

題名	素人のように考え、玄人として実行する：楽しく役に立つビジョンとロボットの研究
Title	Think like an Amateur, Do as an Expert: Fun Research in Computer Vision and Robotics
著者名	金出 武雄
Author(s)	Takeo Kanade
言語 Language	日本語・英語 Japanese, English
書名	稲盛財団：京都賞と助成金
Book title	Inamori Foundation: Kyoto Prize & Inamori Grants
受賞回	32
受賞年度	2016
出版者	公益財団法人 稲盛財団
Publisher	Inamori Foundation
発行日 Issue Date	8/31/2017
開始ページ Start page	108
終了ページ End page	153
ISBN	978-4-900663-32-9

素人のように考え、玄人として実行する： 楽しく役に立つビジョンとロボットの研究 金出武雄

京都賞先端技術部門受賞の榮に浴し大変光榮に存じます。

私の研究は計算機視覚と知能ロボットと呼ばれる分野です。カメラからの画像をコンピュータで処理し・理解する、つまり人間の目に当たる機能を実現し、そしてそれを用いた賢いロボットを作り出すというものです。そんな研究はどんなものか、先ず最近やっている一つの例をお話したいと思います。

話を始める前に、聴衆の皆様にお願ひがあります。私の話には、時々ジョークが混ざることがあります。「これはジョークだな」と思われましたら、ドッと笑っていただけでも盛り上がり楽しくなるのでご協力のほどよろしくお願ひいたします。

雨が消えるスマートヘッドライト

雨の夜に車を運転していると、たくさんの雨粒がヘッドライトに照らされて白い筋のように見えて運転がしにくいことがあります。雨粒は水ですから透明のはずですが、実は丸い水滴は光をよく反射する性質があります。水滴は、ヘッドライトで照らされるとその光を反射して、視野の中で白く光り運転の邪魔になるのです。では、雨の夜に水滴が見えないように、まるで雨が降っていないように感じながら運転することはできないでしょうか。

雨が白く見える理由がわかったのですから、答えは簡単です。雨が降ってくれば、ヘッドライトの光のうち雨粒に当たる光線は出さず、雨粒の間をすりぬけていく光線だけを出すようにすればいいのです。雨粒から反射される光はなくなり、一方すりぬけた光は背景に届いて反射してくるので、運転者の目には背景だけが見え、雨が降っているようには見えません。

そんなマジックのようなことができるのかと思われるかも知れませんが、意外と簡単です。先ず、雨粒の落ちる速度はそれほど速いわけではありません。雨粒はせいぜい直径4 mmから大きくても10 mmで、その速度は最大秒速10 mほどです。反応速度が遅い人間の目には筋のように見えますが、その程度の速さの雨粒は露光時間を短く、例えば1 msec、つまり1000分の1秒に設定したカメラにはこんな風に粒々のように映ります(Fig. 1)。それらのカメラ画像上の位置を知ることはそれほど難しいことではありません。

そんな高速のカメラとプロジェクタを用意し、それらを一つの光学レンズの後ろの同じ位置に置きます(Fig. 2)。これを専門的には共焦点、つまり光の焦点を共有し

Think like an Amateur, Do as an Expert: Fun Research in Computer Vision and Robotics Takeo Kanade

It is indeed a great honor to be chosen as the 2016 Kyoto Prize laureate in the Advanced Technology Category.

As a researcher, I have been involved in the fields known as computer vision and intelligent robots. To put it simply, I work to realize a function akin to what the human eye achieves, by using computers to process and understand camera images and then to develop intelligent robots equipped with that function. I would like to start with a recent project of mine to better illustrate what I do.

Before I begin, I have a favor to ask you in the audience. I occasionally make jokes. If you notice that I'm joking, please feel free to laugh out loud so that we can get this event going and have fun. Thank you.

Smart headlights that make rain invisible

When driving on a rainy night, you might find it difficult to see ahead of you because the falling rain, lit by the headlights, appears to be white streaks. Raindrops are made of water and should be transparent, but these round drops of water are highly reflective. When lit up by the headlights, the drops of water reflect the light and glow white in the driver's vision, thereby making it difficult to see the road. One might wonder if there is a way to make these water drops invisible on a rainy night so that one can drive just as smoothly as if it were not raining.

Now that we know why raindrops appear white, the solution is straightforward. When it rains, we control a headlight so that it only emits light that goes between the raindrops, while turning off light that falls onto them. This way, the light won't reflect off of the raindrops and will pass right between them to illuminate and reflect off of the background. Then, the only thing that drivers will see is the scenery ahead of them, and it will appear as if there is no rain.

You might think, "That sounds like magic! How can it be done?" However, it is surprisingly simple. Raindrops do not fall at a very high velocity. Their sizes are 4 mm or 10 mm in diameter at most and they fall no faster than 10 m per second. Because human eyes do not react quickly, they look like streaks. If you shorten the exposure setting of your camera to, say, 1 msec, or 1/1,000 of a second, to capture raindrops, they indeed appear as droplets in the camera image (Fig. 1). It is not very difficult to determine where they are located in the image.

So, what you would do is to get a high-speed camera and a projector ready, and place them together at the same position behind an optical lens (Fig. 2). Technically,



Fig. 1

た配置といいます。実際には二つの物体を同じ位置に置くことは物理的にはできませんから、一方からの光は通り、逆向きの光は反射するハーフミラーと呼ばれる装置を45度の角度に置けば、二つを同じ位置に置いたことと同等になります。この共焦点配置の特徴はカメラの視線とプロジェクタの光線の方向が一对一に対応していることです。だから、カメラ画像中のある画素に雨粒が見えれば、それに対応するプロジェクタの光線が簡単に決まります。

スマートヘッドライトの動作は次のようです。まず、非常に短い時間だけ、プロジェクタのすべての光線をONにして光を前方全体に送り、すぐにカメラで撮影します。先ほど言ったように雨粒は白く映りますから、それらの位置をすばやくコンピュータで検出して、対応するプロジェクタの光線はOFF、それ以外はONにすればいいのです。このプロセスを繰り返すだけです。実際は、完璧にはいきませんから、少しは雨粒に光は当たります。まったく当たらないことだけが目的なら、ヘッドライトをOFFにするのがベストの方法ですが、それでは運転できません。目標は送った光のうち何%がすり抜けていくかの割合(スルーput)を上げることです。また、いくら「すばやく、速く」といっても、画像を処理するのに少しは時間がかかります。その間に雨粒は少し移動しているはずなので、その分を考慮する必要がありますが、その詳細はここでは説明しなくてよいでしょう。

結局、スマートヘッドライトは共焦点に置いたカメラとプロジェクタを、画像を処理し光線を制御するコンピュータで繋いだものです(Fig. 3)。実際に試作してみました。会議用のプロジェクタを改造します。実は、現在のプロジェクタというのは、百万個ほどの小さな小さな鏡が縦横に並べられたDMDチップと呼ばれるものが使われています。強いランプからの光をDMDチップに当て、それらの鏡の方向を一つひとつ制御することで自由な画像が投影できるのです。しかし、会議用のプロジェクタ

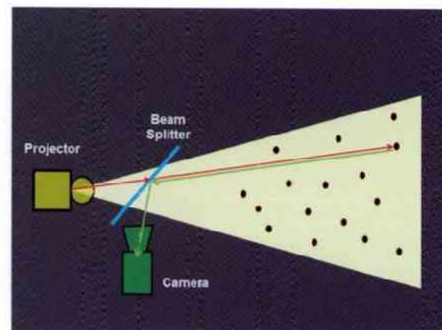


Fig. 2

this setup is called a confocal arrangement, that is, they will be positioned to share the same light focal point. In reality, it is physically impossible to position two objects in exactly the same place, and so you use a device called a beam splitter, which allows the light from one direction to pass straight through, but reflects the light from the other direction, at an angle of 45 degrees to the incoming and pass-through beams. Then you get an arrangement equivalent to that of the two objects being placed at the same location. What is unique about this confocal arrangement is the fact that the camera's line of sight and the direction of light from the projector have a one-to-one correspondence. So, if you can see a raindrop in a certain pixel of the camera image, you can determine the corresponding light beam emitted from the projector.

Now, smart headlights function like this. For a very short time, all of the projector light beams are turned on and emitted forward, and an image is captured immediately with the camera. As I mentioned earlier, the raindrops will glow white, and thus all one has to do is to have the computer quickly detect their locations and turn off the corresponding light beams from the projector, while keeping all of the rest turned on. The process is simply repeated. In reality, avoiding raindrops to be shined can't be perfectly achieved, and some light will inevitably hit raindrops. If the sole purpose were to prevent raindrops from being lit, the best thing you can do is to turn off the headlights completely. That is not good. The actual goal is to increase the throughput, that is, the percentage of light that would pass between the raindrops. Also, no matter how quickly you try to do the above, it will still take a little time to process the images and control the headlight. In the meantime the raindrops will have moved slightly, and that must be taken into consideration, but I won't bore you with all of the details of that, right now.

In summary, a smart headlight consists of a camera and a projector in a confocal position that are connected by a computer, which processes the images and controls on-off of the light beams (Fig. 3). We made a prototype by remodeling a projector used for conferences. Built inside the projectors we use today is a digital micromirror device chip, or a DMD chip, on which about a million tiny mirrors are aligned in an array. By emitting light from a strong lamp onto a DMD chip and controlling the directions of those tiny mirrors individually, it is possible to project any image that you want. With a conference projector, you only have to switch around sixty images per second at most, but we wanted to control the lights much faster, so we modified the projector. Then, using a beam splitter, we placed the camera at a confocal position. We connected them with a computer and put a cover on it (Fig. 4). If you equip a car with the devices we've

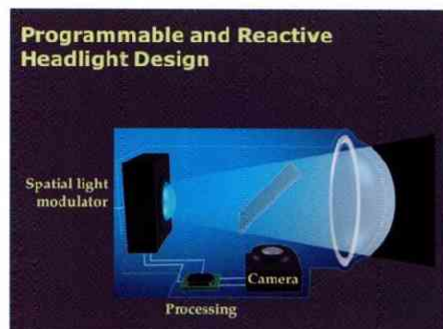


Fig. 3



Fig. 4

は1秒間にせいぜい60枚の画像を入れ換えるだけで十分ですが、われわれはもっと速く制御したいので改造します。次に、カメラをハーフミラーを使って共焦点の位置に据えます。それらをコンピュータでつなぎ、カバーをかぶせます。出来上がった装置を自動車に載せます(Fig. 4)。新しいスマートヘッドライトのついた車の出来上がりです。

そのスマートヘッドライトで雨を見ると確かに雨はあまり見えません。それはそうでしょう。水滴には光が当たらないように設計したのですから、水滴からの反射は少なく人の目にはよく見えないはずですが、雪はどうでしょう(Fig. 5)。これは人工雪でテストしたのですが、確かに雪が薄く見えます。

雨や雪を消すスマートヘッドライトが実際に使われるのは少し先のことでしょう。しかし、環境に応じて光線を自由に制御できるヘッドライトが一旦実現できることがわかれば、もっと多くの実用的な能力を考え付くことが可能になります。直近の応用は夜の運転時におけるハイビームとロービームの切り替えの問題です。夜道、特に照明のない田舎道を運転するときは遠くがよく見えるように、ハイビームで運転したいです。しかし、対向車が来ればロービームに変更する必要があります。ハイビームの強い光線が対向車の運転者の目に入り、グレアと呼ばれる一時的に目の見えない状態を作り出して危険だからです。でも、面倒でときには変更が遅れたり、忘れたりすることもあります。何か解決策はないでしょうか。

ヘッドライトの光線のうち対向車の運転者の目に入る光線だけをOFFにしてやるのです(Fig. 6)。相手は、目に光が入らないので、まぶしく感じるはずはなく、一方こちらはハイビームのままでよく見えます。目の位置を知るのは出来ないかも知れませんが、カメラに大きな明るく丸いものの組が見つければそれは対向車のヘッドライトでしょう。アメリカであればその間の少し上の向かって右側、日本では左側が運



Fig. 5

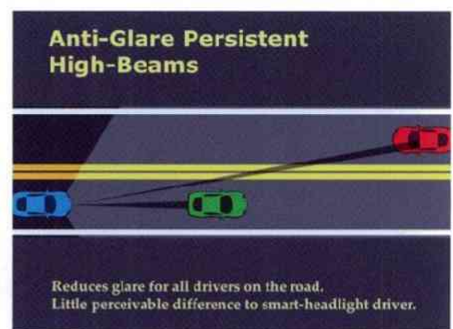


Fig. 6

created, you will have a vehicle with a brand new set of smart headlights.

