



92年9月号 NO. 154 別刷

題 名

ROBOT VISION

1992

著 者 名

TAKEO KANADE

新技術コミュニケーションズ

6

6

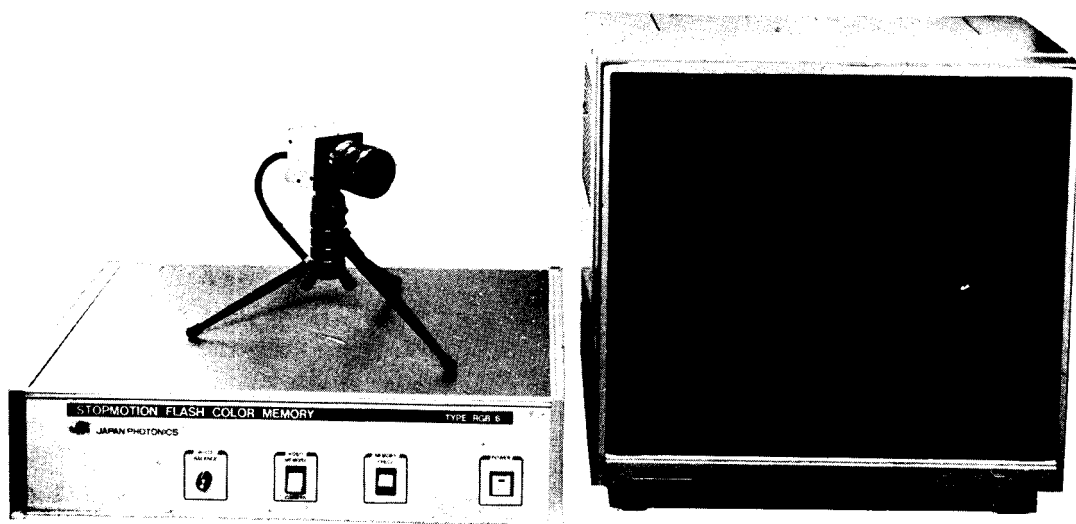
新製品 RGB出力CCD三板式カラーカメラ使用!!

高速度瞬間カラー画像記録装置

STOPMOTION COLOR FLASHMEMORY TYPE RGB-6

高速度瞬間現像の動作分析に!!生産ラインの目視検査に!!

- ☆ 超高速度動作分析システム!
- ☆ 1000コマ/秒の連続多重カラー撮像!
- ☆ コンピューターによる動作分析処理機能!



あらゆる産業の省力化・効率アップ・高品質化に役立ちます。

RGB映像出力のCCDカラーカメラに、RGB映像対応のフレームメモリーを内蔵した、高速度瞬間カラー画像記録装置です。

照明にストロボ光源の採用で、超高速度現像の動作解析が鮮明にしかもカラー映像で表示できます。



止めないで止めて見るロボットの目

株式会社

ジャパンフォトニクス

〒108 東京都港区白金5-12-24 和光ビル ●本社 803-3441-9372 FAX03-3441-9827

●開発研究所 803-3446-2700

ロボットビジョンと自律走行車 ——チップからシステム作りまで 情報入力の原点に戻って

ゲスト：金出 武雄

(カーネギーメロン大学)

プロローグ

昨年まで、「画像処理・認識の実践講座」を全28回、約3年半も続けておきながら、本インタビュー・シリーズでは画像認識・コンピュータビジョンを正面から取り上げてこなかった。一段マクロな視点に立ったこのシリーズでは、研究の最前線的话题を垣上に乗せるか、実利用の実態を鳥瞰するか、人選についても大いに迷った。結局、フロンティアの名前に相応しいゲストをとということで、米国カーネギーメロン大学(CMU)の金出武雄教授にお願いした。学術的なビジョン研究でも超一流であり、かつ大掛かりなプロジェクトの推進役であるという2つの立場が、どういう発言になって飛び出すかを期待した。

金出氏は、小生とは同窓の2年先輩であり、学生時代から秀才の誉れ高いオールラウンド・プレイヤーである。それでも、米国のコンピュータサイエンス御三家の1つであるカーネギーメロン大学へと転身され、ビジョン界の寵児として大活躍されている様子を見るにつけ、いつもわれわれ日本人の誇りと感じている。抜群の語学力に加えて、氏の弁舌爽やかな講演の素晴らしさも特筆に値する。当日は、“Shape and Motion without Depth: A Factorization Method”と題する極めてスマートな研究内容を紹介された講演の後、小生のオフィスでのインタビューに入った。(♣Dr. SPIDER)

出来合いのシステムじゃ役に立たない

♡北村素子 今日、コンピュータビジョン分野の代表として金出先生に来ていただいています。

す。あとで、自律走行車のプロジェクトのお話やビデオもあるようです。まずお2人でコンピュータビジョンとか、ロボットビジョンの現状といったところから議論していただけますか。

♣Dr. SPIDER ロボットやコンピュータの視覚というと一般の人は何を想像しますかね。鉄腕アトムや「2001年宇宙の旅」のHAL¹⁾のあの赤いランプのような目ですかね。

♡ このシリーズの読者は、もうちょっとレベルが高いと思います(笑)。

♣ コンピュータビジョンとか、単にビジョンとかいうのは、物を見て認識・理解する機能の実現をめざしている研究の総称で、理論やアルゴリズムだけの研究も含まれます。ロボットビジョンというと、もうちょっと即物的で、実際に産業用ロボットや移動ロボットに組み込んで使う視覚システムというニュアンスがありますね。

金出武雄 まあ、そういう感じですね。私は、画像理解という言葉も本質的な違いはないと思っています。

♣ ビジョン研究の話は後にして、産業用ロボットのレベルから始めましょう。10年位前にロボットのブームがありました。日本ロボット学会が設立されたり、極限作業ロボットのプロジェクトが始まったのもその頃です。これが本家であるアメリカにも逆流して、ロボティックス(ロボット工学)研究が再燃したという印象があります。ロ

ボットの視覚については、その後どの程度まで実用化されたのか、まだ何が難しいかという点についてはいかがですか。

金出 まず研究レベルと、実際に工場で使われている産業用ロボットの間には、大きな隔りがありますね。工場の自動化に役立っているのは、ほとんどハードオートメーションのレベルでしょう。

♡ ハードオートメーションってどういう意味ですか。

金出 ラインが固定という意味です。決められた順序に従って、機械そのものの能力を発揮しているだけです。ロボットというのは、本来、ソフトオートメーションであるべきです。プログラムそのものをロードすることにより新しい機能が生まれるというのが本当なんです。

◆ 指示通りの動作を繰り返すプレイバック・ロボット²⁾は、まだ汎用性がありますね。実際、日本では、溶接ロボットとか塗装ロボットとか、限られた単純作業には使われています。ブームの当時、プログラムを変更するだけで、多品種少量の生産ラインにどんどん応用できるという期待がありました。特に、センサー付きの知能ロボットの時代がやってきて、人間が目でやっていることをビジョンシステムが代行すると思っていたのでしょ

