

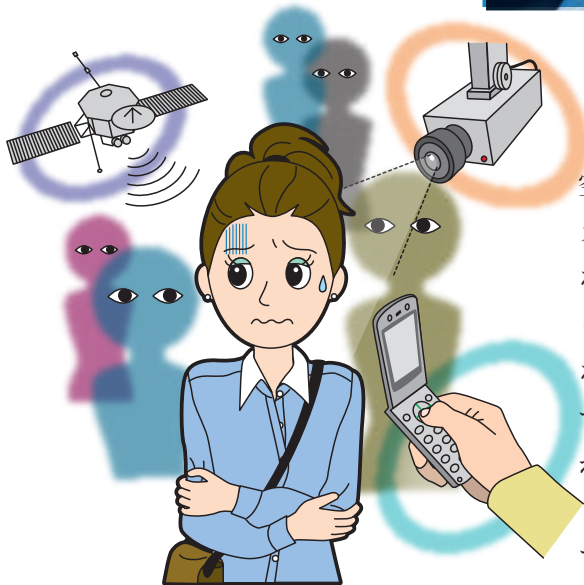
●信号処理は電気系の基礎的な科目としてとても重要視されています。馬場口登先生の研究室では「マルチメディア」…つまり音声・画像・映像に対する様々な信号処理技術の研究がされています。



馬場口 登 先生

人と人との関係に応じて見せる画像を変える

～ 誰に、どのくらい見せるか～は自分で決めたい～



グーグル・ストリートビューや公共空間の監視カメラ、街中の人々が持つカメラ付き携帯。便利な世の中になった一方、思いがけず被写体となりその映像がひとり歩きする…そんな危険性の中を私たちは生きています。このような現代社会の怖さ。それは「自分の情報」さえ自分でコントロールできていないことにあるのではないのでしょうか。画像を“撮られる人と見る人の関係によって、見せ方を変える”馬場口研究室ではそんな技術の研究に取り組んでいます。「隠したい」はずのプライバシー情報も、相手や状況によっては「見せたい」となりうる…という側面があることに着目した研究。プライバシー保護処理の現場をのぞいてみましょう。

「隠したい」はずのプライバシー情報も、相手や状況によっては「見せたい」となりうる…という側面があることに着目した研究。プライバシー保護処理の現場をのぞいてみましょう。



画像処理とは具体的にどのようなことですか？

馬

画像処理にはいろいろあります。ここ 15 年ほどは、画像を見やすくするとか、数時間にわたる映像の中のいい場面を自動的に探す研究に取り組んでいます。朝から家に帰るまでには、一日を通していろんなことがあります。かりにその様子を撮影して、それを自動的に編集できる。それが究極の目的です。中でも現在取り組んでいる研究のひとつに「プライバシー保護処理」があります。名前のように個人が特定できるものをプライバシー情報とよびます。日本にはありませんが、個人にたどり着けるソーシャルセキュリティナンバーのようなものは、すべからずプライバシー情報です。例えば、データベース情報の中にも個人にたどり着ける情報はあります。画像情報の中にも、センシティブなプライバシー情報があります。一つはやはり顔です。顔をみたら誰かわかる。そして名札。文字が読めれば名札を見て誰かということが識別される。今の技術だと、顔がわかり、誰かが分かれば、データベースまでたどり、特定の人の年収・職業・家族までわかる可能性があります。プライバシー面で厄介なことは、その他にもいろいろあります。たまたま道で撮影した写真に家や表札が写っていることがあります。なにげなく見た映像でも、周囲に映る文字から、一定の個人情報とつながり、その家の持ち主や家の中のことまでわかってしまうのです。このように撮影画像には個人のプライバシーにたどり着ける情報がいっぱい散らばっているのです。そ



の代表的なものが顔や姿形です。それに対して、見せたい人には見せてもいいし、見せたくない人にはちゃんと隠す処理が大事になってきます。



具体的にはどのような技術をつかいますか？

馬

基本的に画像処理の技術を使っています。ところで「プライバシーの感覚」は、人それぞれですよ。私がどこかに映った時、そんなのは誰に見せてもいいと思いますが、あなたはどうか。



