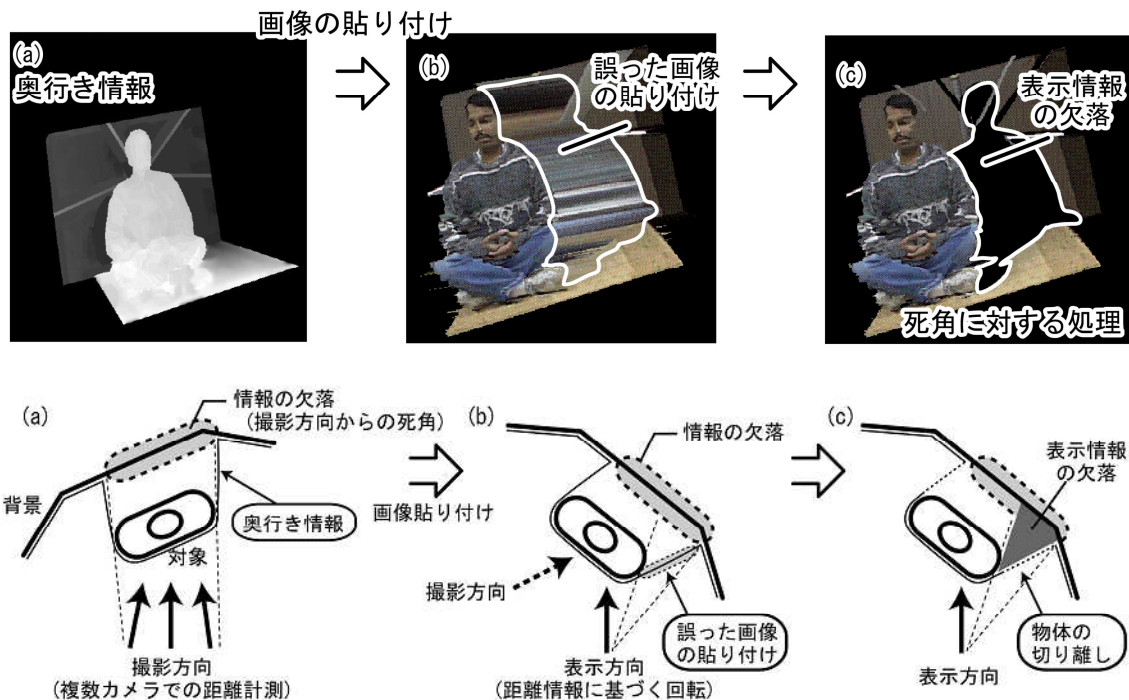


図 8. 推定された奥行き情報と、仮想化された対象

まず、図 8(a)の奥行き情報と対象の模様との対応をとる。すなわち、得られた奥行き画像に通常の画像を貼り付ける処理であり、これによって、シーンの 2 次元的な画像情報と空間的な位置情報とを対応付けることができる。この対応付けは指定された視点からの対象物体の見え方を、角度などを考慮して正確に得るために必要となる。指定された視点からの対象の見え方を取得するとき、撮影方向と同じ視点からの表示であれば問題ないが、図 8(b)のように異なる視点に表示方向を移動すると、撮影点からは見えていなかった死角の部分(情報の欠落部分)が視界に入ってしまう。この死角部分がどこであるかを図 8(c)のように正確に計算し、背景と物体の切れ目をあきらかにする。これらの処理によって、そこに最も近い位置にあるカメラを基準にして対象の仮想化を行い、指定された視点に回転させることができる。その際、基準カメラからの死角になっている部分は欠けてしまうが、次に近いカメラを基準とした仮想化対象を、指定の視点に回転させた画像から補うという処理で、死角領域のない任意視点画像を生成することができるのである。