Using these smart headlights to look at rain, for sure, you will not notice it as much. No wonder. We designed it so that no light would hit the water drops. Because the water drops will reflect only a small amount of light, they will not be very visible to human eyes. What about snow? (Fig. 5) We tested this with artificial snow and the snowfall certainly appeared lighter.

It will take some time before smart headlights that make rain or snow disappear will be commercialized. Once you know that it is possible to create headlights that can control light freely depending on the circumstances, it becomes to come up with many other practical applications. One application that we have been working on is about the issue of switching between high and low beam while driving at night. You want to drive with high beam so that you can see far ahead when driving at night, particularly on country roads without street lights. But, when there's a vehicle coming toward you, you need to switch to low beam. This is because the intense high beam light will strike the oncoming driver's eyes and results in glare, or temporary blindness, which can be dangerous. However, this is a cumbersome task and, sometimes, you might get the timing wrong or even forget to do it. Is there a good solution to this problem?

The simple solution is to turn off only that part of the headlight beam that could strike the eyes of the oncoming driver of the oncoming vehicle (Fig. 6). The driver will not be dazzled because there will be no blinding glare, while you will be able to see just as well as you would with the lights on high beam. It may not be possible to pinpoint the precise position of the eyes of the driver of the oncoming vehicle but, if the camera sees a pair of large, bright, and round objects, whose between-distance is getting larger as the time goes, chances are that they are the headlights of a vehicle coming toward you. In the United States, the driver in the oncoming vehicle would be at slightly above and to the right of the center of the two lights; in Japan, that would be on your left. You would

転席ですから、そのあたりに向かう光線をOFFにすればいいのです。やってみると、こんな風に、自分はハイビームのままでも、相手にはまったくまぶしく感じません (Fig. 7)。つまり、自分から常にハイビーム、相手には常にロービームとなります。

安全に役立つ良いこともできます。現在の車は自動運転の機能によって自分の走行すべきレーンを知っていますから、それを運転者に知らせるように表示できます。道路の境界が写|かれていない田舎道では特に有効です。もちろん、「ここで左折」といったナビ情報などを直接道路に表示することは簡単なことです。さらに、現在では、レーザセンサーや赤外線カメラによって、人の目よりも速く、遠くの歩行者や自転車を検出することができます。だから、より多くの光をそちらに送ってやることで、運転者には見やすくなり安全な走行が出来ます。現在では、このようなシステムをトラックのヘッドライトに組み込めるレベルにまで達しています。こんな風に、次々と新しい機能も思い付くことができるでしょう。

さらによく考えると、カメラとプロジェクタは、光の方向が外から内へ入ってくるのか、内から外へ出て行くのかの違いだけでそれ以外はまったく同じ性質を持っている双対の関係にあります。だから、現在のプロジェクタは光源からの光をDMDで2次元位置と方向に制御する装置ですが、同様に外からの光を2次元位置と方向に制御して画像面のセンサーでとらえるカメラが考えられます。そんな双対のカメラとプロジェクタを組み合わせれば、光線を入出それぞれについて、2次元と、方向の3自由度、合計6自由度制御できるシステム—カメラ・プロジェクタ・システム—を考えることができます。カーネギーメロン大学(CMU)の私の同僚であるスリニバス・ナラシムハン教授はそんなカメラ・プロジェクタ・システムによる新しい、今までになかった機能をもつセンサーを数多く作っています。

素人のように、、、玄人として、、、

お話したスマートヘッドライトの研究開発は、今日の講演のテーマである「素人のように考え、玄人として実行する」をよく表していると考えます。運転時に雨粒が視界を悪くする効果を減らそうとすると、「ヘッドライトの光が当たるのがその原因なら、当たらないようにすれば良い」と考えるのは、これほど直裁的で当然な発想はないでしょう。まさに素人考えです。実は、専門家は、雨粒それぞれの3次元位置を正確に高速に計測するのはきわめて難しいし、ましてや、それに当たらないよう

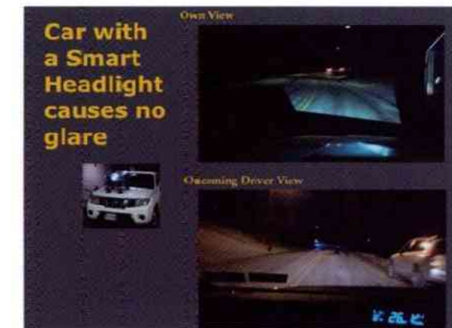


Fig. 7

turn off the part of light beams directed toward that areas, and then the oncoming driver is not blinded by glare, even when your headlights are on high beam (Fig. 7). Thus, the lights are always set on high beam for you, but appear to be on low beam to the other driver.

Such devices can also improve safety. Today's vehicles “know” which lane they should be traveling in, thanks to the vision technology for automated driving. The driving lane can be displayed. This is particularly effective on country roads, where road boundaries are not always clearly drawn. Of course, it would be a simple matter to display information from a car navigation system, such as “Turn left here,” directly on the road. Furthermore, today's laser sensors and infrared cameras can detect pedestrians and bicycles from a distance more quickly than the human eye. By shining more lights in that direction, the driver should find it easier to see far ahead and drive safely. If you think in this way, more new functions, one after another, will come to your mind.

You will find further that the camera and the projector form a dual relationship; they both have exactly the same properties, aside from the direction of light, that is, whether the light comes in from outside or goes out from inside. Current projectors use DMDs to control lights from the source in two-dimensional locations and directions. Likewise, you can think of a camera that controls lights in two-dimensional locations and directions before capturing them with a sensor on the image surface. By combining a camera and a projector in a dual relationship, a camera-projector system could control incoming and outgoing light with six degrees of freedom – two positiona and one directional freedoms, multiplied by two. Professor Srinivasa Narasimhan, my colleague at Carnegie Mellon University (CMU), uses such camera-projector systems to develop many sensors with never-seen-before functions.

に光線を正しく制御することはさらに難しい—実際そうなのですが—ということを知っているで、ほとんど瞬間的にこの素人考えを拒否してしまいます。皆さんにも経験があるでしょう。何かをしようとする提案に対し、最初に否定的な意見を言う人の多くはその道の専門家です。「それは難しい」「そんな風にはやらないものだ」と。私のような大学教授、常に人よりもものを知っているということを知らしめなければならぬという一種の強迫観念をもっている人間は、学生が何か良いことを言うと「君は知らないのかね。それは1965年に誰々がやって失敗している」ということを言いたくなるものです。

しかし、ひとまず、この素人考えを受け入れたとしましょう。その実現方法を実行するのは専門家の知識と技能が必要です。例えば、自動運転において、車が左に寄ればハンドルを右に切る、逆に車が右に寄ればハンドルを左に切るというプログラムを書いたとしましょう。これで車が真っすぐに走るかという走りません。蛇行します。なぜかという、制御理論における利得という概念がプログラムに入っていないからです。スマートヘッドライトのキーとなるのは共焦点というやり方でした。それによって雨粒の3次元位置を知る必要はなく、画像中の位置だけを知ればよいということに結びつけるのは専門家の知識です。さらに、実際にカメラとプロジェクタを極めて高速かつ正確に処理と制御することでその原理を実現するのは専門家の技能がなければできません。ただ、言っておかねばなりません、私は「素人と専門家が集まって、素人がアイデアを出し、専門家が実現しろ」と言っているのではありません。あくまで、われわれ一人ひとりの研究者や技術者は最高の専門家であるべきなのですが、考えるときには「まるで」素人のように、実行するときは「本当の」玄人として、ということです。さらに、このスマートヘッドライトのアイデアには面白く発展性があります。この装置の原理とその意味するところがわかれば、ハイビーム・ロービームなど次々と新しい使い道を考え付くことができました。さらに、この装置の考え方をカメラ・プロジェクター・システムとしてもっと一般化できました。また、皆様もそう思っていたら幸いです。私がこの話をしますとどこでも聴衆の方はほぼ一様に、「面白い。こんな風にすればもっと役に立ちませんか。私のもつこの技術をここに使えませんか」などと反応されます。つまり、人を刺激し考えを広げる効果があるようです。

私自身40年以上にわたって計算機視覚と知能ロボットの研究をしてきましたが、私のアプローチはここで申し上げたように、ある意味極めて単純な考えに基づいたも

Like an amateur. As an expert.

I believe this story about research and development of smart headlights demonstrates a perfect example of the topic of this lecture: “Think like an amateur, do as an expert.” When you want to prevent raindrops from obstructing your view while driving, no idea is more straightforward and natural than to think, “If the obstruction is that the lights from the headlights hit the raindrops, the cure is to prevent it.” This is such an amateurish opinion. Any expert would know that it would be extremely difficult to accurately and rapidly measure the three-dimensional positions of each raindrop and that – and this is actually the case – it would be even harder to control the lights so that they don’t hit the raindrops. This knowledge would make an expert almost instantly reject the above amateur’s idea. I am sure that you have had similar experiences. If you come up with an idea for a new initiative, in most cases it will be an expert in the field that will make a negative comment first. They’ll say things like, “That would be too difficult” or “You just don’t do things like that.” In fact, people like me – professors at universities – tend to have an obsession that we must let people know that we know better than others; so when a student makes a suggestion, we are tempted to say something like, “Didn’t you know? Someone tried that in 1965 and failed.”

Let us, for now, accept this amateur’s idea. You would actually need expert knowledge and skills to make it happen. For instance, imagine that for automated driving you write a program that commands to steer to the right when the car swerves to the left and, in the opposite case, to steer to the left when the car swerves to the right. You might expect that the car would travel straight. It will likely weave back and forth instead. Why? The concept of “gain” in control theory is missing from the program. Recall that the key to implementing the smart headlight was using the confocal placement of a camera and a projector, so that it is not necessary to know the three-dimensional locations of raindrops, but just their locations on images. This is the expert knowledge. Also, you need expert skills in order to put that principle into practice, that is to use a camera and projector to process and control data accurately and rapidly. Mind you, though, I am not saying, “We need both amateurs and experts, with amateurs offering ideas and experts giving shape to them.” Each individual researcher and engineer must be an expert of the highest level, but he or she needs to think without preconceived ideas “as if a novice” and do methodically “as a real” professional. I would also like point out that we found the idea of smart headlights not only interesting but expansive. Once we correctly understood the principle of the smart headlight, we were

のでした。その中で研究結果が世の中に役立つことを一番に考え、また私自身はそういう活動を楽しんでやってこられた幸運に恵まれたと思っています。そういう私の研究生活を子供時代から振り返ってみたいと思います。

私の生い立ち：私の幼年期から大学卒業まで

私は戦後すぐの1945年10月に兵庫県氷上郡大路村栢野(かやの) (現在の丹波市春日町)という土地に生まれました(Fig. 8)。いろいろなことを覚えています。田んぼの中の小さな借家住まいでした。土間の台所、縁側、その前には小さな庭と井戸がありました。家の裏には小さな川があり、夏には遊んだり、どじょうをざるですくったりしました。小学校は家から少し登った県道沿いにありました。1年生として入学したときは1クラスしかなく、10人いるかいらないくらいの小さな学級でした。もっとも、記憶に残っていないこともあるようです。私は姉2人、兄2人の5人兄弟の末っ子なのですが、私をおぶって子守りをさせられたという姉たちによると、私の性格はどうしようもないほど負けず嫌いで意地っ張り、一旦言い出したら地面にしがみついて、てこでも動かなかったと言います。てっきり物分りの良い、大人しい子供だったと思っている私からすると、想像もできないことですが、姉2人が言うのですから多分そうだったのでしょう。1年生の夏には事情で神戸市に移り住み、その後は小学校卒業まで神戸市の中を2ヶ所ほど転々としながら、中学・高校と進みました。

家庭ははっきり申して極めて貧しく困窮状態でした。そのことは小学校に入る頃には子供心にも十分認識しておりましたから、小遣いをねだるということはしたことがない子供だったと思います。母はその娘時代に本格的に習ったという裁縫で、呉服店

able to come up with many other applications, such as the high/low beam headlights. We were also able to generalize the idea of this device in the form of a camera-projector system. Finally, whenever I talk about this story of smart headlights, my audience almost always reacts like, "That's interesting. Wouldn't it be even more useful if you do it like this? Couldn't you use my technology here?" In other words, it seems to have the effect of stimulating people and helping them to expand their own thinking framework.