私はいやです。本当にいや。

馬

ではその「いや」というのは、全員に対していやなのか…場合によっては見せてもいい時もありますよね。例えば、幼稚園にいるわが子を遠くから見る時に、自分の子どもの顔は人には見せたくないけど、親は見たいと思いますよね。撮影した画像には自分の子どもを含め他人もそこに映っています。親が見る時には、自分の子どもの顔はしっかり見えても、よその子はぼやっとぼける。また、別の保護者には自分の子どもはよく見え

ないようにするのです。集合写真でも、見る人によって、表示される画像の状況を変える。たとえば図1のように「画面A」と「画面B」で、被写体の画像を変えるわけです。



図1

そのためには、画像の中に誰が映りそして誰が見るのかを判断しなければなりません。画像に映っている人が誰で、見る人に対しどのような画像処理をすべきかを決めるわけです。その処理の方法はさまざまです。何も処理せず、実写をそのまま映すというのが一つ。ほかにも抽象度によってシースルー、モノトーン、モザイク、シルエット、バー、ドット、透明などの映し方があります。(図2参照)

この「透明」とは、本当はそこに人がいてもあたかも存在しないように隠してしまうのです。透明人間のようににも見せません。また、グーグル・ストリート

ビューでは「ぼかし」が使われています。顔あるいは体に、モザイクをかけてわからなくする。エッジと呼ばれる輪郭だけで表示する方法もあります。棒や点で表す方法もあります。棒で表すとどれくらいの人数の人がいるかがわかるのです。このような画像処理を自動的に行う研究をしています。

アリス：このように「これは出す、こっちは出さない」というのは、自動的に判断されるのですか。それとも法則を覚えさせておくのですか。

馬：ここで最初に開発したシステムは、会社や学校、マンションのようにメンバーが特定されるようなところで使います。特定メンバーを対象にする場合、今、画像に映っているのは誰かということは自動的に判断できます。この被写体を誰に対して、どのくらい見せるかということをあらかじめ決めておきます。例えば、私の家族にはすべて見せるが「知りあいのあの人には服は見せないが、私の顔なら見せてもいいわ」とか、全然知らない人には何も見せない、とかを決めるので

す。そのためにはメンバーを決めておくことが大事なのです。あらかじめ誰に対してはどこまで表示するかを決めておけば、被写体によってコンピュータが自動的に判断して画像処理を行うのです。このとき、固定カメラで撮影していると、背景の映像は分かっています。



先に背景だけ撮っておくということですね。

馬：そう。監視カメラなどで四六時中固定カメラで撮影していて背景がわかっているならば、ここに人か犬かわけのわからない変なものが入ってきても、背景との差から画像処理して取り出すことができます。つまり、背景の上にオブジェクト1、オブジェクト2が重なるといった画像生成モデルをベースに画像の処理をします。そして、この現れたオブジェクトの領域だけ切り取れば、先ほどのような抽象化の処理は比較的簡単です。透明にしようと思ったら、これを抜いちゃえばいいわけです。

アリス：そうすると、この被写体の背景は静止していなければならないのでしょうか？

馬：確かにひらひらと動く背景ではこの種の画像処理は難しいのですが、長時間観測すると背景の統計的なモデルが得られます。木立が風で揺れていても、四六時中この画像を撮っておくと、統計的にここあたりの葉っぱほどの程度揺れるかがわかります。いっぽう、人物が現れて画像が大きく変われば、画像処理によって高い精度で、それが判断できるのです。

アリス：背景の上にオブジェクト1、2と載せていくという考え方はアニメーションの作成過程と似ていますね？



アニメーションの場合はすべて生成された画像だけで構成されていますが、実際の世界ではそうはいかないことが多いのです。昨日まで何もなかったところにオブジェクトを放置されると、それは背景になってしまいます。それも背景のモデルになっていくわけです。画像処理では精度よく背景と前景を分けることが重要になってきます。このような画像処理や画像認識の方法については、まだ完璧なものはありません。次に抽出した被写体が誰であるかを判断することが重要になります。背景の前に現われた物体がAさんかBさんかそれを判断する必要があります。その方法として、①顔を認識して個人を特定する②動き、歩き方などを認識する…というやり方があります。実はこの辺りの認識は、前景と背景に分ける技術よりはるかに難しいのです。例えばデジカメは、撮影時に顔のところにピッと枠が出てくるでしょう。あの“顔”認識は、フレームの中である程度顔が大きいと上手くできるのですが、監視カメラのように上の方から撮影したり、映った顔が小さいと、認識できないことが多いのです。そこで我々は、顔画像や動きで個人を認識する確度を高めるため、RF-ID タグ（ID情報を埋め込んだタグから、電磁界や電波などを用いた近距離の無線通信によって情報をやりとりするもの）と画像処理を併用して、画面に映っている人を特定する研究に取り組んでいます。RF-ID タグを併用すると誰がその