金出 そういうことですね。しかし、この10年を振り返れば、失望論をいう人のほとんどは、ロボットがもっとマニファクチャリングに使えていたと思っていた。こういう人は、ロボットを買ってきて、ボンと置いて、パッとプログラムすれば新製品ができると思っていたようです。ところが、現実にはそれほど甘くなかった。

♡ 産業用ロボットにはビジョンを使ったものは、あまりないんですか。

金出 そう多くないけれど、うまく行っているところでは非常にうまくやっています。例えば、ICのチップ上の電極とパッケージのリードとをつなぐワイヤボンディング作業は、比較的簡単だけれど、精密なビジョンなしではできませんね。その他、プリント基板の検査や部品形状検査、溶接ロボットにおける溶接線の検出や溶接状態のチェック等、用途はいろいろありますが、出来合い

の産業用ビジョンシステムを買ってきて、そのまま使えるというケースはまずない。自分で問題を切り分けて、ターゲットを絞った人が成功しています。

研究成果はどこへ行ってしまった

◆ しかし、アメリカのマシンビジョン産業はほとんど衰退してしまいましたね。かつて、デトロイトの自動車産業向けに、検査や選別用のビジョンシステムを作っていた会社がいっぱいありました。今だめなのは、それだけ役に立つ需要がないということでしょうか。

金出 ビジョンが工場の生産ラインで思ったほど使われないのには、3つ理由があるんです。まず値段、それからスピード、そして本当の理由は、プログラムが簡単にできないということです。

♡ 誰ができないんですか。

金出 現場の人です。物体が与えられても、それを認識するプログラムは、ビジョン屋がせっせと書くしかない。1台装置を売るごとに、専門家が1人つかなきゃならない。これじゃ手間がかかりすぎて商売にならない。そして、マシンビジョンの連中は、皆コンサルティング屋になってしまったんです。

◆ 今おっしゃられた3つの点に加えて、システムの信頼性の問題もあるんじゃないですか。使われている技術のレベルも低いですね。

金出 今でも1970年代の研究成果を使っていますね。これじゃ大したことはできません。

♡ その後の研究成果はどこへ行ったんですか。

金出 それは面白いなあ。どこへ行っただろうね。

◆ コンピュータビジョン研究とマシンビジョン産業のギャップが大きいんですよ。これだけ一

「コンピュータビジョンは狭い世界に閉じこもっているように見えます」





金出 武雄 (かなで たけお)

1945年10月兵庫県生まれ。丹波の山奥で幼児期を過ごした後、神戸で高校まで生活。1968年京都大学工学部電子工学科卒業。1973年同大学院博士課程終了。京都大学工学博士。情報工学科坂井利之教授の研究室の助手、助教授を経て、1980年米国カーネギーメロン大学 (CMU) コンピュータサイエンス学科およびロボティクス研究所の Senior Research Scientist。現在、同教授および研究所長。

京都大学時代は、パターン認識・学習、画像処理装置、コンピュータネットワーク等の研究に従事。「ハードもソフトも、基礎も応用もやらなきゃダメ」という坂井教授の教育を受け、本人の生来の移り気とも相俟って以後の多様なテーマに手をつけることになる。学位論文のタイトルは、“Picture Processing System by Computer Complex and Recognition of Human Faces”。CMU では、DARPA の画像理解プロジェクト、自律走行車プロジェクト、NASA の火星探査ロボット・プロジェクト等の研究代表者を担当。コンピュータビジョン分野のエレガントな理論の提唱から、ロボット分野でのダイレクト・ドライブ・アームの開発等、理論・デバイス・システム作り・プロジェクト管理のあらゆる面に興味をもち、これらを精力的にこなす。米国のAI分野でもっとも成功した日本人と言われている。

情報処理学会25周年記念論文賞, Marr Prize Award, AT&T 財団特別奨励賞等受賞。IEEE 及び AAAI フェロー。主要な編著書は、“Three-Dimensional Machine Vision” (Kluwer Academic Publ.) など。

見かけによらず、スポーツは万能。大学時代は研究室のエースで、絶対に打てない(ストライクが入らない)豪球投手として鳴らす。現在は、渡米後に覚えたゴルフ・テニスに熱中している。来日中に購入したカラオケ装置をかついでピッツバーグまで持ち帰って以来、CMU への日本人留学生・研究生を集めて行われる金出家でのホーム・カラオケ・パーティは有名。最近、市内のホテルやバーに入るまで、ピッツバーグ唯一のレーザーカラオケ設備と言われた。客の居ない日の深夜も1人で熱唱される傍らで、黙って聞いておられる同郷の夫人(可憐な美人で賢夫人と評判!)と一男一女の4人家族。

流の先生方が大勢で研究されているながら、その成果がベンチャー企業でほとんど使われてませんね。他のコンピュータサイエンス、ハイテク分野は、もっとアクティブに研究成果を事業化して、お金に換えています。コンピュータビジョンくらいですね、純粋な研究に没頭しているのは。

金出 そう言われるとそうだなあ(笑)。成果を生かす手頃な対象がなかったんじゃないのかな。

◆ 具体的な問題に対して、何ができて、何ができないかと言うことが難しいんじゃないですか。

金出 そうね、何ができるようになったかという、形の3次元的な意味において、対象物の形状が制限されている物体を認識することはかなり進んだ。ところが、名前を対象物を規定することができない。例えば、単に椅子といっても、形も違ひ、色も材質も大きさも全部違う。だから、机や椅子の識別や検査ができるとは言えない。だけど、形が先に決められれば、これは何とかなる。

◆ 現実の課題を、そのレベルへ1回翻訳してやるプロセスが要りますね。ところが、どうやって翻訳してやるかの方法論がないみたいです。

金出 ビジョンの難しさは、センシングとそのデータのインタプリテーションだといえます。

♡ どう難しいんですか。

金出 重要な情報のセンシングができたとして、最終的な判断の部分は、AIの手法を使えば、かなりのことができるようになった。単純に言えば、医療用のエキスパート・システムがあれば、医者並みの診断はできます。ところが、患者の顔色を見るとか、問診するといったセンシングの部分ができない。だから、薬を投与するとか、注射をするといった最終アクションのロボットも必要ないんです。

火事場のレスキュー・ロボットもそうです。耐熱式マニピュレータはできたけど、火事の現場に倒れている人を見つけることができないじゃ役に立たないわけです。

◆ なるほどね。単純なセンシングだけなら、テレビカメラやレンジファインダがあればいいんですね。

金出 画像だけはいっぱい撮れます。しかし、そのデータをどう解釈するかが問題です。これが一般的に論じられていない。

◆ そうですね。文字や図面なら、対象と目的がクリアだからまだやりやすいですね。誰が考えても、数字よりアルファベット、活字より手書きの方が難しい。音声だってそうですね。孤立単語より連続音声認識の方が難しい。ところが、画像や物体を認識しようとする、とたんに問題のレベル自体が不明確になりますね。何の次に何を片づけるべきかという難易度のレベル分けが、自然にはできてない。