I have worked on computer vision and intelligent robots for over forty years, and my approaches have been, as I have just said, based on very simple ideas in a sense. My priority has always been to find ways of making my research findings benefit the world, and I count myself fortunate to have enjoyed my work so far. Now, I'd like to tell you a little bit about my career as a researcher, starting from my childhood.

My early years: From childhood to university graduation

I was born in October of 1945, just after the end of World War II, in a place called Kayano in Oji Village, Hikami District, Hyogo Prefecture (now Kasuga Township in Tamba City) (Fig. 8). I have many memories about that place. My family rented a small house in the middle of a paddy field with an earthen-floored kitchen and a narrow veranda, which overlooked a small garden and a well. Behind our house was a small stream, where I would often play and catch loaches using a bamboo basket in the summer. My elementary school was located a bit up along the prefectural road from my home. When I started school there as a first-grader, we only had one very small class of maybe 10 students. It seems that there are also things that I don't remember. I was the youngest in a family of five children, with two elder sisters and two elder brothers. My sisters, who were babysitting me, probably reluctantly, by carrying me around on their backs, say that I was hopelessly obstinate. According to them, once I started to demand something, I would dig my heels in and refuse to budge an inch. As a person who remembers myself a very obedient good and quiet child, I can't believe or even imagine such a story, but since both of my sisters insist on it, I suppose I could have been like that. By the summer of my first year in elementary school, my family moved to Kobe City, where I spent the rest of my elementary school years, and went on to junior high and senior high school, while moving with the city twice.

Quite frankly, my family was very poor and badly off. Even as a child, by the time I started at elementary school I fully understood the situation and I don't think I ever pestered my mother to spend money on me. My mother made a living by doing work of



Fig. 8

からもらった仕事で着物を朝早くから夜遅くまで作ることで生計をたてていました。縫った後で縮むのを防ぐために、呉服の生地はあらかじめアイロンがけをしておくことが必要なのですが、小学生の低学年の時にはもうそんなことを手伝っていたのを覚えています。母は丹波の農家の出で、高等小学校を卒業しただけでしたが、静かで賢明な女性でした。大きな声を上げて叱られたような記憶はほとんどありません。故事や熟語などを大変よく知っていて、裁縫しながら、いろいろな昔話、歴史にまつわる話や人物について教えてくれました。「無駄遣いをしない」「過ぎたるは及ばざるが如し」「人はうまくいっているときにどれほどのことができるかではなく、困難にあったときにどうするかで、値打ちが決まるのだ」とか、さまざまな人生哲学はすべてと言っていいほど母から学んだように思います。

よく、「子供の時からロボット工学の研究をすることを考えていましたか」と訊かれるのですが、特にそういうことはなかったように思います。強いて言うなら、子供の頃からいろいろなものをその辺にある材料から自分で作ってみて試すことが好きでした。今も記憶にありますが、小学校に上る前に、魚の釣り針を自分で作ろうと、針金をペンチで曲げて形だけは釣り針の形にし、糸に結び、ミミズの餌をつけて、竿の先から川にたらししました。しかし、餌はとられるばかりで何度やっても魚は釣れません。魚を逃がさないためには、釣り針に戻りという逆向きのとがりが必要であることに気が付くことができなかったのです。その悔しさは今でも忘れません。そのためか、今でも魚釣りは好きになれません。それでも、今も日曜大工などをはじめ、物づくりは大好きです。家庭でも壊れたものは、家具から電気製品に至るまで一度は分解して修理を試みる趣味があります。結構直してしまうので古い製品をいつまでも使い続けるところがあり、私の妻は「あなたに製品を売った会社は、一回きりしか買ってもらえないから、良くない客ですね」と言います。

1964年に高校を卒業し、京都大学の工学部電気・電子工学科に入りましたが、正直あまりよく考えていなかったと言ってよいでしょう。私の兄が「京都がよい」と言うので京都にし、工学部にしたのは、当時は経済成長の時代で猫も杓子も工学部だったからであり、電気・電子工学科にしたのは、募集人員が一番多かったので「競争率は同じでも落ちる確率は低いだろう」という、わかったようなわからないような私なりの確率論でした。この後の研究生活でもそうですが、私は意外と人の言うことをあっさり聞き、影響され、簡単な理由で物事を決めるようです。大学ではずいぶん張り切っていました。京都の紫野とか北白川という歴史のかおりのあるところに下宿し

sewing Japanese kimonos from early in the morning to late at night. To prevent kimono fabric from shrinking after being sewn, it is necessary to iron it, and I remember helping her do this from as early as when I was in the lower grades of elementary school. My mother came from a farming family in Tamba. While she only graduated from higher elementary school, she was a very quiet, wise woman. I hardly remember her raising her voice to scold me. She knew about a great many historical events and interesting sayings, and while she sewed, she used to relate various old tales and tell me all about famous incidents and people from history. In retrospect, I learned virtually all of my life philosophy from my mother: “Don’t waste money,” “Too much is as bad as too little,” and “A person’s real capability makes itself appear not when things are going well, but when things are difficult.”

I’m often asked if I was thinking about doing robotics research from my childhood. I don’t think I was. If I had to think of any connection, I enjoyed making all sorts of things from whatever was at hand around me to see how they worked. There is one thing I remember to this day. Before entering elementary school, I decided to make my own fishhook. I bent a wire with pliers into a hook shape, tied it to thread, put a bait worm on the hook, and dangled the line in the river from a rod. However, I kept losing the bait and couldn’t fish, no matter how many times I tried. That was because I didn’t realize that I needed a barb, which is a backward-facing point, on my hook so that the fish wouldn’t get away. I still remember how vexed I was. Perhaps because of this experience, I never came to like fishing. Anyway, still I truly love making things and doing things myself. At home, I disassemble anything, whether it be furniture or electrical appliances, and try fixing it if it’s broken to the point that it is my hobby. I would say my success rate is quite high, and I end up using old products forever. My wife says to me, “You’re not a good customer for salespeople, because you buy from them only once.”

I graduated from high school in 1964 and went on to enter the Kyoto University’s Electrical and Electronic Engineering. To tell the truth, it was not necessarily a decision of any particular thought. One of my brothers suggested that I should go to Kyoto University, instead of the University of Tokyo, which I did. I chose the Engineering school because it was everyone’s top choice at the time when the Japanese economy was experiencing rapid growth. I picked Electrical and Electronic Engineering because its quota was the largest. I figured that the chances of me failing there would be low even though the competition rate was the same with other departments with smaller quota. It might sound like nonsense but this was the probability theory that I made up.

ていましたので、京都の神社仏閣、郊外の散歩道は大体制覇しました。その意味では、特段変わった学生ではなかったと思いますが、強いて言えば当時の学生の必須科目とも言うべきマージャンはまったくやらなかったことと、授業で習ったことで面白く感じた理論や分野があれば、名著とされる本や英語の教科書を買ってきては一人で読破して悦に入っていたことでしょうか。例えば、朝永振一郎博士の名教科書『量子力学』やウィリアム・フェラーの確率論の名著など、手垢のついた本は今も書斎の本棚にあります。それらから得た知識自体もさることながら、そういう本を読むことができたという自信はその後の研究者生活に良い影響を与えたことは間違いのないでしょう。

京都大学大学院で研究者の道へ

1968年に京都大学を卒業して、その大学院に進みました。選んだのは坂井利之教授の研究室でした(Fig. 9)。坂井先生は、「コンピュータは数を計算するだけの機械ではない。画像や音声処理したり、認識したりするのに使うべきだ」と強調されていました。パソコンや携帯で画像や音声処理が扱える今では当たり前ですが、当時としては斬新な考え方で、大きな影響を受けました。さらに、デカルトがどうこうというほどの深い哲学的思考の故ではありませんが、私は人間機械論をまったくそのとおりで思っていましたから、アラン・チューリングの書いた知能に関する随筆や、マービン・ミンスキーの“*Steps Toward Artificial Intelligence*”といった論文を読んだりして、「人と同じか、人より賢い人工知能ができるのは自明の理」という信念が固まったのもこの頃でした。

この坂井研究室での経験が私のその後の研究者生活の骨格を作るものになったとい



Fig. 9

Even in my later years as a researcher, it seems to me that I'm influenced easily by what others say to me, and make decisions based on simple and probably not well-thought out reasons.

I was brimming with energy at university. My lodgings were in historic places in Kyoto, like Murasakino and Kitashirakawa, and I visited most of the shrines, temples, and promenades in the surrounding areas. That was I was a typical student. If I were to think of anything that made me different from others, I would say that I never played the game of mah-jongg, which seemed to be a “mandatory” subject for college students at that time. Also, when I came across an interesting theory or field in class, I made it a rule to buy classic books or English textbooks on the subject and read through them, something that I found personally very satisfying. Some of those books included *Quantum Mechanics* by Shin'ichiro Tomonaga and William Feller's classic book on probability theory, and those well-thumbed volumes are still on the shelf in my study. Certainly what I gained from that experience was not just knowledge but also the confidence that I was able to read them, and this definitely had a positive impact on my subsequent life as a researcher.

Beginning a research career at the Kyoto University Graduate School

After completing the undergraduate course at Kyoto University in 1968, I moved on to the university's graduate school. I chose Professor Toshiyuki Sakai's laboratory (Fig. 9). Dr. Sakai already stressed that, “Computers aren't just for calculating numbers – they can and should be used to process and recognize images and sounds.” We now take it for granted that images and sounds can be handled using our computers and mobile phones, but his idea was extremely novel at the time, and it had a profound impact on me. I would never presume to be as deeply philosophical as Descartes, but I completely agreed with the theory of the human machine, and read Alan Turing's essays on intelligence and Marvin Minsky's paper “Steps Toward Artificial Intelligence.” It was around this time when I became convinced that, “it is self-evident that we will someday have artificial intelligence on a par with or even surpassing human intelligence.”

The experience I gained at Dr. Sakai's laboratory provided a framework for my subsequent career as a researcher. I came away with several things from there, especially my attitude toward research. Dr. Sakai's favorite phrases were “first in the world” and “interesting.” At that time, Japan was far behind the United States in the



Fig. 10

えるでしょう。いくつかのことを学びましたが、特に研究というものに対する姿勢です。まず、「世界で初めて」と「面白い」が坂井先生の口癖でした。情報工学の分野において、日本は当時明らかに米国に大きく遅れている時代ではありましたが、それでも世界一流の、そして面白い、つまり先進性と発展性のある研究を目指す姿勢は研究室全体の意欲を高めていました。これも坂井先生の口癖でしたが、「ハードしか知らない人間は、ソフトでちょっとやればもっと柔軟なものが作れるのにしない。ソフトしか知らない人間は、ハードでちょっとやればもっと高速なものが作れるのにしない。理論と応用も同じだ。応用のない理論は意味がないし、理論のない応用は当てにならない」。つまり、理論と応用、ハードウェアとソフトウェアの区別なく、すべてを考え、システムとして捉えるやり方です。最後の一つは、当時坂井研究室の助教授であり後に京都大学総長になられた長尾真先生に教えられた、「研究は具体的な目標を持ったものでなければならない」ということです(Fig. 10)。そんなことは当然だろうと思われるかも知れませんが、これが結構難しく、研究者は往々に忘れるものだという点をわかっていただくには、私が博士論文の研究で随分苦労した話をする必要があるでしょう。

数学の試験問題を解いたりするのは結構得意な方でしたから、理論の論文を読んでは、「それならこういう定理や理論があるのでは」と思って取り組みました。しかし、研究は試験と違って、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かどうかかわかっていませんから、大学院生がちょっと考えたくらいでそう易々とできるほど甘くはなく、すぐに行き詰まります。これはいかんというので、また次の課題、そしてまた次と、いろいろやっているうちに、随分いろいろな論文を読んだので、物知りにはなりましたが何もできません。単に、「この論文のここを拡張したい」とか、「ここをもっと数学的にして格好良い理論にしよう」といったことそれ自体が自己目

information technology field, but his attitude of aiming for “world-class” and “interesting” research – that is, “advanced” and “expansive” research – seemed to motivate everyone at the laboratory. Dr. Sakai also often said, “If people know only about hardware, they won’t use software even when they could make something much more flexible by using a few lines of software. If people know only about software, they won’t use hardware even when they could make something much faster by using a small piece of hardware. The same holds true for theory and application. Theory without application is meaningless, and application without theory cannot be relied upon.” To sum up, you need to think about everything, irrespective of theory and application or hardware and software, and instead see them all as parts of a system. Another thing that I learned at the laboratory is from Dr. Makoto Nagao, then an associate professor in Dr. Sakai’s laboratory, who later went on to become President of Kyoto University. He said, “Any research must have a concrete goal (Fig. 10).” You may think that this is obvious, but it is much easier said than done. To help you to understand how often researchers forget about this simple rule, perhaps I need to tell you how much difficulty I had doing research for my doctoral thesis.