図 9 の画像を例にとると、ある基準カメラから撮影された画像と距離情報をもとに、任意の視点からの映像を生成する。このとき、図 9(a)のように、対象の左後ろの背景部分が基準カメラからは死角となっているため、情報が欠落してしまう。この欠落部分を補うために、別のカメラから取得した画像と距離情報を用いる。図 9(b)のように、異なる位置にあるカメラから撮影した場合、視点の変更によって欠落する情報の場所が異なる。たとえば、図 9(a)では対象となる人物の左背後が欠落しているのに対して、図 9(b)では右背後が欠落している。この 2 つの

合成画像から、図 9(c)のように、任意の視点からの画像を合成する際に、図 9(a)の欠落部分を図 9(b)の情報を用いて補完すれば、実際にはカメラが設置されていない任意の視点からでも、情報欠落がない自然な合成画像を作ることができる。

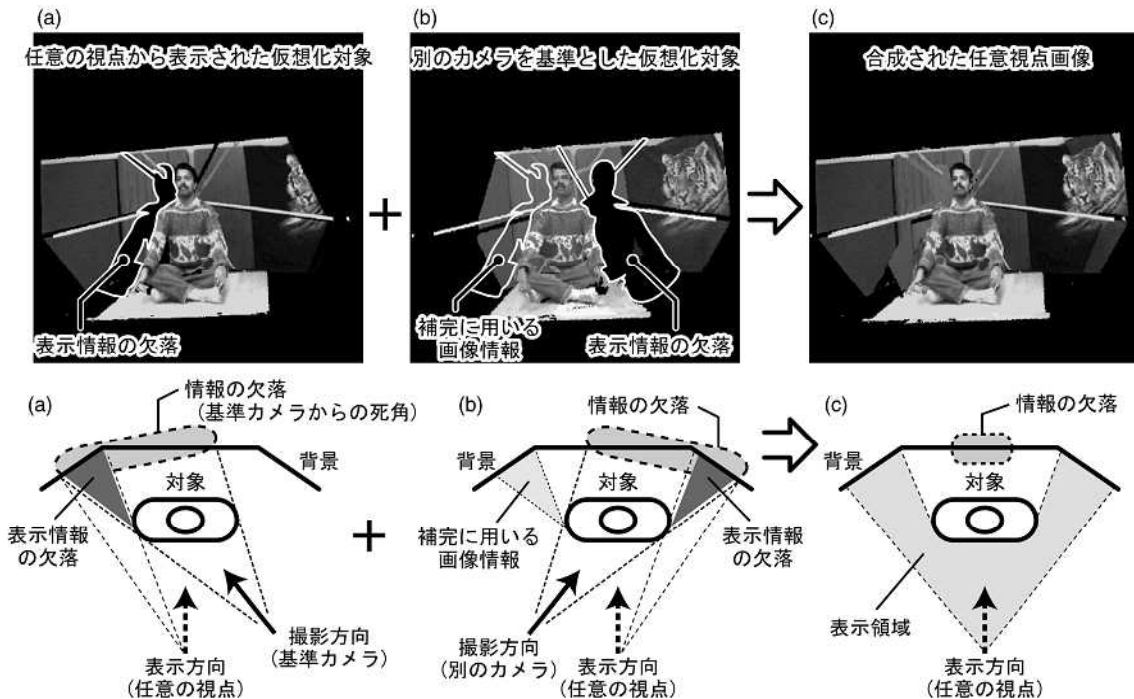


図 9. 死角による欠落個所のテクスチャ補正

このようにして、1990 年代の後半、金出らは仮想化現実という概念の実現と完成を目指し、多くの理論的・実験的成果を重ね、この仮想化現実の技術が、スポーツ観戦、人のコミュニケーション、デモを用いた訓練などの分野で、従来とは全く異なる新しい方法を生み出すことができるという確信を抱くようになる。また、金出らの仮想化現実に関する数多くの研究成果報告は、その応用分野と実現可能性を示すのに十分なものとして受け入れられ、世界中で多数の同様な研究活動を生み出していった。金出がこの研究に託したメッセージは多くの研究者たちに届いたといえる。

「EyeVision」プロジェクトが SuperBowl 中継の歴史を変えた

1999 年の秋、米国の主要テレビ放送局の一つである CBS⁸が、「SuperBowl XXXV」中継の打ち合わせを行っていた⁹。この打ち合わせが、当時、金出らが取り組んでいた仮想化現実の実例が多くの人の目に触れる大きなきっかけとなった。「SuperBowl」は米国で最高の人気を

⁸ Columbia Broadcasting System, Inc.