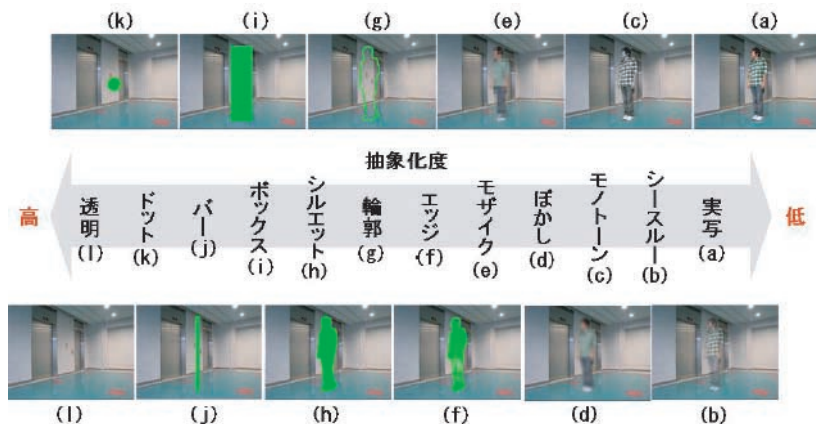


図2

エリア内にいるかという情報はわかりません。それを組み入れた我々のシステムを PriSurv と呼んでいます。

馬 このシステムでは“誰にどのくらい画像を見せるか”をどのように決めているのですか。

馬 開示レベルは、映っている被写体に対し、見る人がどのくらい親密感や見守り義務感を持っているか…に大きく影響されます。そこで、私たちは各種の映像を見せ、それぞれにどのような表示画像を設定したいかをアンケート調査しました。もちろん親しい人から見ず知らずの人まで、親密度に応じた数レベルに分けて実施しました。例えば、見守ってもらいたい保護者には実画像を見せるが、見ず知らずの人に対してはどの程度の画像表示が良いか…といった調査を行いました。その結果、各人のプライバシーポリシーと状況に応じ、画像表示を変える方法をとりました。

アリス：ユーザーの人たちの関係によって決めたルールを設定するわけですね。

馬 そうです、あらゆる年齢層を含む100人にアンケートして、設定すべき適切なルールを決めました。



図3

馬 このシステムは公共の場でも使われるのでしょうか。

馬 公共空間は想定していません。個人個人の関係を明確にすることが難しい公共空間では、デジタルジオラマというプライバシー保護の別のやり方をご提案しています。公共空間では撮影された個人情報本人の了承なく流通するかも知れません。だから強めのプライバシー保護処理をして人を“棒”で表示します。(図3参照)さらに、家族が画像を見る時には“棒”の色を変えて家族の居場所を特定したり“棒”の代わりに実画像にするなど様々なやり方があります。

あるモールでデジタルジオラマの実証実験をやりました。公共空間のモールには誰が来るかわかりませんし、対象のメンバーとして認識できるのは家族くらいしかいません。そこでRF-IDを所有する人はその家族メンバーというグルーピングをして、それによって表示画像を変える方法をとりました。

アリス：事前にRF-IDタグなどで家族を登録しておけば、人が大勢いても「あ、この人は仲間だ」と、色が変わって位置が把握できるのですね。

馬 そう、これはそういうイメージです。だから公共の間では、デジタルジオラマの技術を使います。いっぽう、構成メンバーがだいたい決まっていて、自分を誰になら見せてもよいか…がある程度決まっている場合は、前述のPriSurvの技術を使います。柔軟にプライバシーポリシーを設定できるPriSurv技術は、表示のコントロールが簡単になるという特徴があります。