これまでいくつもプロジェクトを進めてこれて、こういう話は話題になりませんでしたか。

金出 アメリカでは、何ができて何が難しいかなどという哲学的な議論はめったにしませんね。日本人の方が好きです。

◆ 確かに日本人は、今どこまで来たかとか、何が最新動向かとかを気にしますね。それだけ自分に自信がないからかなあ（笑）。

サイエンスは必ず役に立つ

◆ では、世俗を離れたコンピュータビジョン研究の方ですが、これは今、かなりサイエンス寄りに振れているんじゃないですか。重箱の隅をつついたような研究とか、自分で設定した問題を解いて狭い世界に閉じこもっているように見えます。

金出 この何年間か、そういう傾向はありますね。

◆ サイエンスとしてのビジョン研究の功罪をちょっと考えてみたいと思います。罪というのは言い過ぎかも知れませんが……。

金出 いやいや、サイエンスというのは、本来すべて功であるはずで、基本的には役立つんです。エンジニアリングと結びつかないのが罪ならば、それは、本来やるべきことをやってないからです。

◆ 今のビジョン研究のほとんどは、似非サイエンスであって問題の本質を突いていないということですか。

金出 そうです。本来のビジョンの目的は、知能機械の情報入力チャンネルなんですよ。

♡ どういう意味ですか。

金出 視覚というのは、文字通りロボットの目に過ぎなくて、トータルシステムの一部として機能すべきだということです。この原点に立ち戻るなら、研究者としては、応用、アルゴリズム、そしてそれを装置とシステムに組み上げることまでやるべきで、この3つ全部をやらなければビジョン研究じゃないんです。うちなんかは全部やっている。システムも作っているし、専用チップまで作っています。

◆ アメリカのビジョン研究社会は、その3つの真ん中ばかりをやり過ぎですね。

金出 確かに、アルゴリズムばかりやっていて、最近ではシステム作りをやる人が少ないね。論文ばかり書いている連中は、何を対象として研究しているのか知らないんですよ。例えば、動きのパラメータをやっている研究者は、どれくらいの精度が出て、どれくらいの処理時間でできて、どういう装置があれば実現できるのかを考えようとしな。どういう目的で、運動を解析して、何を抽出したいのかを忘れてしまっているんです。

◆ CMU の金出グループは、例外的にバランスがいいんでしょう。

金出 一応そう思っていますが、SRI^① の連中なんかも、昔からそういうバランスは良くて、よくシステム作りをやってますよ。ただ、システム作りをする時は、既知の方法の単なる応用・変形になる危険性もある。だから、研究として進展させるには、なるほどシステムというのはこういう具合に作るのかという所を見せるべきです。

◆ 応用システムをやるのはいいけれど、シス



「顔の金出と言われたもんです」

♡ 「それはそれは……」

テム構築の方法論が要りますね。似た問題を同じように解く方法を示す必要があると思います。

金出 ただ作ったというレポートだけじゃ誰もついて来ない。方法論のデモンストレーションがないと説得力がない。われわれの自律走行車は、その辺をかなり意識してやってるつもりです。

♡ 画像分野は、途中の処理結果でも最終結果でも人に訴えるものが大きいですね。画像認識やコンピュータビジョンでも、もっとデモでアピールできる研究が出てきて欲しいと思います。

環境問題でロボット再待望論

金出 画像認識全体で考えるなら、実用的な成果も少しずつ出てきてるでしょう。昔やった問題がまた形を変えて取り上げられている。たとえば、ペン入力タブレット。あれなんか、昔でいうオンライン手書き文字認識ですよ。

♣ 今流行のペン・コンピュータですか。確かにそうですね。テレビ電話が実用に近づき、顔画像の認識もあちこちで取り上げられるようになりました。知ってますか、金出先生は昔、顔の認識の大家だったんですよ。顔で博士号を取られたんです。

♡ へえー、そうだったんですか。いつ頃のことですか。

金出 20年くらい前かな。頭じゃなく顔で売り出したんです。「顔の金出」と言われたもんです(笑)。

♣ 私は、細線化のアルゴリズムをやっていたから、「細めの田村」と言われた。今は太目ですけどね(笑)。

♡ 「顔の金出」と「細めの田村」ですか、それはそれは(笑)。

♣ 完全な顔の認識は難しいけれど、圧縮して伝送するための特徴抽出とか、人の同定のための顔の識別などは手頃なテーマですね。誰かがいい問題を持ち出すとそれが活性化の起爆剤になります。

金出 移動ロボットもそうですよ。本来のロボットとしての用途は、アメリカでは環境問題での利用がクローズアップされてます。

♡ 環境ですか。

金出 ええ。特に核廃棄物ですね。昔はよく考えずに実験後の放射性物質を捨てていて、アイダホあたりの岩盤に穴を掘ってホンホン放り込んでいた。こういう所がアメリカに500か所くらいあります。ひどいところでは、大雨が降って洪水が起こり、水がその穴に流れ込んで、廃棄物のドラム缶がブカブカ浮いてきた、なんて話もあるんですよ。

♣♡ へえーっ。

金出 あと、地下タンク。昔、核物質を作ったという工場は皆閉鎖されている。でも、それをそろそろ壊さなきゃいけないけれど、とても人間がやれる仕事じゃない。日本じゃこれを極限作業というんですか。そんな所でロボットの利用に大きな期待が寄せられているんです。

♣ いわゆる3K作業にぴったりなわけですね。

金出 スリーマイルズ島の原発事故は知っているでしょう。

♡ ええ。

金出 これはその事後調査に使った移動ロボットです(写真1)。いろいろな部屋の破損状況を調べて回ったんです。これにはまだケーブルがついていて、人間が遠隔操作していた。これを自律的に行動できるロボットにしようというんです。

♡ 自律的ってどういうことですか。

♣ 外界からの刺激に対して、自分で判断して適応的に行動するということです。視覚センサや触覚センサで、周囲の状況や障害物を判断するんですよ。

金出 そのとおりです。その自律システムの例として、1984年から運転手がいなくても自分で走

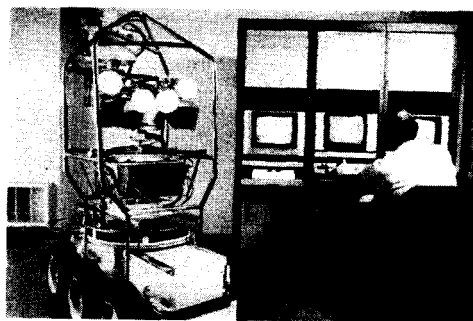


写真1 スリーマイルズ島の点検ロボット

○ plus E・1992年9月

れる車、自律走行車の研究をしているんです。

自分のプログラムと運命を共にする

♡ では、今日のメインテーマの1つであるその自律走行車のお話を詳しくしていただけますか。

金出 ALV といって、Autonomous Land Vehicle の略で、野外を自律的に走る車の開発プロジェクトがDARPA⁹⁾ で始まった。それ以来、CMU では本格的な野外自律走行車の研究を8年やっています。