⁹ “Instant replay gets ‘Matrix’ – like view of key plays,” USA TODAY, Jan. 24, 2004.

誇るスポーツであるアメリカンフットボールのチャンピオンシップゲームであり、全米中が注目するイベントである。全世界でその視聴者が1億人を超えるとされる「SuperBowl」の2001年1月の放映に向けて、視聴者が眼を見張るような新しい映像放映技術を開発するというプロジェクトが立ち上がった。

打ち合わせに参加した CBS の技術者たちは、当時全米で話題となっていた映画「マトリックス」のような、映像のあるシーンで停止させ、自由な視点にぐりと動かす表現が可能なのではないか考えた。つまり、通常は位置的に固定されたカメラで撮った映像が時間軸に添って流れるが、これを逆にして、時間軸を止めてカメラ位置を動かすことで生まれる特殊な映像効果を、フットボール中継に応用しようと考えたのである。しかし、それはコンピュータグラフィックスを駆使した映画の世界ではなく、ライブ放映のリプレイとしてである。

このプロジェクト「EyeVision」の開発は、CBS スポーツ代表の Sean McManus による250万ドルの予算承認を受けて、CBS 技術者たちによって検討が始められた。しかし、対象の周囲を自由に視点が動き回るのに、いったい何台のカメラを、スタジアムのどこに配置したらいいのか。この新しい取り組みは彼らにとって多くの問題が山積みという状況であった。

そこで、カーネギーメロン大学ロボット研究所で当時所長をしていた金出に、秘密裏に「EyeVision」開発に関する協力の話が舞い込んできた。CBS からの話はさほど詳細なものではなかったが、アメリカンフットボールの試合で映画「マトリックス」のような、カメラ(視点)が自在に動き回れるような映像表現を実現できないかという趣旨のものであり、「仮想化現実」に取り組んでいた金出は、その経験から技術的にはすぐさま実現可能であると判断し、「そんなことならやってやる」と請け負うことにした。

最初の6カ月間はシステムの実現に必要な機材を探すのに費やされた。というのも、広いスタジアムで注目する出来事を追いかけるには、離れた場所に設置された多くのカメラを1つの対象に向けて高速・正確に追従しフォーカスできるようカメラの向きなどを自由に操るロボットが必要だったからである。また、対象物との距離に応じてズームをすることで、各カメラの映像が同じ大きさにならなくてはいけない。特に、被写体の場所や対象は時々刻々変化し、事前の予測が不可能であることが問題をさらに難しくした。基本的な制御は、マスターカメラがプレーを追いかけ、他のカメラはマスターカメラが注目している場所を自動的に追従するという方法である。しかし、図10のように、マスターカメラの注目している場所を他のカメラが自動的に追従しようとしても、マスターカメラの向きからだけではその視線上のどこに注目点があるのかが分からない。つまり、マスターカメラの向きからだけではプレーの高さ方向に対して追従ができない。金出らは、マスターカメラのフォーカスや、カメラで撮影された物体の大きさなどから高さを予想するプログラムを実装することで問題の解決を図り、さらにカメラマンに高さ方向のレバーを持たせることで不意な状況にも対応できるようにした。