As an undergraduate student I was fairly good at solving math exam questions. So, when I read theoretical papers, I could think, “Ah, I could modify it into such and such big theorems.” However, unlike math exams, you never know beforehand if there is an answer to it in the first place, let alone if a research question is worth solving. It was not easy for a graduate student with a little experience to make a big discovery, and it wasn’t long before I started running into dead ends. I began to grow anxious and tried a next topic. And fail, and next, one topic after another. As I attempted various topics, I read many papers, and became knowledgeable, but I still couldn’t do anything. What I did was actually things like “expanding on that point in this paper” or “making this point more mathematical in order to make the paper look better.” However, such attempts easily became ends in and of themselves, but were devoid of concrete objectives. This form of approach made it difficult to determine the course of my research. About two years passed almost in the twinkling of an eye, and I became quite nervous thinking that I would end up achieving nothing.

It was a piece of advice above mentioned from Dr. Nagao that saved me during this time of crisis. He said to me, “Mr. Kanade, you should set a concrete goal for your research.” He also told me about the existence of digital image data containing as many as 1,000 human faces. “If only you could make a program that would process this data properly, that in itself would be a superb achievement,” he continued. Those image data

的化した活動には、これといった具体的な目的がないので研究の方向が定まらないからです。あっという間に2年ほど経ってしまい、このまま何もできないで終わってしまうのかと焦りが出てきました。

そんな危機を救ってくれたのが、先に述べた長尾真先生の「金出君、もっと具体的な目標を持った研究をするべきだ」というアドバイスでした。それとともに、1000人もの顔のデジタル画像データの存在を教えてくださいました。「これらが正しく処理できるプログラムができれば、それだけでもう立派なもの」と言うわけです。その画像データは1970年の大阪万博で私の同級生であった木戸出正継氏(後に(株)東芝勤務を経て奈良先端科学技術大学院大学教授)が作ったプログラムによるアトラクションが出品され、参加した入場者の顔画像データを記録したものでした。デジタル画像が簡単には手に入らない時代において、1000枚のデジタル顔画像はまさに破格のデータベースだったのです。確かに、具体的な目標を持つと、あれよあれよと研究が進み、結果として「顔画像の処理と認識の世界で最初の本格的な研究」と認められるものになりました(Fig. 11)。

助手になり、そしてアメリカ留学へ

1973年に大学院を卒業して、京都大学の助手になり、しばらくして助教授になりました。1975年、後に私の人生を大きく変えるきっかけになる方が訪問されました。人工知能研究の元祖の一人とされる、CMU計算機学科のアレン・ニューウェル教授です(Fig. 12)。当時、博士課程大学院生だった大田友一氏(後に筑波大学教授)とやっていた屋外風景の認識プログラムを見せましたら、「うちでも似たようなことを



Fig. 11

were collected for an attraction at the Expo '70 held in Osaka, which used a program designed by my former classmate Masatsugu Kidode (later professor at the Nara Institute of Science and Technology after working at Toshiba Corporation). The dataset contained facial images of attendees at the event. In an era when such digital images were not easily available, the digital images of 1,000 faces was a very substantive database. True to his advice, once I had determined a concrete goal, my research got going, and my thesis research was eventually recognized as “the world’s first full-scale research in the area of processing and recognition of facial images (Fig. 11).”

From an assistant to studying in the United States

In 1973, I finished graduate school and became an assistant professor and then soon an associate professor at Kyoto University. In 1975, a person who would later change my life significantly came to visit my university – Professor Allen Newell from the CMU School of Computer Science. He is one of the pioneering figures in artificial intelligence research (Fig. 12). I showed him a program for recognizing outdoor scenes that I was working on with Yuichi Ohta, who was at that time a doctoral student and would later become a professor at the University of Tsukuba. He told me that researchers at CMU were doing similar research. I asked him if I could go to CMU to conduct my research and, in 1977, he invited me to come to CMU as a visiting research scholar for a year. I was hosted by Professor Raj Reddy, who was an authority on speech recognition and was at that time expanding the scope of his research to include video-related topics. By that time, I was already married and had my first son. We all moved to Pittsburgh and rented an apartment near the campus. Driving on the highways, seeing huge mansions on the way to the university, and shopping at massive supermarkets –

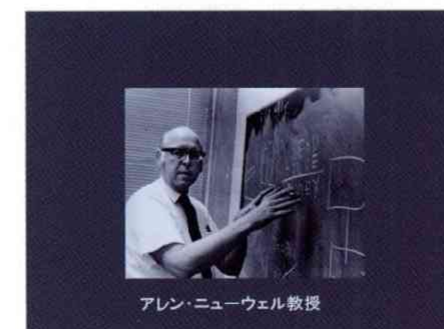


Fig. 12

やっている」と言われたので、「アメリカに行ってみよう」と尋ねたのがきっかけで、1977年から1年間の予定でCMUに客員研究員として招待してもらえることになりました。ニューウェル教授を通じて紹介された、音声認識の権威で画像関係にも研究を広げつつあったラジ・レディ教授がホストでした。その頃には既に結婚し、長男が生まれておりましたので、家族3人でピッツバーグに滞在し、大学の近くにアパートを借りて生活しました。高速道路の運転、大学に行く途中の大きな屋敷、とてつもなく大きいスーパーマーケットなど初めてのアメリカ生活はすべてが驚きでした。

CMUに行ってみると、アメリカでは計算機環境が格段に違っていました。日本では大学全体の計算機センターでのバッチ処理が普通で、Time Sharing System (TSS)は限られた形で始まったばかりの時代でした。一方、CMUでは、学科に大型TSS計算機が何台もあり、24時間稼働していました。オフィスには一人ごとにCRT端末があり、いつでも好きなときに計算機が使えましたし、家で使うCRT端末とモデムも支給してくれました。50ボー(Baud)という今では考えられないような超低速な回線でしたが、オフィスに行けないピッツバーグの極寒の日でも昼も夜も家でプログラムを書いていた。また、現在のインターネットの原型であるARPAネットというのがアメリカの主要な大学と研究機関を結んで稼働しており、メール、ファイル転送、トークと会話機能なども既にありました。ビットマップディスプレイや自由フォントや濃淡画像プリントも使えるようになっていました。要するに現在われわれが計算機を使うやり方が、もちろんスピードや機能は大きく違うものの40年前に全部行われていたのです。

さて、私の客員研究員としての研究です。私たち人間は、ここにあるような線図形から、その表す形がわかります(Fig. 13)。そういう「線図形からの形復元の方法」を研究することにしました。いろいろやってみるのですが、あまりすっきりとした結

my first stay in the United States was full of surprises for me.

At CMU, I was surprised to see how advanced the research environment for computer scientists was in the United States. In Japan, batch processing by the university computer center was standard, and the time-sharing system (TSS) had only just started in a limited fashion. At CMU, on the other hand, the CS department had several large TSS computers that were operating around the clock. In the offices, cathode-ray tube (CRT) terminals were assigned to each individual researcher, allowing them to use a computer whenever they wanted. They even provided another CRT terminal and modem for use at home. The communication line was extremely slow – 50 baud – which is unthinkable nowadays, but I wrote programs at my home in Pittsburgh both day and night, even on freezing-cold days when I couldn't get to the office. Also, there was already the ARPANET, a prototype of the Internet, which connected major universities and research institutions in the United States and allowed researchers to send e-mails, transfer files, and talk and chat. There were also bitmap displays, free fonts, and grayscale picture printing. In other words, all of the functions that current computers have now were available forty years ago, albeit at very different rates of speed and performance.

Now, let me talk about what I did as a visiting research scholar. When we as humans see a line drawing like this, we can perceive its three-dimensional form (Fig. 13). This was my research topic as a visitor, namely, “How can we recover a three-dimensional shape from a line drawing?” I tried many different approaches with some success, but I became increasingly frustrated when I was unable to get clear, systematic and convincing results. One day, when I bumped into Professor Newell in the hallway, he asked, “How are you doing, Takeo? Why don't you come to my office and tell me what you're working on right now?” I replied, “I don't have particular results yet. I'll tell you when I get something interesting.” But he told me that I should simply come and talk. I visited his office and explained what I was doing. Like when he visited us at Kyoto University early, he always listened very carefully to what young people like me had to say and then ask questions. During our discussion, he challenged me: “What you're trying to do reminds me of the theory of David Waltz; what is the difference?” Waltz did well known work for a mathematical theory on the understanding of line drawings. It was one of the most successful artificial intelligence researches at the time, and I certainly knew about it. In fact, I had tried to apply his method to my example drawing of a chair, but it didn't work. On the spur of the moment, I replied, “I'm trying to find a more flexible method.”



Fig. 13

果は出てこず、フラストレーションがたまり始めていました。たまたま、ニューウェル教授に廊下で出会ったら、「タケオ、元気でやっているか？私の部屋に来て研究していることを教えてくれないか」と言われました。「特にたいしたことはできていません。できたらお話しします」と言うと、「やっていることを話せばいいのだ」とのことなので説明に行きました。前に、京都大学を訪問されたときもそうですが、私のような若造の話を全身全霊を傾けて聞いて、さまざまなことを質問されます。そのうち、「君のやろうとしていることはデイヴィッド・ウォルツの理論と似ているようだが、どこが違うのか」と訊かれました。ウォルツの理論というのは、線図形解釈に関するかなり数理的な理論で、当時の人工知能研究で最も成功した研究の一つだったので、私ももちろん知っていました。しかし、私の例題である椅子の図に彼の方法を試したところ、うまくいかなかったのです。ですから、「私はより柔軟な(more flexible)方法を目指しているのです」とひとまず答えておきました。

自分の部屋に帰ってきてから考えました。確かに、私の例題はウォルツの理論ではうまく処理できなかったのも、「それより柔軟」な方法を目指しているはずですが、しかし、その本当の意味はなんだろうと考え始めました。すると、アイデアが浮かびました。ウォルツの理論は、物体は体積のある要素からできていることを陰に仮定しているが、私の必要としている理論は、物体は厚みのない面からできていることも許すべきだということに気が付いたのです。この差を最も直裁的に表す例はこの「箱」の図です(Fig. 14)。「箱」を描くときにわれわれは普通右図のように描くけれど、ウォルツの方法は左図のように厚みも描かれることを期待しているのだと。実際、ウォルツの方法は左図はうまく認識できるのですが、右図を入力すると、「この図はありえない」と出力します。すなわち物体を構成する基本単位を体から面に変更

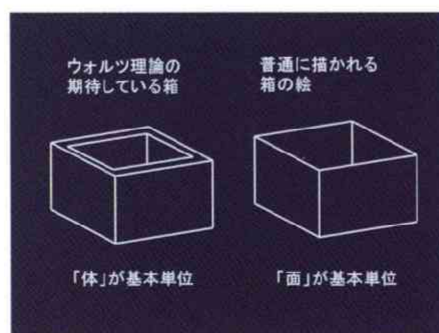


Fig. 14

Back in my office, I began to wonder. To be sure, the Waltz's theory could not understand my example, so what I am looking for is a "more flexible" way of doing things. Then I started thinking what this "more flexible" really meant. And an idea came to me. Waltz's theory implicitly assumes that an object is made up of elements with a certain volume, but I realized that the theory that I needed should also allow an object to be made up of only surfaces without any thickness. The example drawings of "boxes" most directly depict the difference here (Fig. 14). If we were to draw a "box," we would normally draw one like the box on the right, but Waltz's method expects us to draw thickness as well, like you see in the drawing on the left. In fact, Waltz's method recognizes the left drawing correctly, but if we input the right drawing, Waltz's method says that the drawing is "impossible". In other words, what being "more flexible" really meant was to change the basic unit that constitutes an object from a dimensional "body" to a "plane". With this small "discovery," despite some mathematical and programming difficulties, my research was rapidly set into motion toward a new theory of three-dimensional form recovery. I named it the "theory of the origami world," and I could develop its extension theory as well, which became my second successful work after my PhD thesis work. I still remember the excitement that I felt at the time of the "small" discovery. One of the lessons that I learned from this experience is the importance of talking and sharing your work with others. This doesn't mean that you always expect them to give you a good idea. Rather, when you speak to others, you can organize your thoughts, notice deficiencies, get feedback, and think through again. When I met Professor Newell later, I said to him, "Thanks to your comment, I was able to accomplish something great." His reply was, "Did I really ask you that question?"