♣ 軍需産業のマーティン・マリエッタ社が八輪駆動の大きなクルマを作っていましたね。もうあのプロジェクトは終わったんじゃないんですか。

金出 ALV プロジェクト全体としては、4年で予算打切りになったけれど、CMU はよくやっているというので、うちだけはまだ予算が続いています。

♡ 予算規模はどれくらいなんですか。

金出 毎年、2 ミリオン・ドルだから、これまでで22億円くらい使っていることになりますね。

初期のタイプのものは、コンピュータは搭載せず、普通のTVカメラやアンテナを積んでいたんです。ちゃんと電波管理局の許可をとって、UHFの電波で画像を研究室のコンピュータに送っていた。この画像をデジタル化し、解析し、コンピュータ側からクルマがどう動くべきかの指示を送り返す方式を取っていた。小さな放送局が移動しているわけだから、これだと家に居てもテレビのチャンネルを合わせれば、今、何処を動いているかを確かめることができたんです。

♣ 家でビールを飲みながら、研究のチェックができるんですね(笑)。

金出 ところが、これだとセンサや動力部は積んでいるけれど、意志決定をする部分がない。それで、コンピュータを積んだり、いろいろなセンサを載せようとする、と、重心がだんだん高くなってきて不安定になってきた(写真2)。そこで、しっかりしたものをということで作ったのがこの実験走行車(写真3)で、Navlab と呼んでるんです。

♣ バンを改造して、コンピュータ制御にした

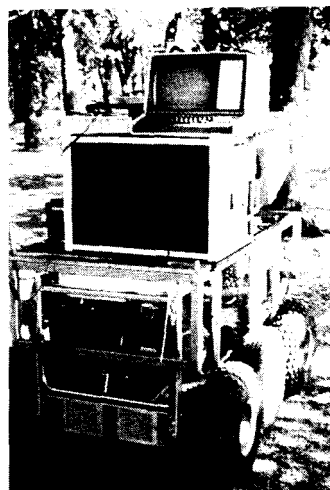


写真2 初期の自律走行車



写真3 Navlab の外観

んですね。

♡ かなり大きなクルマですね。これなら沢山積めますね。

金出 自己充足型といっています。必要なものはすべて搭載してある。重要なコンポーネントは3つあって、① on-board sensor, ② on-board computer, ③ on-board researcher なんです。さらに、最近になって、Navlab II というもっと高性能な新しいモデルを作りました。これはその構成図です(図1)。

♡ 人間も乗るんですか。

金出 ええ。以前の車だと、無線で連絡し、「さあ行くぞお」とか「そこで停めろ」とかやってたのが、これ1台で全部同時にできる。この車に乗っていると、研究者にとっても、どういう状況で自分のプログラムがどう働くのかその場でチェックできる(写真4)。

♣ 欠点もすぐ分かりますね。

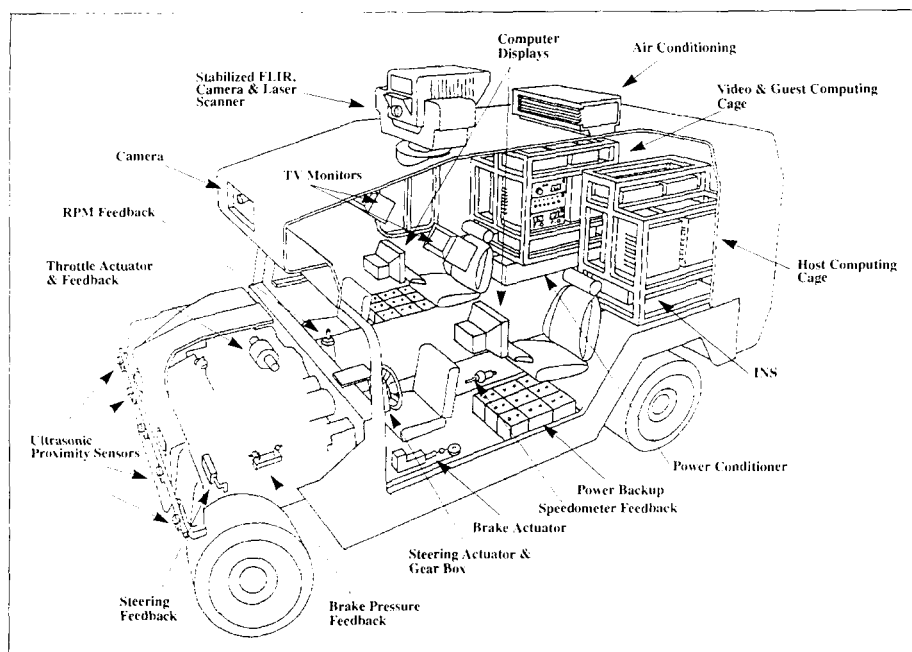


図1 Navlab II の構成



写真4 Navlab の内部

金出 プログラムもその場で修正できます。何しろ、自分の書いたプログラムと運命を共にするんだから、下手なプログラムは書けない(笑)。

♠ うっかりすると衝突したり、崖から落ちたりしますね。

金出 これをやって以来、プログラムの信頼性がぐんと上がった(爆笑)。じゃ、どうやって走るのがビデオを見ましょうか。

[Navlab のビデオ・デモ]

自律走行車 Navlab の動く様子は、以前に NHK 教育テレビの科学技術番組で見たことがあります。大学のキャンパス内の林の中の道をゆっくりと進む様子は、なるほど研究が一步一步着実に進んでいるのだなという印象を受けました。それが今回は公道を走るビデオというので、危くないかしらと心配しながら興味深く眺めました。

まず、装備の説明です。カラー TV カメラとレーザ・レンジファインダを搭載しています。それぞれカラーの画像と距離画像を得るため、こうした複合データを利用するのが「センサ融合」(sensor fusion) というそうです。もちろん、コンピュータも搭載されていて、ハンドルやアクセル、ブレーキ等もコンピュータ制御されています。運転席は無人ですが、コンピュータによる画像解析の結果が表示されているディスプレイがありました(写真5)。

Navlab が走り出しました。一般道を走っています。道路両端のガイドラインやセンターラインを検出して走っているようです(写真6)。次に雨の中をワイパーをかけながら走っているシーン。へえー、こんな時でも走れるんですね。今度は、ハイウェイです。うわーこれは速い。ちょっとビックリするくらいのスピードです。