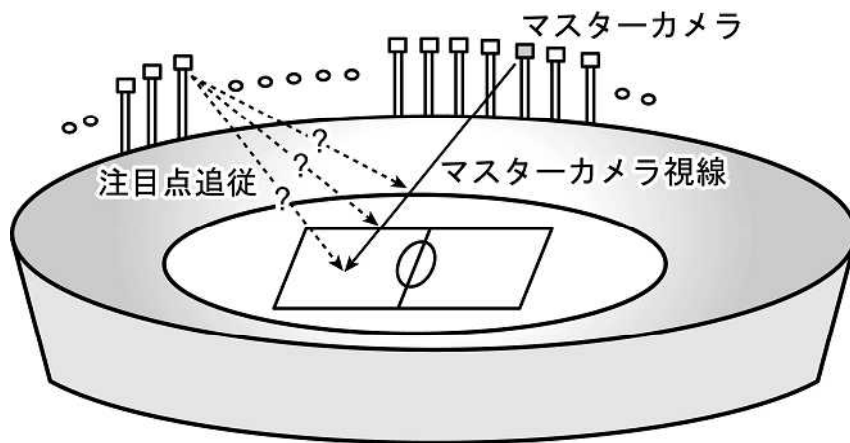


図 10. マスターカメラ視線上の注目点に追従するカメラの問題

さらに、もう一つの大きな問題は全てのカメラが 1 点を注視することを制御できるようにカメラの位置と視線方向の関係を正確にモデル化する作業だった。これまで実験的に用いてきた「3D Room」では、個々のカメラを順番に補正してから対象を撮影していた。あるカメラの位置がずれた場合は、そのカメラを補正し直してシステムをリセットする。プロトタイプである「3D Room」では、実際問題として補正後にカメラの位置がずれることはほとんどなかったし、このような補正方法でも問題はなかった。システムのリセットに 10 分程を度要しても構わなかったからである。しかし、アメリカンフットボールの試合はそうはいかなかった。大勢の観客がいるスタジアムは、時間とともに歪んでいく。ほんの 10 分の 1 度でもカメラの角度がずれれば、全てのカメラが 1 点を注視することができなくなる。また、再補正の度にシステムをリセットしては、時々刻々と展開していくゲームの様子を伝えることは到底できない。金出らは、システムの観点からの見直しを迫られることになった。金出は、数 10 台というカメラを用いて実現する「EyeVision」システムの特徴を生かして、仮に数台のカメラが動作していなくても、他のカメラで動作していないカメラの方向を補うことが可能であると考えた。つまり、各カメラを独立に切り離せるようにシステムを再設計することで、カメラの大半が稼動していれば、数台のカメラを再補正している最中でも任意の視点からの映像を提供できるようにしたのである。常時 8 割ほどのカメラが稼動している状態で、位置がずれたカメラを次々と切り離し、再度補正をした後でシステムに加えることで、時間とともに変わっていくカメラの位置関係の問題を解決した。

これらのさまざまな検討とこれまで金出らのグループで培ってきた技術をもとに、33 台のロ

ボットカメラ¹²をスタジアムの上部デッキに設置し、DVD録画装置と9万フィートの光ファイバケーブルが期待通りに稼動すれば、視聴者が選手のある瞬間の動きを270度の視点から再生し、観ることができるようになるという結論に至った。そして、これらの条件を満たすためのロボットカメラに使うロボットアームは、車の製造ラインに使われていた最も精度の高い1台2万ドルのものを三菱重工から調達した。2000年秋に12台のロボットカメラを用いて「EyeVision」のプロトタイプを設置、見事に任意視点の仮想化された映像を提供することに成功した。試合中のタッチダウンの判定などにおいても決定的な判断材料を提供できることを実証し、「SuperBowl」の主催者であるNFL(National Football League)からも承認を得られることになった。