I ended up staying at CMU as a visiting research scholar for 18 months until March 1979 before returning to Japan. On our way home, my wife, son, and I traveled around, visiting the Grand Canyon, California, and Hawaii (Fig. 15). I dropped by research centers and universities here and there to deliver lectures and I met many interesting people. It certainly was a luxurious trip.

Moving from Kyoto University to Carnegie Mellon University

Some time after returning home to Japan, Professor Reddy contacted me (Fig. 16). He told me that he was planning to set up a new research center called the Robotics Institute, and he wanted me to take the opportunity of its establishment to officially transfer to CMU. I was interested in the offer and consulted Dr. Sakai about it. He

する、これがまさに「より柔軟な」の意味だったのです。たったこれだけの「発見」で、数学的、プログラムの困難はあったものの、後は研究が一気に進んで形復元に関する、私が「折り紙世界の理論」と名付けた新しい理論とその拡張理論ができ、少しは知られる研究となりました。そのときの興奮は今でも覚えています。この経験から学んだことの一つは自分の考えを人に話すことの重要性です。必ずしも、人に話して相手に良い考えを教えてもらうというわけではありません。話すことで、自分の考えをまとめ、不備に気づき、反応を聞いてまた考えることができるからです。後でニューウェル教授に、「先生のコメントのおかげで良いことができました」と言うと、「そんなこと訊いたかね」との反応でした。

CMUでの客員研究員の生活は結局1979年の3月まで1年半続き、日本に帰国しました。帰国の途には、家族3人でグランドキャニオン、カリフォルニア、ハワイを旅行しました(Fig. 15)。また、私はそここの研究所や大学にも寄って、講演をしたり、いろいろな人に会ったり、贅沢な旅行でした。

京都大学を辞してカーネギーメロン大へ

帰国してからしばらくして、レディ教授から連絡がありました(Fig. 16)。ロボット研究所という新しい研究所の創設を計画しているので、それを機にCMUに正式に来ないかという誘いでした。「行ってみたい」と思い、坂井教授にご相談すると最初は反対されましたが、何度かお話しするうち「そこまで言うなら行ってみたらよい」と言ってくれました。そして、「行くからにはここに日本人ありと言われるようになること」という目標と、「日本人であることを武器に使ってはならない」というアドバイスをいただきました。後者は、反対のように聞こえるかも知れませんが、実に適切なアドバイスで、私はその深い意味を後によく理解できるようになりました。

こうして1980年の4月から、CMUに移籍しました。私の研究者生活の中で最大の決定であり、転換であったといえるでしょう。与えられた職は高等研究員というタイトルで任期5年の約束でした。安定した日本の大学のポストと比べれば不安定な職でしたが、私はそれほど深刻には考えていませんでした。客員で滞在したとき、「ここでは3倍の密度で研究できる」という感覚を持ちましたから、「5×3で15年分と思えばそれで十分」と言うほどの考えでした。というわけで、妻には「5年で日本に帰るだろう」と説得して出かけました。



Fig. 15

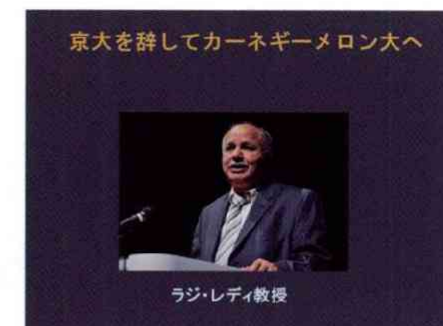


Fig. 16

opposed the idea at first. After several meetings, he finally gave in, saying, “If you really insist, you may go.” He told me, “If you’re going to go, make your presence felt as a Japanese researcher,” and also advised me to “Never use the fact that you’re Japanese as your advantage.” The latter advice may sound somewhat contrary to the common sense, but it turned out to be very appropriate advice. I did not fully grasp its deeper meaning until later.

I moved to CMU in April of 1980. I must say that it was the biggest decision I had made and a turning point in my career as a researcher. I was given the title of Senior Research Scientist with a five-year contract. Compared to the life time stability of a position at Kyoto University, the offered was rather tenuous. Honestly, I did not give much serious thought to the matter because, when I was at CMU as a visiting researcher, I came to be convinced that I would be able to do my research there with three times more intensity than I would have in Japan. I casually assumed, “five times three equals fifteen years, which is enough.” Thinking this, I convinced my wife that we would be back in Japan after five years and set off for the U.S.

And so, my life as a full-time researcher and educator at CMU began. Having been there once until the previous year, I thought that I knew most of the things that I needed to know. At the same time, I was a bit nervous and also fired up about the new opportunity, knowing that I had no university to go back to in Japan. Having become a regular faculty member, I discovered some things that I had not known or that puzzled me. For instance, about one week after my arrival, a graduate student visited me in my office and said, “I heard that you’ve just arrived here. I want to do research on computer vision. Will you take me as your student?” I thought it would be a nice thing to do, so I said, “Sure.” Then, after a week, another graduate student stopped by and said, “I want to be your student, too.” I began to feel good about my unexpected popularity and said, “Okay, of course” to that student. In time, I ended up having three students and began

CMUでの常勤の研究者・教育者の生活が始まりました。1年前までいましたから、様子は大体わかってはいましたが、今度は日本の大学に戻るところはないということで緊張し、張り切っていたのは事実です。正規の職員ということで、初めてわかったことも、戸惑ったこともありました。例えば、私が到着して1週間も経つと大学院生がオフィスに訪ねて来て、「あなたが今度来たことを知った。計算機視覚の研究をしたい。あなたの学生にしてくれるか」と訊きます。良いことだと思ったので、「もちろん」と言いました。しばらくすると、別の大学院生がやって来て、「私もあなたの学生になりたい」と言います。結構人気があるものだと思分が良くなって、彼にも「OK」。そうこうするうち3人ほどの学生ができて一緒に研究を始めました。すると、少し経って学科の経理担当オフィスから、「あなたの新しい3人の学生の学費はどのアカウントに請求するか」というメールが来たのです。「えっ、学費は私が払うの?」とびっくりしました。授業料と生活費は自前の日本と違って、アメリカの大学院生の授業料と生活費は、教授が獲得してきた研究予算から払う、つまりのところ雇っているのです。しかも、CMUのようなアメリカの私立大学の費用は高いので相当なお金です。就いたばかりの私にそんな研究予算があるわけはありません。困ってレディ教授に相談すると、彼のアカウントにつけておけということで助かりました。

当時、アメリカでは画像理解とよばれる大きなプロジェクトが遂行されていました。CMUもその重要なメンバーで、レディ教授が研究代表者(PIと言います)でした。私が、到着して2ヵ月ほど経ったとき、プロジェクトの会議で彼は「今日からタケオがPIだ」と宣言しました。研究プロポーザルを書かずして大きな研究プロジェクトの代表者になってしまったのです。日本から来たばかりの私にそんな思い切った決定をしてくれたレディ教授に感謝せずにはおれませんでした。このように、レディ教授は大胆な決定をし、大きな構想やプロジェクトを考え、それに人を巻き込んで実現させていくのが得意な人でした。私にはそのスケールはとても真似できるものではないと感じましたが、少なくとも研究は人より一歩先、スケールは一回りくらいは大きくする努力をすべきだと学んだ気がしました。

こんな風に戸惑ったり、びっくりしたり、感心したりすることもいろいろとありましたが、CMUでの教官と研究者生活にもおいおいと慣れていきました。アメリカ国防総省高等研究計画局(DARPA)やアメリカ国立科学財団(NSF)という研究支援機関を通じて、研究資金を得て、人を雇ってチームを作りプロジェクトを遂行していくアメリカの研究活動のやり方も学んでいきました。そんなダイナミズムを好ましく

working together with them. A little while later, I received an e-mail from the accounting office of my department, which asked me to which account they should bill the tuition of my three new students. My jaw dropped, and I said, "What? I have to pay their tuition?" Unlike common practices at Japanese universities, where students have to take care of their own tuition and living costs, professors at universities in the United States pay their graduate students' tuition and living costs out of their research budgets. In a nutshell, professors are their employers. And this was no small sum of money – CMU is a private institution in the U.S. Having just arrived, I did not have such any research grant. In my confusion, I spoke to Professor Reddy, and he kindly suggested that I could charge the expenses to his account, and thus I was saved.

At that time, a big project on image understanding sponsored by DARPA was underway in the U.S. CMU was one of the key members of the project, and Professor Reddy was serving as a principal investigator (PI). At a meeting for the project held about two months after I arrived, he suddenly declared that I would be a PI from that time on. I became the head of such a huge research project without even writing a research proposal. I couldn't thank him enough for his bold decision of assigning me to such an important position. Professor Reddy is a visionary and bold decision maker. He conceives major concepts and projects, involves others in them, and makes them happen. Though I could not possibly emulate him, at the very least, I believe that I learned from him that you have to make an extra effort to stay one step ahead of others in your research and think on a grander scale than others.

Constantly being perplexed, surprised, and impressed by one thing or another, I gradually and eventually got used to teaching and researching at CMU. I also learned how research activities work in the United States, a process of gaining research funds from research support organizations such as the Defense Advanced Research Projects Agency (DARPA) and the National Science Foundation (NSF) and hiring people to form a group to carry out the project. I came to like this kind of dynamism in idea development and project execution, and felt that it suited me well, so much so that I ended up having stayed at CMU for a total of 35 years, contrary to my original plan of five years. As shown on this slide, I have worked on a diverse range of research topics (Fig. 17), including basic methods and theories of computer vision like stereo, color and video image processing, and their applications in face detection, virtualized reality using numerous cameras, surveillance camera systems, and microscope image processing for cell tracking. In robotics, I worked on the invention and development of autonomous robots, including direct-drive robot manipulators using no gears, an autonomous

思い、それが自分にも合っているようにも感じました。こうして、5年間の予定が結局、その後35年間CMUに在ることになりました。そして、このスライドにまとめたようにさまざまな研究課題に次々と挑戦してきたように思います(Fig. 17)。ステレオ・カラー・ビデオ画像処理など計算機視覚の基本的課題、応用分野として、顔の検出、数多くのカメラを使う仮想化現実、監視カメラシステムや細胞の顕微鏡写真処理、ロボット分野ではギアを使わない直接駆動型ロボットマニピュレータの発明と開発から、ヘリコプター・ヒューマノイド・自動運転などの自律ロボットの分野、手術ロボットや生活の質工学として人に働くロボットシステム、などです。思いつくままと言っても良いほどです。これらのうちの2、3について、それらがどんな風に始まり、どんな方々と仕事をして進展し、また苦労したのかを少し話したいと思います。

自動運転 — “No Hands Across America (手を離してアメリカ横断)”

渡米して4年ほどの1980年代の中頃に始めたのが、自動陸上走行車というプロジェクトでした。カメラとコンピュータを積んで自分で道や障害物を見つけ、道路や一般的な地形を自動運転する車です。当時、米国が始めたStrategic Computingというプロジェクトの一環でした。その仕様を決めるのに、MIT、スタンフォード大、メアリーランド大の教授、スタンフォード研究所(SRI)の研究者、そして私の5人でDARPAの会議室で何日も会議したのを思い出します。アメリカのDARPA型プロジェクトの目標の設定の仕方—十分に未来先駆的ではあるが、実現可能性の感じられるレベルに設定する—を学んだ気がしました。

最初の頃は小さい荷車のようなものでやっていたのですが、1986年にはバン



Fig. 17

helicopter, humanoids, and automated driving robots. I also worked on surgical robots and robot systems that work for people, which I named the quality of life technology. You could say I do whatever comes into my mind. Now, I'd like to take this opportunity to talk about two or three of these topics: how they started, whom I collaborated with, and what difficulties and fun I encountered along the way.