かなり画面が変わって、住宅地の中の道路をゆっくり

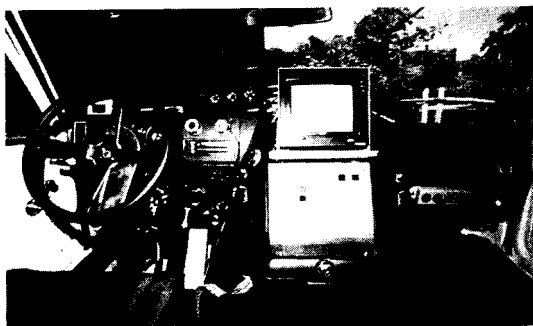


写真5 Navlab と運転席

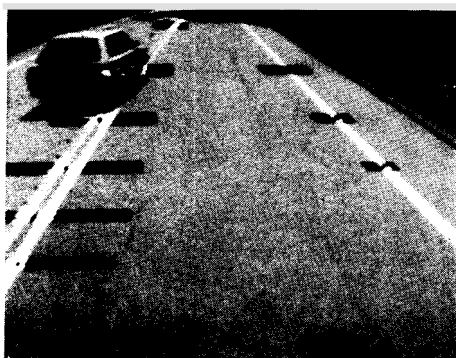


写真6 道路シーンの解析

走っています。道路の傍のガイドラインはないので、これは画像処理としてはかなり状況が異なり、難しそうです。まったく勝手に走るのではなく、あらかじめ地図がコンピュータの中に入っていて、どこをどうドライブするかを自律的に決めているようです。アメリカの住宅地には、各戸ごとに大きな郵便受けが道路沿いにあり、この位置の検出も目的地を探す重要な手掛かりになっています。

Navlab の前に自転車に乗った子供が飛び出しました。あっ、危ない！ でも、チャントこれを判断して停止しました。エライ、エライ。デモビデオとはいえ、本当に停まるというのは大したものですよ。

交差点にさしかかりました。さあ、どちらに行くんでしょう？ ランドマークを検出して、コンピュータが提示したドライブプランに従って正しく曲がりました。いやはや驚きです。もうここまで進んでいるのかと感心しました。♡

我が子を使って安全性を証明

◆ もう随分進んでいるんですね。こんな感じで一般道を走っているとは思いませんでした。

No. 154・O plus E

金出 だんだん大胆になって、いろいろなところを走らせて実験しています。普通の人がこれがあるといけないから、運転席に人を座らせたりもしていますが……。

♡ かなり速く走れるんですね。

金出 時速55～60マイルだから、80～90 km くらい出していますね。

◆ 画像のセンシングは距離データと濃淡データの併用ですか。

金出 そうです。駐車場で障害物を避けたり、人が飛び出してきたのを検出するのは主に距離が頼りです。レーザの距離センサで、1/2 秒で 64 × 256 画素の1画面が入力できます。

♡ 人を感じてちゃんと停まりましたが、あれは本当に子供が突然出て来たんですか。

金出 いや、あれはヤラセです（笑）。あの子はおね、プログラムを書いたスタッフの子供なんです（笑）。それでも、信用してなきゃできない。

◆ まず自分の子を使うんじゃない、まるでウィリアム・テルかジェンナーみたいですね（笑）。それにしても、かなり複雑な動きができるんだから、全体として相当なプログラム量ですね。

金出 これは大変な作業量です。今は、地図の中にいろいろな指示を書きこんである。交差点に近づいたら何を注目しろとか、その場合にはカラーカメラの方を使えとか、手法はニューラルネットを使えとかが書いてある。こういう風にどんどん増やして行っているの、膨大なシステムになります。

♡ 自律的に行動するといっても、どこへでも行けるということじゃないんですね。

◆ まだ、それは無理でしょう。さっき話したブレイバック・ロボットのように、まったく同じ動作を反復するのはレベルが違うけど、だからといって、あらゆる環境で動けるわけじゃないですよ。

金出 最終的には道のない所へも行ける、というのが目標です。いま、森の中を木や大きな石を避けて通る実験をやっています。

♡ どんな処理をされているんですか。

金出 主に距離画像を使ってます。Terrain Map と呼ぶ地面の高さの凹凸や木・岩・溝等の

障害物の形・大きさ・位置を表現したものを自動的に作り出して、その中で、車がどこまで傾くか、車輪が溝にはまらないかなどを考えて、安全な経路を計算するんです。

◆ 他の機関では、せいぜい道路のガイドラインを見つけて進む程度ですね。

♡ 他でもやっているんですか。

金出 日本では、日産と富士通が共同研究していますね。それからドイツでもやっている。基本的には皆レーンマークを見て走るというスタイルです。

◆ こうした研究開発は、Navlab が火つけ役になり、これに触発されて自律走行車をあちこちでやり出したんじゃないですか。

金出 そうだと自負しています。クルマそのものは、専門メーカーの方がしっかり作ってあるでしょうけど。

♡ そろそろ実用になるんでしょうか。

金出 実用にもいろいろレベルがあるし、技術的には可能でも、社会的に受け入れられるかどうか分からない。ただし、工場内の限られた環境などでは可能だと思いますよ。ゴルフ場でなら、自動カートとして使える。

◆ 今のリモコン電動カートの代わりですか。確かに、自分で判断してついてきてくれるのはいいけど、あまり右へ左へOBしているとALVも大変ですよ（笑）。

金出 私の場合は、うまいからそんなことは起こらない（笑）。

（注：このインタビューの翌日、一緒にラウンドしたが、意に反して金出先生は絶不調、久々に100をオーバーするという、ご本人としては惨々なスコアであった。）

火星を歩く6本足ロボット

◆ 原子炉内や海底とかの極限作業用のロボットとしては期待が大きいんじゃないですか。そういう応用は……。

金出 もちろんやっています。まず、大きなプロジェクトとしては、NASAの火星探査計画で、火星の岩石を採集して地球へ持ち帰ろうというものです。1992年から2000年にかけての壮大な計画

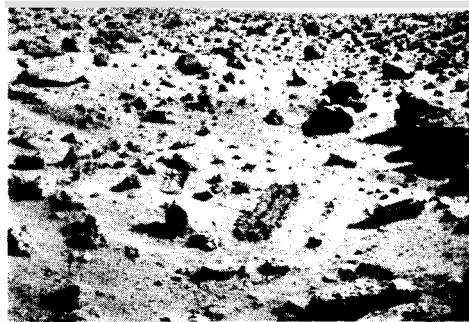


写真7 火星表面のシーン



写真8 Ambler

です。

♡ 井口先生のときにお話がありました。Mars Rover とかいう……。

金出 そうです。サンプルを見つけて取ってくるには、火星表面を移動できるロボットが要る。これは全自動である必要はなく、地球から採取すべき目標地点などを指示してもよい。だけど、動作をいちいちリモコンすることはできない。地球ー火星間の通信は往復1時間くらいかかってしまうからです。