CBS スタッフとの共同作業は、時間が差し迫っていたこともあり衝突も多かった。特に、放送業界における新しい技術に対する慎重なグループの存在や、プロデューサーの意志に沿わないものは一切受け入れられない風土などに戸惑うことも多かった。技術的な可能性をどれだけ説明しても、その場で見せたデモンストレーション以上のことは受け入れてもらえなかった。例えば、基本的な機能だけでなく、一層面白いと思われる機能も開発し、使ってもらいたかったのでデモンストレーションをした。ただ、そのプロトタイプでのデモンストレーションはいったん放映されたビデオ映像をデジタル化していたため画質が落ちていた。すると、「実際の放映では元となる映像を直接扱えるため、高い画質が実現できる」と何度説明しても、「こんな画質では使い物にならない」と拒絶された。完成度の高いものを見せなければ納得してもらえないことを実感した。2001年、「SuperBowl XXXV」の会場であるレイモンド・ジェームス スタジアムに設置された33台のカメラによる「EyeVision」システムは、小規模な視聴者を対象とする放映で試験を行うことなく、直接1億3千万人の視聴者に向けて、仮想化現実による全く新しい任意視点での映像を提供することに成功した(図11)。任意視点のスポーツ中継システム「EyeVision」は、CBSによる1963年の「インスタント リプレイ(瞬時再生)」登場以来の、スポーツ中継における最も飛躍的な進歩であるといわれている。そして、この「SuperBowl」での放映を皮切りに、NBA(全米プロバスケットボール)や欧州のサッカーなど、さまざまなスポーツ中継に「EyeVision」システムが利用され始めることとなった。

¹² 上下左右に自由に首を振ることができ、マスターカメラの動きに追従して対象を追いかけることができるカメラ

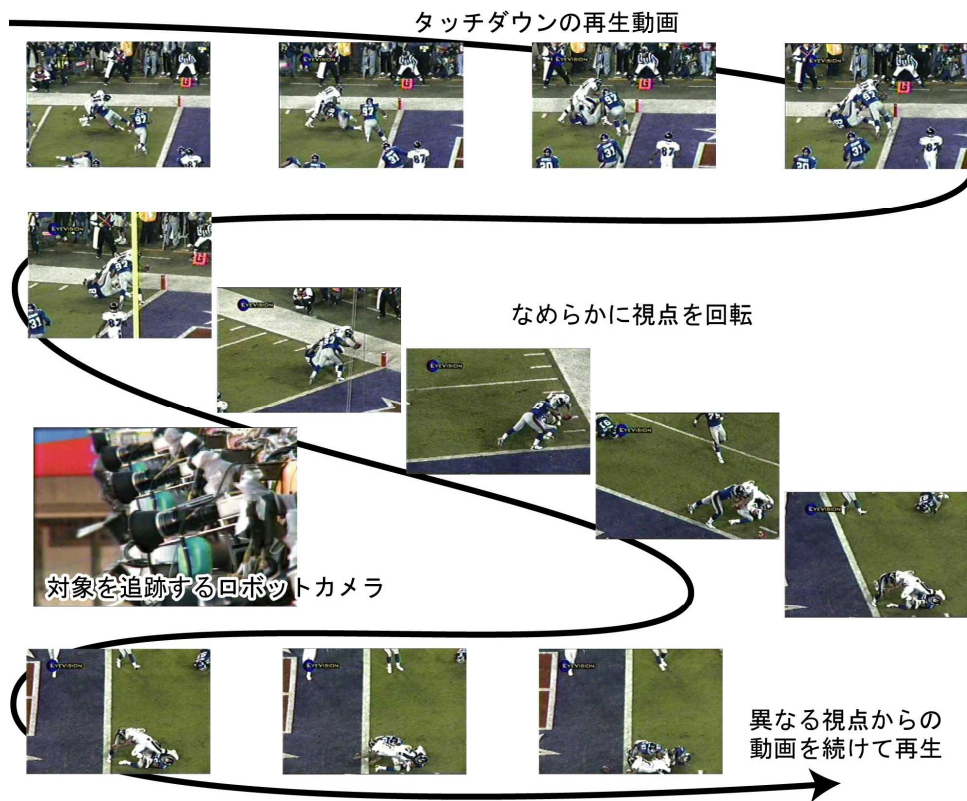


図 11. 「SuperBowl XXXV」において「EyeVision」を用いて放映された任意視点再生¹³

神宮の森に「EyeVision」登場

「SuperBowl」での放映で成功を収めた「EyeVision」システムは、ロボットカメラを提供した三菱重工とフジテレビにより、金出らの技術指導のもとで日本上陸を果たす。2001 年 7 月、「ヤクルト vs 巨人」の 3 連戦に備えて神宮球場で準備が進められた。「SuperBowl」と同様、コントロールルームから遠隔操作で向きを自在に変更でき、任意に選択されたマスターカメラに追従して、他の全てのカメラが同一被写体をあらゆる角度から同時に撮影する。例えば、マスターカメラが一人のプレーヤを追えば全てのカメラがそのプレーヤを追ひ、ボールであればそれを追うことができる。