Automated driving – “No Hands Across America”

In the mid-1980s, about four years after I moved to the States, the Autonomous Land Vehicle Project started. The project involved a vehicle that would drive on roads and over natural terrains by detecting streets and obstacles using onboard cameras and computers. It started as a part of a larger U.S.-government-initiated project called Strategic Computing. I remember that five of us – from the Massachusetts Institute of Technology (MIT), Stanford University, and The University of Maryland, the Stanford Research Institute (SRI, now SRI International) – spent several days discussing the vehicle's specifications in a conference room at DARPA. From that experience, I believe that I learned how they set goals for DARPA-model projects, that is, setting goals at a level that is forward looking and pioneering, but still deemed feasible.

For autonomous navigation research we started with a small wagon-like vehicle, but by 1986 we had been convinced that we need a self-contained robot vehicle for research. We converted a van into a robot vehicle equipped with four workstations, a camera, and a special laser distance sensor (Fig. 18). We named it Navlab, a portmanteau of “navigation” and “laboratory.” We started with slow cruising at about 5 km per hour along a path in a park next to CMU, then moving on to suburban roads and highways. This project lasted for more than 15 years, and we built Navlabs 1 through 10. Thanks to the research of many active doctoral students and capable staff members, we experimented with almost all of the functions required in today's automated driving vehicles, such as keeping the vehicle in the lane or changing lanes by detecting lane markings, parallel parking, detecting other moving vehicles and pedestrians, and detecting vehicles in the blind spot. Navlab had all of these, although they were still not perfect. The biggest event of the Navlab Project that has had a profound impact on automated driving research overall was the “No Hands Across America” tour of 1995, a demonstration of automated driving with Navlab 5 across the North American continent from Pittsburgh in the east to San Diego on the West Coast. For 98% of the approximately 5,000-km-long journey, our vehicle drove by itself at about 100 km per

にワークステーション4台、カメラ1台、特別なレーザ距離センサーを搭載した「NAVLAB」というロボット車を作り上げました(Fig. 18)。NavigationとLaboratoryを組み合わせて作った言葉です。はじめの頃は公園の中の道を時速5 kmくらいという速度でゆっくり運転していました。それから郊外の道路、高速道路とレベルアップしていきました。このプロジェクトは15年以上続き、NAVLABも1号機から10号機まで作りました。多くの博士論文研究やスタッフの活動によって、レーンマークの検出によるレーンキープやレーン変更、パラレルパーキング、他の走行車検出、歩行者検出、ブラインドスポットの後続車検出といった、今日の自動走行車に必要な機能はほとんどすべて手掛けられ、完全とはいえませんが持っていました。しかし、なんと言っても、NAVLABプロジェクトが自動運転の研究全体に大きな影響を与えたのは、1995年にNAVLAB5で、「No Hands Across America」といって、アメリカ大陸を自動運転で横断するというデモでした。ピッツバーグから西海岸のサンディエゴまで約5000 kmの98 %以上を、時速100 kmほどで自動運転して行ったのです。今から21年前のことです。このプロジェクトのいわばハイライトで、自動運転の可能性を一般に示すことができたのでした。

NAVLABのごく初期の頃に、一太郎で有名なジャストシステムの浮川和宣社長が訪問されました。ある新聞の記事に氏の思い出話として、私がノロノロと走る車のビデオを見せながら、「『いやそのうちにどんどん速くなって、高速道路でもスイスイ走れるようになるでしょう』と自信たっぷりに話した」と書かれていました。そんな昔にそれほどの自信があったかどうか自分でも覚えていませんが、きっとそんな偉そうなことを言っていたのでしょう。今となっては多少大風呂敷を広げたのも許されたかと思っている次第です。現在では、皆さんご存知のように自動運転は現実の技術となっ

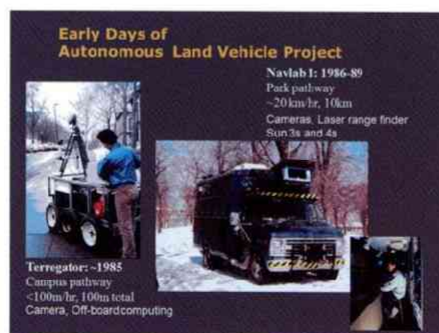


Fig. 18

hour. This was 21 years ago, and it was essentially the highlight of the project, as it made people aware that automated driving was a possibility.

In the very early stages of the Navlab project, we received a visit from Mr. Kazunori Ukikawa, then president of JustSystems, which is best known for the Japanese word processor Ichitaro. I read about his recollections of that visit in a newspaper article, in which he said, "While showing me a video of a very slow-moving vehicle, [Dr. Kanade] spoke with great confidence, saying that it would soon become faster and faster and even cruise on highways." Honestly, I do not remember sounding so confident back then, but I must have talked big. In retrospect, I suppose I was modestly justified in speaking in such grand terms. Automated driving is becoming practical technology, and I am very proud that the Navlab project paved the way toward making it a reality.

Directorship at the Robotics Institute

In 1990, when the Robotics Institute marked the tenth anniversary of its foundation, Professor Reddy, who had served as the founding director since its establishment, came to me and said, "Takeo, you'll be the next director." By that time, the Robotics Institute had become a huge, globally renowned research center with over 150 staff and an annual research budget of over 3 billion yen. Having been in the States for only ten years or so, I was unsure whether I could fill such a big job but Professor Reddy urged me to take on the directorship, saying, "You can do it, you can do it," and so I decided to accept the offer.

After taking up the position, I discovered two contradictory things: it was somehow easier than I had expected, but at the same time it was more difficult than I had thought. In U.S. universities, individual professors and researchers follow their own thinking, conviction, and vision to seek out research funding from national research support organizations or private businesses, and then use the funding to hire people and buy what they need to carry out their research. It's like a conglomerate of independent operators. What this means is that you do not have to act politically or make backroom deals so often. What the director is expected to do is surprisingly simple: to reduce the number of unnecessary procedures and tasks imposed on researchers and keep the organization moving forward quickly and flexibly so that researchers can perform their work efficiently, even eliminating the distractions originating from the director himself. Most importantly, the director must hire "star" professors and researchers who can attract funds for their research, or young researchers who have the potential to become

てきつつありますが、NAVLABプロジェクトによって、それにつながる貢献ができたように思っています。

ロボット研究所の所長になる

ロボット研究所ができて10年になった1990年に創設以来研究所長をやって来られたレディ教授に「タケオ、おまえが所長になれ」と言われました。そのときには、ロボット研究所は既に150人以上の陣容で年間30億円以上の研究費を使う大研究所になっており、世界的な評価を受けていました。そんな研究所の所長をアメリカに来て10年ばかりの自分ができるかと不安でしたが、レディ教授が「できる、できる」と言われるので引き受けることにしました。

就任してわかったのは、意外に簡単であるということと、思った以上に大変であるという二つの相反する事実でした。アメリカの大学は、個々の教授や研究者たちが自分たちの考えと信念とビジョンに従って国の研究支援機関や企業から研究資金をとって来て、それで人を雇い、物を買って研究を進めるいわば個人経営者の集合体のようなどころがあります。だから、政治的に動くとか根回しといったことは意外と多くはありません。所長がすべきことは、簡単と言えば簡単で、そういう活動をしやすいように、もっと言えば邪魔にならないように、彼らにとって無駄な手続きや仕事を減らし、組織を迅速柔軟に動かすことにつきます。そして、一番重要なのはそういうことができるスター的教授や研究者たち、あるいは将来そういうスターになりそうな若手研究者を雇うことなのです。だから、職場を移ってもよいと思っていることがわかった他機関のシニアレベルの研究者を講演に招待してその日のうちにポジションをオファーしたり、MITでポストドクを済ましたばかりの機械系の有望な若い女性研究者を機械学科と交渉して兼任ポジションを作って雇ったりしました。そして、しばらくすると、その研究者が機械学科との兼任が却って自分の自由を束縛しているというので、今度はあっさりこちらの専任に替えました。

一方、難しいのは自分の研究所がそういう良い組織であることを外部に発信でき、注目させられるかということです。所長自身が研究者としてよく見える存在であるべきこと以上に、研究所全体での成功例をうまくパッケージにして、機会を選んでいろいろなところで魅力的に話せる発信力があるかとなります。その点、私のように外国人で必ずしも英語が完璧にしゃべれるわけではない人間には、さまざまな会合での挨

stars in the future. As such, for example, I invited a senior-level researcher from another institute – who I knew was thinking about a move – to give a lecture and offered that researcher a position the same day. I also asked our Mechanical Engineering Department to create a joint position in order to hire a promising young female mechanical engineering researcher who had just completed her post-doctoral research at MIT. When that researcher came to me after a while to say that the joint position rather hampered the flexibility of her research activities, I readily switched her to a full-time position at the Robotics Institute.

Now, the difficult part is promoting your Institute to outside organizations and making them aware of the merits of your Institute. Without this, it is difficult to show the vision of the institution and to attract supporters of the necessary new initiatives. It is important that the director is seen and respected as a researcher but, more importantly, you need to be able to present a package of successful projects being undertaken by the entire organization and talk about them in a charismatic way on various selected occasions. As a foreigner in the U.S., my English, thought not bad, was naturally not perfect, which put me under considerable pressure when I had to make remarks at various meetings. Looking at it another way, however, it might have been an advantage for me, since I had to make a conscious effort to construct my stories and give an enjoyable speech that included jokes. Occasionally, I would be stuck for a word during lectures and speeches and, while trying to explain my ideas with other words, I would often ask the audience for help, saying, “Hmm, what’s the word?” This wasn’t intentional and just came to me naturally, but somehow it appeared to people that these instances were very finely timed, and one of them once told me, “I know you really know the word, but pretend not to. That helps you to establish a dialogue with the audience.” That kind of generosity from people in the U.S. helped me a lot.

I held the position of director until 2000, by which time the Robotics Institute had doubled in size, with 300 staff and an annual research budget of 6 billion yen. In 1995, the Wall Street Journal selected Pittsburgh as one of the hottest high-tech cities. In an article describing how Pittsburgh had earned the moniker “Roboburgh,” the CMU Robotics Institute was mentioned as having made significant contributions.

The only professor to ever appear at the Super Bowl?