◆ それで、ある程度自律的に動く必要があるわけですね。

金出 そう考えて、すぐプロポーザルを出した。だけど、Navlab じゃ無理なんです。空気はないし、砂と岩の世界を移動しなければならない（写真7）。

♡ この写真は実際の画像ですか。

◆ 前にバイキング⁹⁾ が撮ってきたものです。

金出 そうです。この環境で3mの段差、30°の勾配を歩行できるように設計したのがこのロボットで、Amblerと呼んでいます(写真8)。

♡ 随分大きそうですね。

金出 高さが約4mで、重量は2t以上あります。火星で、1.5~2mくらいの岩をつかむことを想定しているからです。これは、6本足で、18自由度の構造をもっています。

♡ 6本足なんて昆虫みたいです(笑)。でもどうして6本なんですか。

金出 これはよく聞かれるんです。実際は3本ずつの制御系に分かれていて、1番後ろの足を軸の周りにグルッと回して前へ出す。1本が移動している間、残り5本は地面に固定されています。よく知られているように、安定して重量を支えるには最低3本要りますね。火星では不測の事態を考えて、急に足が壊れたりとか地面が抜けて穴があいたりとか、そんな時にも重心が、残りの足で作る多角形の中に入るように検討した結果です。

♣ かなり安全を見込んであるわけですね。視覚センサは何を使っているんですか。

金出 これはね、歩行用とサンプル採集用で2種類使います。どちらも距離センサです。歩行用は、Navlabと同じようにレーザ光の位相差を利用したもので、水平・垂直方向におおの60°の視野、奥行きに2mから40mくらいの距離を12ビットの精度で測れます。サンプル採集用は、液晶フィルタを使った光投影型で距離データと同時に濃淡画像も撮れます。

♣ Amblerの歩行練習はどうするんですか。まさか、これは公道を歩かせるわけには行かないでしょうから(笑)。

金出 ゴルフ場で、タマ拾いでもさせましょうか(笑)。実際は、大学の建物内に砂場のような場所を作り、岩を置いたり、穴を掘ったりして、練習させています。さらに今では、外へ出して近くのウェスチング・ハウス社の敷地で数か月動かしています。

♡ こんな大きなロボットを宇宙に運べるんですか。

金出 この程度の大きさは問題じゃないんです。だけど、われわれは、このAmblerが本当に

そのまま火星に行くとは思っていない。宇宙に送るには、space qualificationという検査にパスした部品だけで作られていないと駄目なんです。今のAmblerはあくまで基礎技術のデモンストレーション用なんです。

♠ でも、アイデアや技術は、別の目的に生かされるでしょう。これまでに、NASAの開発した技術は、かなり実社会で生きていますよ。

金出 われわれの目標は、技術的可能性を示すことです。自律走行という点では、海底ロボットは、Autonomous Underwater VehicleでAUV、Air VehicleはAAV……という風にどんどん発展できます。こうしたシステム作りをやっているのは楽しいけれど、あまりやり過ぎるとサイエンスとしての進歩がないので、程々にしておこうと思っています。

♡ じゃあ、逆に、コンピュータビジョンの理論的な研究の成果は、NavlabとかAmblerなどには活かされているんですか。

金出 うーん、これは答えにくいなあ。オフィシャルには、イエスなだけ……(笑)。実際には、最先端の論理は、なかなか直接には使えませんが。開発型のプロジェクトは答えを急ぐので、根本的な原理に立ち戻っている暇はない。

♠ NavlabやAmblerで使っている手法も、かなり前の研究成果が、今では当たり前のように使われていると言えますね。

金出 やはり使うフェーズが遅れますね。むしろ逆に、開発プロジェクトが、どういう問題を解決すべきかを基礎研究側に語ってくれている面の方が重要とも言えます。

バーチャルリアリティはビジョンの活躍の場

♡ 今日は自律走行のお話ですが、シス

「もっとデモでアピールできる研究が出てきて欲しいと思います」





「うまく行くものなら何を
やってもいいんです」

代表的なものとしてはテレプレゼンス^①やテレイグジスタンスなども話題ですね。現実のシーンを人間が立体的に眺めながらロボットを操作するのは、極限作業向きだと思うんですが。

金出 これは自律ロボットとは別の意味で大事ですよ。

◆ 知能ロボットは本来自律的なものですよね。テレイグジスタンスのように、それを人間が介在して動かすというのは、ある種の逃げじゃないですか。AIの観点からすると邪道じゃないかと。

金出 それは違うなあ。自律システムといったって、何をするかは結局人間が指示する必要がある。テレプレゼンスやテレロボティクスがやさしいと思えるのは、現在のAIでできないデジジョンの部分人間がやるからであって、その点に注目するからです。本当は、人間というわけのわからない要素がシステムに入ってくるテレプレゼンスの方が難しい。

◆ 人間という不安定で不確実な個人差があるものが入ってくると、システム全体としては複雑になるということですか。

金出 例えば、物を掴む場合、対象物が判っている場合は、直接アームをもって行く方が早いです。それを、人間がスイッチを入れ、テレビ画面を眺めて判断して、別の世界にある機械系とのコミュニケーションをとる。これは、簡単な話じゃなくて、動作が複雑になればなるほど、コミュニケーションのギャップが出てくる。限られた対象に対しては、自律システムの方がやさしいことも多いんです。

◆ なるほど、そういう考え方もありますか。自律システムの方がやさしいとは思わなかったですね。でも、ロボティクス研究所の所長さんが言うんだから、そうなんでしょう（笑）。

♡ テレプレゼンスは、バーチャルリアリティ（以下、VRと略）の一部とも考えられていますが、ビジョン分野から見てVRについてはどうお考えですか。

金出 VRは大いにビジョンの活躍できる場だと思いますよ。いまのVRは、絵を出す方にすごく重点が置かれていますね。

♡ CGでステレオ画像を作るんですね。

金出 ところが、なぜVRを作るかとよく考えてみたら、そのもとになっているreal realityがあるからですよ。もちろん、まったく架空の世界を作ってその中に入ることもあります。大抵は体験したい現実の世界がある。

VRで体験するのは、金がかかりすぎるとか、危険すぎるといった理由からなんだから、リアルワールドをバーチャルワールドに反映させるのが望ましいんです。それを私は、“putting reality into virtual reality”と呼んでいるんです。

◆ そこにビジョンの活躍する場があるというわけですね。だけど、それが難しいんじゃないんですか。このシリーズでも、原島先生の顔の知的符号化も、岸野さんの仮想空間通信会議も、この入力側がネックだったんです。画像認識・理解側のレベルが思ったほど高くないとの結論でした。