アリス：カメラが固定されていない場合はどのようにするのでしょうか。

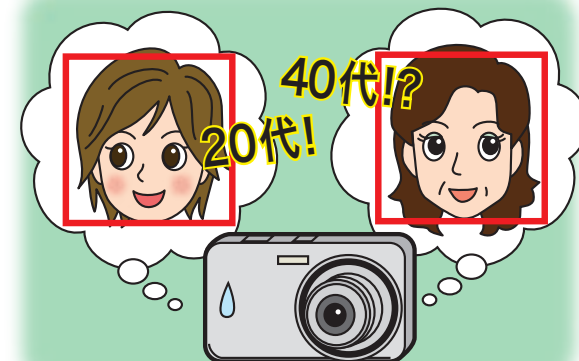
馬 カメラが動くとき背景も動くので、取り扱いがとても難しい。その点、定点カメラで背景のモデルが精密に定義できるという、PriSurv技術とは大きく違います。これはまだ研究の途上です。

馬 モバイルの場合は、被写体と背景との区別が難しいですね。

馬 そうですね。だからこの場合は、カメラを動かしてターゲットの物体を追跡することになります。その意味では、被写体を中心とした画像処理の考え方です。逆に、PriSurvは背景を中心とした画像処理なのです。

アリス：前者のPriSurvは、背景の前に何か知らないものが入ってきたという考え方。いっぽう、モバイルの方は被写体を中心に置く考え方なんですね。ところで、先日集合写真を撮影したとき、たくさん顔が並ぶデジカメの画面の中で、カメラの持ち主の顔だけを枠付けする機能があり驚きました。

馬 あれは、カメラが「私」の写真をいくつもいくつも貯めて覚えているんです。今までの写真では、「顔」を認識する技術が中心でしたが、最近のカメラでは、映ってる「顔」が誰であるかを同定する機能まで入っています。例えばあなたの顔を何度も写真に撮り、「あなたの顔」の画像をカメラにどんどん貯めていきます。そこで、集合写真のように



顔がたくさん映っていても「あなたの顔」はすでにカメラが認識しているので、どの顔が一番近いかを判定して、「これはあなたの顔」と決めることか出来るのです。その辺まで今の画像認識の技術は進んでいます。ほかにも、一番いい顔の時のシャッターチャンスを見逃さず撮影する研究も進んでいます。また、性別の判断も顔からできるのです。年齢に関しても「20代前半か後半か」くらいまでは区別できます。

アリス：事前に登録していなくても、わかるのですか。

馬 カメラはあらかじめ「20代の女性の顔の特徴」をたくさん持っていて、それをもとに年齢を推定することができるのです。この技術は、ある程度実用段階に達しています。タバコの自販機の「タスポ」もそうですね。顔で年齢を推定するので、顔をしかめたら未成年の子も年をごまかせるのです。しわが映ると年齢を重ねて見えるというデータによって判断しているわけです。このように、顔に関する情報処理はかなりのレベルまで進んでいます。

アリス：そのしくみは、よく似た姉妹でもきちんと認識できますか。

馬 これはどうなるかわかりませんが、双子のようにかなり似た子を写真から判別するのは難しいのです。ただ、今の技術では、似た顔をたくさん撮って、トレーニングを繰り返すことによって、顔を判別する学習法がかなり進んでいるので、可能になるかもしれません。その辺の研究は、オムロンなどで行われており、それに関するソフトのシェアも全世界でかなりの部分になっています。かれらは20代、30代、40代の顔の膨大なデータを持っていて、顔認識のノウハウをきちんと持っているのです。

馬 国際化すると、日本人の顔のモデルだけでは対応できないのでは？

馬 20代と言っても人種が異なるとお互い年齢がわからないことも多いのです。そのため、20代の日本人の顔をトレーニングすると「日本人の20代」

が認識できるだけです。認識の範囲を広げたいときにはグローバルなデータを多く集めてトレーニングすることが大事です。

もともと「顔を認識する」ことは難しいとされてきましたが、実は文字の認識よりも簡単だと最近分かってきました。文字はその時々によって変化するのでバリエーションが多いのに対し、人間の顔は、目の上にまゆげ、下に鼻、口があるように、そのバリエーションは限られています。顔は意外と構造的だったということです。