“EyeVision” is a system that was used for the broadcast of Super Bowl XXXV, which took place on January 28, 2001. The development of the system itself was a major

扱は大変なプレッシャーでした。が、逆にそれゆえにストーリーを考え、ジョークも入れて楽しく話す準備をする努力を意識的にしなければならないことは却ってアドバンテージだったかもしれません。たとえば、講演やスピーチの中で言葉に詰まったとき、その言いたいことを別の言葉で説明しながら、私が、「ええーっとなんと言うのでしたかね」と聴衆に訊いて助けてもらうことがあります。私としては自然にしているのですが、タイミングが絶妙らしく、アメリカ人に「知っているくせに無理にやっているのだろう。あれであなたの話を聴衆が対話として受け入れるのだ」と指摘されたことがあります。そういうことを受け入れるアメリカのふところの深さは助けになりました。

2000年まで所長を務めましたが、その間にロボット研究所は2倍の規模、300人の陣容と60億円の研究費を使う研究所になりました。また、1995年のウォール・ストリート・ジャーナル紙は注目すべき市としてピッツバーグを「ロボバーグ」と名付けた記事の中にCMUロボット研究所の貢献が大であると書きました。

スーパーボウルに出演した唯一人の教授？

2001年1月28日のアメリカン・フットボールの王座決定戦スーパーボウルXXXVの放送で使われた「アイビジョン」というシステムの開発は一つの出来事でした。2000年の春にCBSから連絡がありました。映画『マトリックス』のように決定的プレーを時間を止めてグルッと回って見せる映像再生システムをフットボールの中継で実現できないかと言います。皆さん、映画『マトリックス』はご覧になりましたか。その撮影には、100台以上のカメラがスタジオの中心を取り巻くように設置されます。その中央で演技する俳優を、ここだという場面で、これらのカメラで同時に写真を撮ります。それらの画像を順番につなげると、映画を見ている人には時間が止まって、その周りを回っているように見えるのです。それと同じことをアメリカン・フットボールでやろうと言うのですが、グラウンドの場合は広いし、どこで良いプレーが起こるかわからないので、あらかじめカメラを特定の場所に向けて用意しておくことはできません。そこで、カメラをパン(左右)、チルト(上下)、フォーカス、ズームという撮影条件などを制御できるロボットカメラに変え、グラウンドを取り囲むようにスタジアムの上部に設置するのです(Fig. 19)。そして、それらをスタジアムの外に置かれたコントロール室から、すべてのカメラが同時にプレーの一点を追っかけるよ

event. In the spring of 2000, CBS contacted me to ask if it was possible to develop a video replay system for their broadcast of this annual American Football Championship game. They wanted a system that would let you stop the frame on a decisive play and revolve the camera viewpoint around it, like freezing a fired bullet in time as in *The Matrix*. Have you all seen the film *The Matrix*? In the film, more than 100 cameras were arranged around the center of the film studio. Actors and actresses performed at the center, and all of the cameras recorded key scenes simultaneously. By putting together those still images in a sequence, the audience would feel as if time stops and they were spinning around the target. They wanted to do the same for the football game. The problem was, it would be possible to preset the spinning point, since the football field is huge and you cannot know where a good play is going to take place. So instead of static cameras, we used robot cameras that could control their panning (left and right), tilting (up and down), focusing, and zooming, and place them up around the top of the stadium to surround the field (Fig. 19). By automatically controlling all of the cameras so that they would follow the same play simultaneously from a control room outside of the stadium, we were able to create *Matrix*-like effects at any point on the field we chose. We set up a system like this by placing 33 cameras on the second level of a stadium in Tampa, Florida. This is a video of the game. I was asked to appear in the broadcast of the Super Bowl to introduce the system, which made me the “only university professor to ever appear at the Super Bowl.” It was estimated that as many as 200 million people watched the game throughout the world, which featured videos taken using EyeVision and my explanations. After that, a similar system came into common use in broadcasting and other attractions, and I believe that you have probably seen this kind of video since then.

Actually, the origin of EyeVision can be found in one of my research projects that began around 1993 (Fig. 20). We built at CMU a hemispherical dome with 51 cameras all facing inside. In the dome, it felt like being inside the compound eye of a dragonfly. Images taken from 51 different angles were processed to create three-dimensional data of each moment and its sequence. I dubbed this four-dimensional digitalization, or virtualized reality. If you “virtualize” the real world and broadcast the data, viewers can freely move around the software-driven virtualized world and observe synthesized images from arbitrary angles and positions that are different from the original camera angle. This was the original idea. My slogan for the system was, “Let’s watch NBA on the court.” If we covered the basketball court with a veil of cameras, we would be able to watch the game as if we were freely moving around the court. When I assembled the

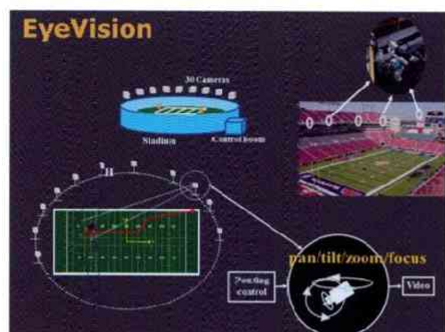


Fig. 19

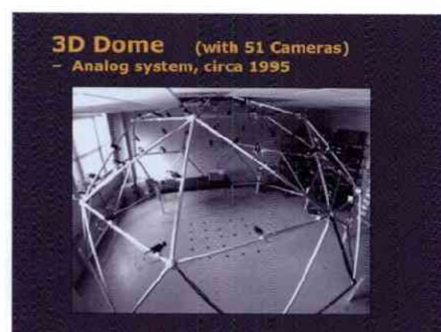


Fig. 20

う自動コントロールすることで、グラウンドの任意の位置で『マトリックス』の状態が起こっているようにします。このようなシステムを実際に、フロリダのタンパというところのスタジアムで、2階席にカメラを33台設置して作りました。その結果がこのビデオです。こうしてスーパーボウル中継放送に出演し、システムを紹介しましたから、私は「スーパーボウルに出演した唯一の大学教授」となりました。「アイビジョン」の映像とその説明のショーは全世界で2億人ほどが見たとされます。その後、「アイビジョン」に似たやり方は、放送やアトラクションで多く使われるようになり、皆さんも見られたことがあると思います。

実は、「アイビジョン」のもととなったのは1993年頃からやっていた研究です (Fig. 20)。CMUには50台以上のビデオカメラが、中を半球状に取り囲むように設置されたドームがありました。ちょうど、トンボの複眼の中にいるような感じです。50の角度から撮影された映像を処理して、各瞬間の3次元の情報とその流れとして取り入れられます。私はこれを4次元デジタル化、あるいは、「バーチャライズド・リアリティ (仮想化現実)」と名付けていました。こうして現実世界を仮想化してしまい、その情報を放送として送れば、視聴者がソフト的に仮想化された世界の中を自由に動き回って、元のカメラアングルにはない角度や位置からの映像も合成し、観察することができるというのがアイデアです。“Let's watch NBA on the court” (NBAのバスケットボールゲームをコートで見よう) がスローガンでした。バスケットボールコートカメラを覆ってやれば、私たちはコートの中に自由にだって入ってゲームを見ることができるはずだというわけです。1995年に最初のシステムを作ったときは、「こんなたくさんのカメラを使うのは現実的ではない」といわれました。しかし、今では、似たような多数カメラシステムがあちこちの研究所や企業で作られ、多くのカメラを使ったメディアは普通になってきました。CMUの現在のシステムは



Fig. 21

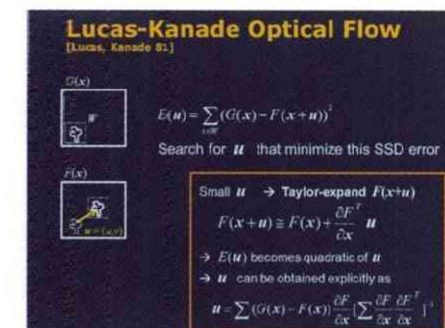


Fig. 22

first system in 1995, some people criticized that it was not very practical to use so many cameras. However, by now many research centers and private businesses have put together similar multi-camera systems, and now such media using many cameras have become common. The current system at CMU is made up of 480 cameras (Fig. 21). I guess one secret to “doing something interesting” is to “go a bit extreme.”

The Lucas-Kanade method – Students are right

One of the most fundamental visual abilities is to track the movements of objects in videos (dynamic images). This is necessary, for example, when following the movements of individuals recorded in surveillance camera video or when a self-driving vehicle determines the moving direction and velocity of another vehicle or pedestrian. More precisely, it is the capability to find vector u , which expresses the movement or displacement of a characteristic pattern between an initial image frame $G(X)$ and the next frame $F(X)$ (Fig. 22). Mathematically speaking, you need to find the u that minimizes the error function $E(u)$, shown here, which is the sum of squared difference of pixel values between $G(X)$ and $F(X+u)$ over a certain window W . Most simply, you could calculate $E(u)$ for every possible value of u – an all-point search – and find the answer of u that generates the smallest E value. The problem of this simplistic method, however, is that no matter how fast computers are, it will cost a considerable amount of time to do this for every single characteristic pattern in an image. The question here is how to find the best u accurately and fast. Actually, I was working on this problem around 1980 with Bruce Lucas, one of my first graduate students whom I mentioned before. One day, Bruce came to me and said that he had come up with a good solution. He said, “The movement of u is generally small. So, we apply the Taylor expansion to the function F and use its first order approximation. By substituting F with

480台のカメラからできています(Fig. 21)。「面白いこと」をする一つの秘訣は「少し極端にやってみる」ということでしょうか。

Lucas-Kanade法—学生の言うことは正しい

視覚能力のなかで最も基本的なものの一つに、ビデオ(動画)中での物体の動きを追跡することがあります。監視カメラで人の動きを追跡するときや、自動運転で他の車や歩行者の移動方向や速度を知ろうとするときに必要です。もう少し正確に言うと、最初の画像フレーム $G(X)$ 中の特徴あるパターンが次のフレーム $F(X)$ の中でどの位置にどれだけ移動したのかを表すベクトル u を求めたいということです(Fig. 22)。数学的には、ここで書いた、誤差の関数 $E(u)$ を最小にする u を求めるということになります。単純には、全点探索と言って、可能な u の値のすべてについて $E(u)$ を計算し、そのうちで最小になる E を与える u を答えとすれば良いのですが、それではいくら計算機が速いといっても、多数の特徴点すべてに対して行うには多大の時間を要します。どうすれば、 u を正しく高速に計算できるかというのが問題なのです。この問題について私は先ほど言った最初にやってきた学生であったブルース・ルーカスという学生と一緒に1980年頃研究をしていました。ある日、ブルースがやって来て、良いことを思いついたと言います。「移動量 u は一般に小さい。だから、関数 F をテイラー展開してその第一項までで近似して、誤差 $E(u)$ の式に代入すると、その最小化は簡単に解けてこの式が得られる。つまり画像の微分を計算するだけで、 u を直接、高速に計算できる。これはすばらしいアイデアだから論文を書こう」と言います。私は、「テイラー展開は300年も前から知られている方法だし、それを使った式の導出も高校生でもできるほとんど自明のことで、別に新しくもなんともない。そんな論文を出せば、私の名前に傷が付いてはいけなくて書いてはならない」と言いました。しかし、彼はそれでも書きたいと主張するので、仕方なく「それなら、なるべく目立たない会議で発表しろ」と言って発表しました。実際、この研究は一度も学会のジャーナル誌に投稿しませんでした。ところが、案に相違して、この会議発表論文は結果的に私が著者として入っている500ほどの論文の中で最も数多く1万回以上参照された、「超」と言っているほど有名な論文になったのです。そして、この論文の方法は、後にLucas-Kanade法と呼ばれるようになり、現在、ビデオの伝送や記憶に使われるMPEGなどの圧縮符号化の中で重要な画素の移動量を求めたり、自動運転の中

its its first order approximation the error function $E(u)$ becomes a second order function of u , whose minimization can be easily solved and to produce this explicit formula. In other words, just by applying differentiation to the images, u can be computed directly and quickly. This is a great idea, and we should write a paper on this.” I replied, “Bruce, the Taylor expansion is a 300-year old method, and the derivation of the formula of the minimum is almost trivial to the level that even a high school student could do it. There is nothing novel about it. If we publish a paper on it, it could ruin my reputation. I don’t want that to happen. You shouldn’t write it.” Since he kept on insisting, however, I said to him, “All right, you can publish, but only at an obscure conference.” And, in fact, we did not bother to contribute that research paper to any academic society’s journal. Contrary to my expectations, the conference paper ended up becoming a “super” popular paper, which has been cited over 10,000 times and became the most referenced among some 500 papers to which my name is associated. The method later became known as the Lucas-Kanade method, and is currently used to compute the image movements or optical flow. Its application ranges from compression coding technologies like MPEG, for efficiently transmitting and storing video data, to tracking feature points, in the simultaneous localization and mapping (SLAM) method for calculating one’s own movements from images for automated driving. The method is also used for image processing in the fields of biology and medical science. It has gone on to become one of the most extensively used vision algorithms. If Bruce had followed my instructions, we would not have the Lucas-Kanade method, or even my current reputation. Since then, I make it a rule to say to my students, “In most cases, what a professor says is wrong.” Papers, or research for that matter, are valuable not because they offer superficial newness but because they solve certain real problems and offer some advantages. And, those points need to be strictly assessed. I certainly dramatized this story a bit to make it more amusing, but this is a good example that indicates the important roles that graduate students play in research projects at universities. I have several other examples in which I worked with extremely talented students, and together produced methods and theories bearing their names as well as mine. Like the case of the Navlab system, project accomplishments tend to be attributed to principal investigators, but the actual work of developing a system is performed by many staff members and students, and no project would be successful without their brilliant efforts. In this sense, I have been more than fortunate to have had consistent support from such great staff members and students for as long as 35 years. I would not even be here today to share this story if it weren’t for them. I cannot thank each one of them