金出 もっとこういう視点からのビジョン研究をやれば、かなりのレベルまで行けると思います。こうして作った仮想世界で得た成果を、また現実世界に戻して役立てるというサイクルも大切ですよ。

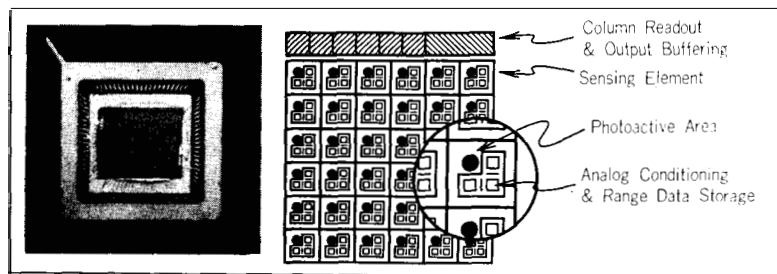
昔の名前じゃ流行らない

♡ まだ伺ってないのは、専用チップのお話ですね。その辺を少しお願いします。

金出 これは、高速で距離データを得るためのアナログ素子なんです。VLSIを使った3次元センサで、1秒間に800枚から1000枚の距離画像が撮れる。

◆ それは速いですねえ。今のレンジファインダとは桁違いでしょう。

金出 今はせいぜい1秒に2枚から10枚くらいだから、2桁くらい違います。素子とプロトタイプの装置の方はもうできています。ただ、今のと



ころ、ワークステーションとのインタフェースがネックで、これがせいぜい1秒に20~30枚くらいしか送れない。それももうすぐ解決するから、そうしたら3次元のスローモーション・ビデオを作ることができます。

♠ それは面白そうだな。でも、どうやってそんなに高速に距離センシングできるんですか。

金出 専用のアルゴリズムと、それを実現する素子を結びつけたところがミソなんです。

♠ あんまり、素子、素子って言わないで下さい。活字にすると素子(そし)は、素子(もとこ)になっちゃうんです。

金出 ???

♡ 私の名前です。あとで読んだら、何か呼び捨てにされるようで……。

金出 えっ。あ、そうか、素子さんか。これは失礼（笑）。じゃあチップで行きましょう。

VLSI のセンサへの応用という点、従来は、すでにあるアルゴリズムをチップ化し、高速化するというのが発想だったんですね。例えば、画像からエッジを抽出するのに Sobel オペレータ¹⁾という簡単な微分演算があり、これを LSI 化してリアルタイムにビデオ信号からエッジ抽出をやる。これはね、別にチップを作らなくても、コンピュータでも実行できるし、どんな画像になるか結果が判っていた。これは VLSI だからできたというんじゃない、VLSI でもできたにすぎない。

♠ ソフトでやってもハードでやっても、違いはスピードだけというわけですね。最近では、DS P³⁾でやってもリアルタイム処理できますね。

金出 われわれのチップは、意味が違います。このチップの光学的、光幾何学的な原理は、これまでのものと同じなんです。ところが、アルゴリズムが違う、オペレーションの方式が違っていて

No. 154. O plus E

[illegible]

図3 センシング要素の回路図

す。詳細は省きますが、このチップによって初めて実現できるオペレーションなんです。

♠ プログラムで実行するアルゴリズムとは、
違うということですか。

金出 違うんです。このチップには、光をセン
スする能力と、そのセンスした光の情報を処理す
るプロセッサが、各ピクセルごとに備わっている
んです。

♡ 普通のカメラだと、光を検出するだけですね。

金出 そうです。光を受けるフォトダイオードの周りに、アナログ回路やデジタル回路がつながっている(図2, 図3)。これでこのチップ独自のアルゴリズムが完全並列型で実行でき、1秒800~1000枚の入力が実現できた。このアルゴリズムをそのまま普通のコンピュータで実行したら、従来のアルゴリズムよりもかえって遅いんです。

♠ VLSI 化して初めて意味があるアルゴリズムということですね。

金出 この考え方が随分当たって、DARPA や NSF からもお金をもらってます。私はこれを computational sensor と呼んでいます。

♠ 処理機能のあるセンサというなら、昔からあるスマートセンサとは何か違うんですか。

金出 同じと言えば同じですよ。スマートセンサは、「賢いセンサ」という意味ですね。

♡ そのスマートセンサというのは、使われているんですか。

金出 アイディアは、相当古くからあるけど、本当の意味で役に立つものはあまり実現してませんね。だから、そんな古い名前じゃ受けないから、本当の実現はこれが初めて、ということで新しい名前をつけた。こういう風に、システムのレベルからチップのレベルまで、新しい考えでインテグレーションするというのが、私の昔からのやり方なんです。

♠ 昔から名前をつけるのも好きですねえ。このインタビューの中でも、もう随分金出用語が登場しましたよ（笑）。

金出 うまい名前づけですが、アイディアの売り込みに必要なものということで、私は多少揶揄的に buzz word⁹⁾ engineering と呼んでいます。しかし同時に、本来の意味を一語でとらえるという努力を要することから、アイディアを研ぐことに役に立つと言えますね。

これからのビジョンは Kiss から

♠ 今日は、ビジョン産業論から、自律システムやチップの開発まで、実にいろいろなお話をしたのですが、金出グループとしての今後のビジョン研究の方向づけはいかがですか。

金出 私自身、気の移る方だから今でもいろいろなテーマを手がけています。画像理解のアルゴリズムをやるかと思えばシステムもやる。移動ロボットの他にマニピュレータから、さっきの3次元センサまで雑多です。今はまだ統一的理論のない時代ですよ。問題は、うまく行くかどうかです。うまく行くものなら、何をやってもよい。

♠ 研究者人口が多い割に、役立つものをやるという気運が少ないのは残念に思います。

金出 問題を難しくしすぎないことが秘訣でし



よう。私の研究の共通点として、Kiss アプローチというのをかかげているのです。

♡ えっ。

金出 Kiss というのは、アメリカの軍隊用語で、“Keep it simple, stupid!” の略なんです。軍曹あたりが部下に使う言葉で、文字通り訳すなら「バカ者。簡単にやれ」といいますか……。

♠ 軍曹と部下が Kiss なんて、エイズの関係かと思いましたよ（笑）。

金出 物事をやる根本原理として、アメリカでよく使われる言葉です。私は、単純な発想のアイディアが大切だということで使っています。複雑なコンピュータビジョンの問題も、1つずつこつこつと解決して行くことで、だんだん役に立って行くでしょう。

♠ それを期待しています。

♡ どうも有難うございました。

エピローグ

金出先生は、このインタビュー・シリーズを大変気に入って下さり、CMU まで取材に来るようにと勧めて下さいました。さすがにアメリカまで行く余裕はなく、一時帰国中の貴重なお時間を頂いてのインタビューになりました。私は、事前勉強不足もあり、ただですら難しいビジョン分野が、なおさら難しく感じてしまいました。