神宮球場では、図 12 に示すように、ネット裏に 6 台、1 塁側からバックスクリーンにかけて 24 台の計 30 台のカメラが設置され、バッターがボールを打つ瞬間を一塁側からぐるりとバックネット側からの視点まで、まるでクレーンでカメラを移動させたかのような映画のような映像を提供することに成功した。盗塁をしたときのスタジアム全体の雰囲気や、ホームランを打つ瞬間の打者のタイミング、クロスプレーでの各選手の動きなど、よりリアルに表現された映像で視聴者を魅了した。それは、フィギュアスケートやラグビー、劇場や監視システムなど、幅広い応

¹³ カーネギーメロン大学 <http://www.ri.cmu.edu/events/sb35/tksuperbowl.html>

用へ広がっていく可能性を示すのに十分なデビューであった。

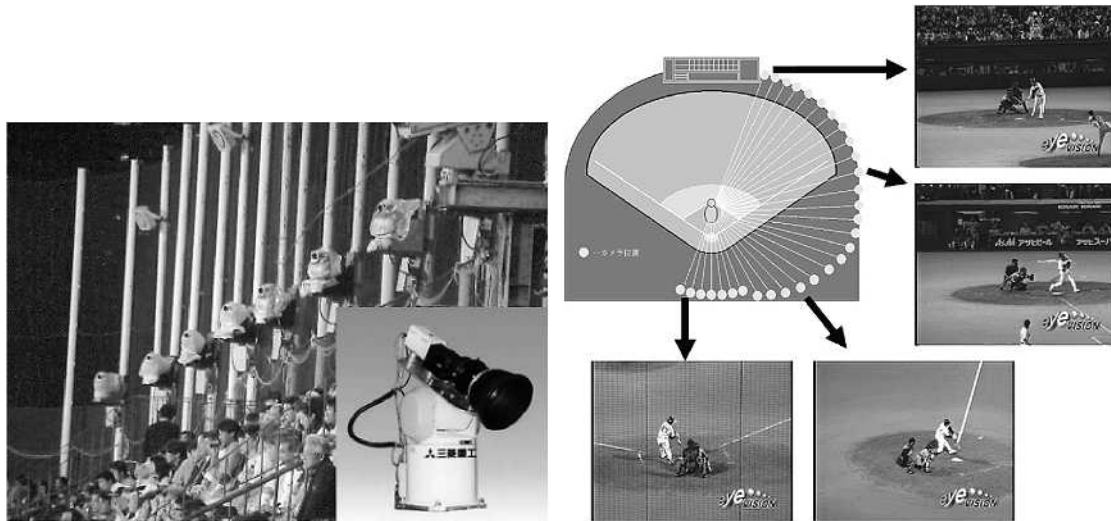


図 12. 神宮球場に設置された EyeVision システムと放映された映像¹⁴

EyeVision は、今後も注目の集まるスポーツ競技での魅力的な映像の提供が期待されている。EyeVision をはじめとする「仮想化現実」システムは、NBA のコート上で観戦する、演劇の舞台の上から観賞する、動物園の檻の中に入るといったエンターテインメント用途だけでなく、映像を用いた教育や、医療用途、そして新しい映像記録の在り方を提唱できる可能性を秘めている。