で移動物体の追跡や自身の動きを画像から計算するSLAMといわれる方法の中で必要な特徴点を追跡したりすることに使われています。また、生物学や医学の画像処理でも使われています。ビジョンのアルゴリズムの中でも最も広く使われるものの一つになったのです。もし、ブルース・ルーカスが私の言うことを素直に聞いておれば、今日Lucas-Kanade法どころか、私自身の今日の評価もなかったかもしれません。以来、私は「教授の言うことは、多くの場合間違っている」と、学生に言うことにしています。論文、つまり研究の価値は表面的な新しさにあるのではなく、何の問題が解け、どういう利点があるかであって、それらをしっかり見極めなければならないということなのでしょう。この話は面白くするために多少脚色が含まれていますが、大学の研究において大学院生の役割が如何に大きいかの例と言えるでしょう。ほかにも、私は極めて優れた学生のおかげでできた仕事で彼らと私の名前が並んで付けられた方法があります。また、NAVLABの例のようなシステムの成果は大枠としてどうしても研究代表者に帰されがちですが、システム開発は多くのスタッフや学生が実際に行っており、彼らのすばらしい仕事なしには絶対に成功するものではありません。これらの意味で、私は本当に優秀なスタッフや学生のサポートを35年間にわたって受けることの出来た幸運の持ち主だったといえるでしょう。今日ここでこんなお話が出来るのも、明らかに彼らのおかげで一人ひとりの名前を挙げることはできませんが、ここから感謝いたします。

楽しく、役に立つ研究を

すばらしい研究をした人と話してみても、いつも感じることは、彼らは楽しく研究の話をすることです。「よくあるこういう問題を考えてみよう」と身近にある問題を例にとって話し始めます。そして、「この問題はやさしそうだけれど結構難しい。まずこう考えられるけれど、それだとこんな風に変なことになってしまう。そこで私はこう考えた。すると、こんな面白い解法が見つかった」と話が進みます。そして、「それによって、この問題も解ける、この問題も。しかも、それらからこんな新しい応用が生まれるのではと思われる」などと言います。そんな人の話を聞いていると、話を一方的に聞いているというよりは、聞き手である自分もつい引き込まれて、その問題の謎解きを話者と一緒に考えようとしているのに気がつきます。その研究ストーリーにある、何がどうしてどうなってという「物語性」に惹かれるのです。研究の物語性

individually here, today, for their dedication.

Fun in Research and its Usefulness

Whenever I speak with researchers who have done brilliant work, I always feel that they enjoy talking about their research. They typically start a story with, “Let’s think about this common or familiar problem.” Then they go on and say “This problem may look easy to solve, but it’s really quite difficult. You can think about it like this at first, but the solution you’d end up with would be really odd. So, I thought about it like this to discover this interesting point and solution.” “In this way, we can solve this problem and that problem, and we could have new applications like this and that.” Listening to such people, we do not hear them speak one-sidedly, but find ourselves being drawn into their stories before we know it and trying to solve the question together with the speaker. We are drawn into the story because of its narrativity, that is, what happens and how. Technically speaking, the story-like nature of research involves clear ideas about the goals, setting methods, methodologies, and expandability. I personally believe that at the very root of all this attraction is a feeling of joy – you get excited about the prospect of your research, that is, how your research will turn out and how it will become useful for society. Now, I may need to add some comments on the concept of “useful.” When I emphasize the point of “usefulness,” one might say, “So, you mean applied research, right?” Others might say, “You shouldn’t evaluate basic research from what use it offers or from a short-term usefulness perspective,” and some extremists might say, “Basic research is doing something without any idea of how it will turn out.” I wouldn’t say that these comments are completely off the mark, but I believe that such arguments are mistaken in most cases from either or both of the following two perspectives.

Firstly, we shouldn’t interpret the term “usefulness” only in the narrow sense of the word – in other words, “can be put to use tomorrow,” “can be applied to commercial product development,” or “can make money.” “Useful” research refers to research that has given birth to a systematic method for something that previously had no solution, or that it reveals something that has remained hitherto unknown. Reading the true story of how DNA’s double helix structure was identified, which is one of the most important basic research findings of the past century, we can see that those who were involved were completely aware that the discovery would lead to a fundamental elucidation of hereditary phenomena. Put simply, research is “useful” when researchers have a clear

とは、技術的にはその目的、設定法、方法論、その発展性などについての明確な考えでしょう。しかし、私はそのもととなるのは、その研究の行き着く先としてどのように社会に役に立っていくかを思い浮かべて楽しくなる、そんな気持ちではないかと思うのです。この「役に立つ」という言う点に少し補足が必要かもしれません。私が「役に立つ」と言う点を強調すると、「それは応用研究のことですね」とか、「基礎研究は何の役に立つかといった短期的な評価をするべきでない」とか、もっと極端には「基礎研究はどんなことになるかわからないことをするものだ」と言う人がいます。私は、そのような場合がまったくないとは言いませんが、その議論は次の二つの観点の両方あるいはどちらかから、大体は誤っていると思っています。

一つは、「役に立つ」という言葉を「明日にも使われる」、「明日にも製品を作る」とか「お金になる」とかいう狭い意味だけにとるべきではないということです。「役に立つ」とは、その研究によって、今までやり方がなかったことに組織的方法が生まれるとか、今までわからなかった何かが明らかになるといったことも含まれるのです。世紀の基礎研究であるDNAの2重螺旋構造の特定に到る実話を読んでも、それに携わった人たちがそれが遺伝現象の根本的解明につながるということを完全に意識していたことがわかります。つまり、「役に立つ」とは研究における問題意識なのです。その意味において基礎研究ほど役に立つ研究はありません。

もう一つは、革新的技術や現象の発見のほとんどは極めて具体的な要求を持っていた人々から生まれたという事実です。トランジスタはかさ高く消費電力が大きく、寿命の短い真空管を置き換える要求から生まれました。インターネットのもととなった計算機ネットワークはミサイル攻撃から軍事用コンピューターシステムを守ることがもともとの動機であったとされています。計算機ネットワークは単に技術のつながり合わせといわれるかもしれませんが、そうではありません。パケット通信とか多層通信プロトコールによる通信のバーチャル化といった極めて基礎的な方法の発明が必要だったのです。人工知能の元祖の一人でノーベル経済学賞受賞者でもあるCMUのハーバート・サイモン教授はこのような現象を「The Sciences of the Artificial (人工物の科学)」と呼んでいます。人工物にも、自然界と同じように科学があつてしかなるべきというのです。それから、私は研究を楽しくまた面白く感じるべきものだと思います。科学や先端技術研究の面白さはどこから来るのでしょうか。私は、研究は「解くべき問題や自然を相手に知的ゲームをやっているのだ」とイメージしています。一般にゲームやスポーツでは相手を正面から突破する手、隙を突く手などさまざ

awareness of the meaning, future impact, and potentials of the issues at hand. In this sense, no other form of research is more “useful” than genuine basic research, and activities without any such awareness is not research, let alone basic.

Secondly, it is a fact that most innovative technologies and discovery of phenomena were created by people who had very concrete demands. Transistors were invented for replacing vacuum tubes, which are bulky, consume much power, and have a short lifetime. The computer network that later developed into the Internet was originally created to protect a military computer system from missile attacks. You might think a computer network is developed by merely linking together different technologies, but this is not the case. It required the invention of fundamentally basic ideas, such as virtual communication, packet switching, and multilayered communication protocols. One of the progenitors of artificial intelligence and a Nobel Prize laureate in Economic Sciences, Professor Herbert Simon of CMU calls such phenomena the “sciences of the artificial.” He claims that there should be sciences for the artificial things just as the traditional sciences for the nature. I mentioned earlier that we should enjoy and have fun with our research. So, where does the fun of science and advanced technological research come from? I imagine that doing research is like “playing an intelligent game with problems to be solved or nature herself.” Generally speaking, when you play games or sports, you compete with your opponents by using all sorts of means, sometimes attacking them by a frontal offensive move and other times taking them off guard. This process between us and them is what makes it “fun.” The same goes for research. When you work on something as a researcher, you try out your ideas in a bid to solve a problem. Then, the natural law or logic of nature, which governs the problem or guards it from being solved, responds: “Hold on, that’s not an essential solution. Things won’t work that way. I’m going this way to escape.” Then you say, “I see. Then, how about this?” In this way, you drive the problem into a corner. It’s fun to keep taking steps that will make your opponent – in this case, a research problem – eventually concede “Hmm, that’s a good move. It gets to the truth (essence). You won.” So, what about the outcome of the game? Undoubtedly, for a game, the primary goal is to win. Athletes and professional chess players are hailed as being top-notch or masters because they win. Likewise, we researchers need to have a competitive spirit with which to win the intellectual game against a problem, that is, to obtain an essential solution to a problem. This requires education, training, and constant study. Once you do those sufficiently, you will come to take natural pose to these intellectual games. If you do so, you’ll feel at ease, happily enjoy doing your research, and increase your winning

まな勝負手を放って相手と対峙します。そのプロセスが「面白い」のです。研究でも、研究者である自分は問題を解こうとして、「こんな考え方はどうだ」と手を指す。するとその問題を仕切っている、あるいは解かすまいとガードしている自然の摂理というか論理が、「どっこい、それでは本質的な解ではない。こういう場合はだめだろう。だから、私はこう逃げるよ」と言う。「なるほど、これならどうだ」とさらに問題を追い詰める。その過程で、相手の問題が「ウム。それは良い手だ、真理(本質)についている」と痛がってくれる手を放つことが面白いのです。では勝ち負けはどうか。ゲームだから勝つことが一義的目標であることは疑う余地はありません。スポーツ選手や棋士も勝つから、一流や名人と言われるのです。だからわれわれ研究者も同様に、問題との知的ゲームに勝つ、つまり本質的に解くのだという競争心を持ち、それを支える練習を常にする必要があります。それが教育であり、訓練であり、常の勉強なのです。しかし、それらを十分した後は、知的ゲームを楽しむのだという自然体が気持ちを楽にし、研究を面白く楽しくでき、そしてそれが勝率を上げるのだと思うのです。

「世界には、どうしてなのだろう？こんなことができたらいいのに・・・などという問題がたくさんある」。アレン・ニューウェル教授は学生たちに熱っぽく語るのが常でした。「そういう問題の一つひとつが『解いてくれ、解明してくれ』と、われわれ研究者を待っているのだ」と。私たちは、そういう問題を仕切っている摂理と勝負するのだという自然体で、楽しく研究に臨めば良い研究が生まれるのではないのでしょうか。そのプロセスの中で、問題の意味を抽象化できたときは、より一般的で高次の概念に到達するし、そうでないときは、特定の問題の特別解を得るでしょう。そのどちらが起こるかは、能力にもよるし、運にもよります。それはどちらでも良いのではないのでしょうか―役に立つというインパクトさえあれば。

percentage.

“In this world, there are so many problems that make you wonder why things are the way they are, or those that make you want to do this and that,” I remember, Professor Allen Newell spoke passionately to his students. “Each of these problems is eagerly waiting a researcher to come and ‘solve or elucidate’ them.”. If we face with those research problems in a calm, relaxed manner, as if we are playing a game with the natural law that governs them, and if we enjoy doing the research, we should end up with a good outcome. If we can abstract the fundamental meaning of the problem that we are working on, we should then reach a more universal and higher-order concept. And if not, we should at least arrive at the particular solution of that specific problem. Whichever happens probably depends on your ability and luck. However, either way is a positive outcome – so long as it makes an impact of being useful.

稲盛財団2016——第32回京都賞と助成金

発行 2017年8月31日

制作 公益財団法人 稲盛財団

〒600-8411 京都市下京区烏丸通四条下ル水銀屋町620番地

Phone: 075-353-7272 Fax: 075-353-7270

E-mail press@inamori-for.jp URL <http://www.inamori-f.or.jp/>

ISBN978-4-900663-32-9 C0000