この CIF シリーズでは、CG 表示の圧倒的な迫力と効果に魅せられながらも、コンピュータビジョン分野への期待の大きさも肌で感じてきました。反面、ビジョン研究はあまり役立たないと聞かされてきたせいか、Navlab や Ambler のビデオを見せていただいて、もうここまですでているのかと驚くと同時に、大変心強く思った次第です。「サイエンスは本来役に立つものである」という金出先生の持論が、見事に実践・証明されている様子は、デモンストレーションの大切さを物語っているようです。人間にとって危険な極限作業を代行するようなロボットの需要は、今後ますます大きくなって行くでしょう。

う。次は技術的可能性を現実のものとする段階なんだと実感しました。

金出先生は、ご講演もインタビューも大変分かりやすく、かつエネルギーに話して下さいました。Dr. SPIDER も学生時代からのお付き合いとあって、ガッチリ議論もかみ合い、傍らで聞いていてもっとも有意義なインタビューになりました。明解な主張のあるご発言は大変気持ちがいいですね。特に、「うまく行くものなら何をやってもよい」というポリシーは、まさに Kiss アプローチです。金出先生率いる大掛かりでシンプルなプロジェクトに触発されて、ビジョン分野全体が世の中の要望に応えられる真のサイエンスになって行くことを期待しています。(♡北村素子)

用語解説

- 1) HAL: アーサー・C・クラーク原作, スタンリー・キューブリック監督の SF 映画「2001年宇宙の旅」に登場する大型コンピュータ。宇宙船の乗組員の唇の動きを読み、人間に反乱を企てた。人間の言葉を理解し、学習する未来の人工知能システムの代表として語られることが多い。名前の由来は、IBM からアルファベットを1文字ずつずらしたものとされている。
- 2) プレイバック・ロボット: 産業用ロボットの作動方式の1種。あらかじめ人間が、コンピュータ等を動作させて作業内容を教示し、位置・姿勢情報を時間情報と共に記憶させ、運転時に教示内容を再現する。

- 3) SRI: SRI インターナショナル社 (旧 Stanford Research Institute) Dr. Fishler, Dr. Bolles らが率いるビジョン研究グループは、移動ロボット、画像理解研究でも長い経験と実績をもつ。
- 4) DARPA: Defense Advanced Research Project Agency (米国国防総省高等計画局)。ARPA ネットワークの構築を始め、コンピュータサイエンスの基幹研究にも多額の研究予算を投じてきた。音声理解、画像理解プロジェクトのスポンサーでもあった。
- 5) バイキング: 1976年に実施された NASA の火星探査計画。火星表面のカラー画像を地球に送信してくることに、多数の岩石を採集し地球に持ち帰った。
- 6) テレプレゼンス: telepresence. 時間や空間を隔てて現実存在する世界を人間がディスプレイ等を通して知覚し、ロボット等を遠隔操作する技術。テレイクジスタンス (tele-existence), 遠隔現実感 (tele-reality) ともいう。
- 7) Sobel オペレータ: 3×3 近傍マスクを用いた微分オペレータ。画像処理でしばしば用いられるエッジ検出のためのもっともシンプルな手法。
- 8) DSP: digital signal processor. 積和演算を中心としたデジタル信号処理用に開発させた LSI チップ。最近のワークステーションでは、主 CPU 以外のコプロセッサの1種として DSP を標準装備している機種も増えている。
- 9) buzz word: バズワード。ひびきが良くて、話題となるはやり言葉。fuzzy, neural, hyper, virtual... などその類い。

ホスト&アシスタント

Dr. SPIDER (田村秀行) / 北村素子

[キヤノン(株)情報システム研究所]



■立体視パターンの手作り

裸眼立体視により、ステレオ写真やいろいろなパターンを立体的に見ることができるという話題が、あちこちで取り上げられるようになりました。これは知識の普及という点からは歓迎すべき事柄ですが、以前からこの分野にかかわってきた者としては、多少注意して対処する必要があると感じています。というのは、多くの新しい人を巻き込んだ結果、専門知識のレベルが大幅に低下してしまい、見栄えの悪い画像が多数作り出される。この結果として、立体画像全部が程度の低い物と考えられ、急速に飽きられてしまう。ステレオ写真に限らず、こんな現象をたびたび経験してきているからです。

画像にかかわる話題には、いつもかなりの混乱があります。その原因のひとつとして、論じている人が自分自身の手で作品を作ってみないため、誤解に基づくまったく間違った主張をしてしまうことがあります。多くのステレオ写真の識者を悩ませ続けている『ステレオ写真の撮影基線長』に

関する議論などは、この典型的なものです。

そこで今回は、手書き図形による立体視パターンをまず取り上げ、最終的にはワープロによるランダム立体視パターンを実現することといたします。

●最初の立体画は手描き図形から

ホイートストンが立体鏡 (Stereoscope) を発表した1838年は、ダゲールが銀板写真 (Daguerreotype) を発表する前年に当たります。このときの画像は手描きによるもので、大小2重に描かれた正方形図形などが記録に残っていますが¹⁾、このような単純な幾何学図形は、方眼紙の上にテンプレートを使って簡単に描くことができます。これはいかにも幼稚な方法のように見えますが、なかなかどうして、本気で取り組むと結構楽しむことができます。図1は不可能図形を立体画にしたもので²⁾、方眼紙上に描いたものですが、多少の無理はあるものの、なかなか魅力的な画像が目の前に開けてくるとは思いませんか？

複雑な立体画を描く道具は、多くの人により考えられ、改良が加えられていますが、なかでもウラディミール・タマリ氏の3次元描画装置³⁾は有名です。どのような道具でも、要するに自分の描きたい題材に適した構造を採用すれば良いので、いろいろお試し下さい。

●貼り込みによる立体視パターン作製

良く見える立体画を描くには、左右の図形の対応する部分が一致している必要があります。手描きの場合には、線の微妙な曲がり具合や線の太さをそろえるのに苦労します。この問題を解決する

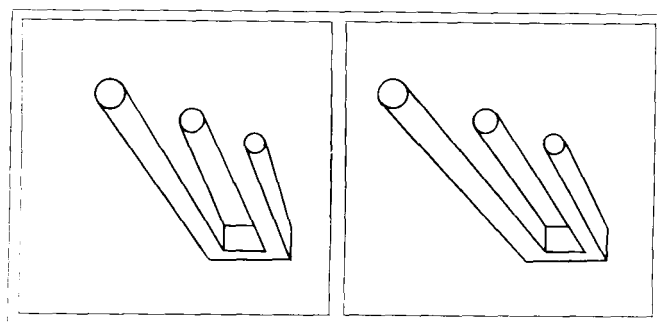


図1 立体視パターン
「3つ又のサス」

,