産総研で人を優雅に助けるロボットの実現を目指す

カーネギーメロン大学 ロボット研究所での研究活動の他にも、金出は3社のベンチャー企業を立ち上げている。1993年には周囲を360度スキャンして3次元画像を作り出す高速センサを開発し、当時金出の学生とともに K²T という会社を設立。設立当初は、開発した装置の約4000万円という価格が高すぎて伸び悩んだものの、開発した装置を使って建物などを計測するサービスを始めることで軌道に乗り、現在では40人ほどの会社となっている。他にも、1998年には3次元メディアシステムなどを開発・販売するザクセル社、2000年には手術用ロボット会社のボーン・クラフト社を起業するのに関わっている。

さらに2001年、独立行政法人 産業技術総合研究所にデジタルヒューマン研究ラボ（2003年より現在、センター）を設立し、ラボ長（現在、センター長）に就任した。長年ロボットを研究してきた金出は、人自身があらゆるシステムの中で最も理解されていない要素で

¹⁴ 辻 洋, 下山 公宏, 藤田 淳, 川内 直人, 「Eye Vision™ 用パンチルトに応用されたロボット技術について」 三菱重工技報, Vol. 40, No. 5, 2003 年 9 月.

あり、今後ロボットを進化させ、人と共存させていくためには人間の研究が欠かせないという考えに至っていた。デジタルヒューマン研究センターで金出の掲げた目標は、人間の複雑かつ洗練された機能とその行動に関して、記述・分析・予測などのモデルを開発することである。このような人間のモデル化技術は、人間に関わるさまざまなシステムを設計・運用するにあたって、個人との高い親和性を持たせることができるようになる。金出は産業技術総合研究所で「人を知るデジタルヒューマン」、「人を見守るデジタルヒューマン」、「人に合わせるデジタルヒューマン」、「デジタルヒューマンプラットフォーム」の4つのテーマを設定し、図13のように、人の動作のシミュレーション、超音波の無線タグを用いて人の生活を管理するシステム、人の形や動きをつぶさに計測することでの衣服など人の使う製品の設計手法、そして人の機能を実現するヒューマノイド技術など、数多くの「ヒト」と「モノ」のリンクに注力して研究開発を進めている。金出の目指すロボットは「人をgracious(優雅)に助けるロボット」である。ホテルのドアマンは、急いでいる訪問者には素早くドアを開け、そうでなければゆっくりとドアを開ける。自動ドアにそのような気の利いた挙動をさせることは現在は難しい。しかし、複雑な人の動きや心理をロボットが理解し、人の存在を最優先に考えた優雅なロボットの実現を目指しているのである。

日米のロボット研究分野での掛け橋として、日本と米国を行き来する多忙な日々の中で、最近ではさらにクオリティー・オブ・ライフ(生活の質)工学研究所を設立したいと奔走している。金出が「優雅なロボット」と表現するヒューマノイドロボット(人型ロボット)の新しい境地を探りたいという思いからだ。現在も精力的に研究に取り組む金出にとって、現実の中にある問題に挑戦する知的冒険は今も続いている。



図 13. 産業技術総合研究所 デジタルヒューマン研究センターでの取り組み例¹⁵

略歴

- 1968 京都大学電子工学科卒
- 1973 同大情報工学科 助手
- 1974 京都大学にて工学博士取得
- 1976 同大情報工学科 助教授
- 1980 米国カーネギーメロン大学 ロボット研究所 高等研究員
- 1982 同大 テニユア付準教授
- 1985 同大 正教授
- 1992 同大ロボット研究所所長
- 1993 同大 U.A. and Helen Whitaker 記念 教授
- 1995 Joseph Engelberger 賞 受賞
- 1997 JARA 賞 受賞
- 1997 米国工学アカデミー 外国特別会員選出
- 1998 米国カーネギーメロン大学大 U.A. and Helen Whitaker 記念 全学教授
- 2000 C&C 賞 受賞
- 2001 産業技術総合研究所 デジタルヒューマン研究センター センター長（兼任、非常勤）

¹⁵ 産業技術総合研究所 <http://www.aist.go.jp/>

2004 情報科学技術フォーラム船井業績賞 受賞

2004 全米芸術科学アカデミー 会員選出