アリス：画像に情報を埋め込んで発信元を特定できると聞きました。

馬 もとの画像データに、情報ハイドニングとかウォーターマーキングという技術を使って、情報を埋め込めばそれができます。画像の中に情報を埋め込んでネットワークに流すと、暗号の鍵を持っている受け手だけが埋め込み情報を取り出せるのです。埋め込みには画像以

外にも「誰が好き」という文字情報でも可能です。テロの際の指示にこの技術が使われた、というふしがあります。ウォーターマーキングとは、お札にも使われる「透かし」の意味です。デジタル情報にウォーターマーキング(付加的な情報)を入れると何かちがう効用を持たせることができます。お札の場合、「本物」であることを保証する意味がありますね。そのように、データの所有者や、著作権などをウォーターマーキングの形で埋め込んで、改ざんや悪用を防ぐ技術です。例をあげますと、封切り前の映画を盗み撮りし海賊版として販売されると映画会社は大きな損害を受けます。私たちは情報ハイドニング技術を使って、音声に埋め込んだウォーターマークから、どの映画館のどの位置で盗撮したかを判定する技術も開発しました(図4参照)。マルチメディアの信号処理に取り組んでいる我々の研究室では、様々なことをやっています。

アリス：“お札”や“映画”…身近な場所や生活シーンで役立っているんですね！ところで、映る側と見る側の関係によって、表示する情報を変えることは、以前から必要だと感じていました。ブログやSNS、ツイッターで何か発言する時でも、この情報を誰が見るのか、それによってどこまで書きたいかが変わると思っています。

馬 まさに、私たちはそれを映像・画像情報に対してやったのです。自分の書いた文章についても同様であるべきだと思います。友達が見るならここまで書くけど、一般の人が見るなら“ピー”

と鳴るようなものです。

馬 相手との親密度によって、“ピー”の程度も変わってくる。

馬 まさにそのとおり。それを画像でやったのです。親でも知らせたいものと知らせたくないものがある。プライバシーというものは、そのように相手と時間と場所によるわけです。情報処理の分野の中で、このプライバシー情報、パーソナル情報の問題はかなりホットな研究テーマです。

馬 これだけ多くの人が携帯電話を持つようになると、そのカメラで誰もが思いがけず被写体となり、その情報が自分の知らないうちにネットに拡がる可能性がありますね。

馬 その安全策ですね。ただ、その反面、自分のプライバシー情報でも開示したい場合も実はあるんです。自分の日々の健康状態はものすごいプライバシー情報です。本来、自分の血圧とか体温とか生理的な状態を公開するのはイヤなはずですが、それをお医者さんには見てもらい、適切な診断・診療をすばやくしてもらいたい。それは情報開示によるメリットです。

今まではプライバシー情報は何でも隠してきたけれど、それも見せる対象によってはメリットもあります。この考え方で、人に応じて表示を変えるシステムにしているのです。例えば、福祉施設や保育園などでは、子どもの画像情報を親には開示してもいいわけです。

アリス：本人が情報を見せる対象者を選ぶのは、すごくいいと思いました。

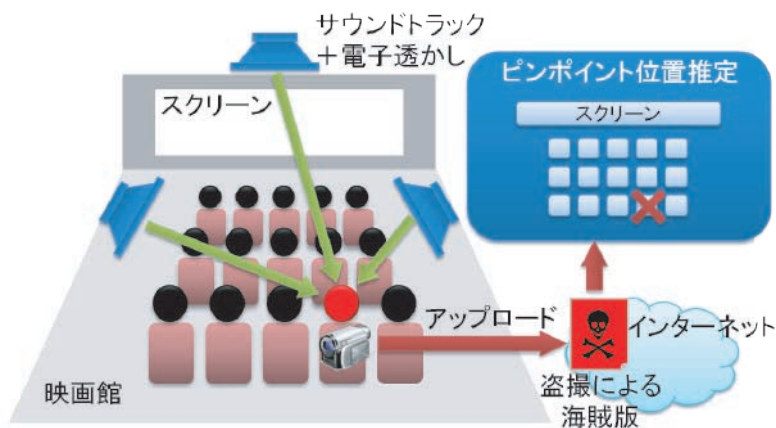


図4

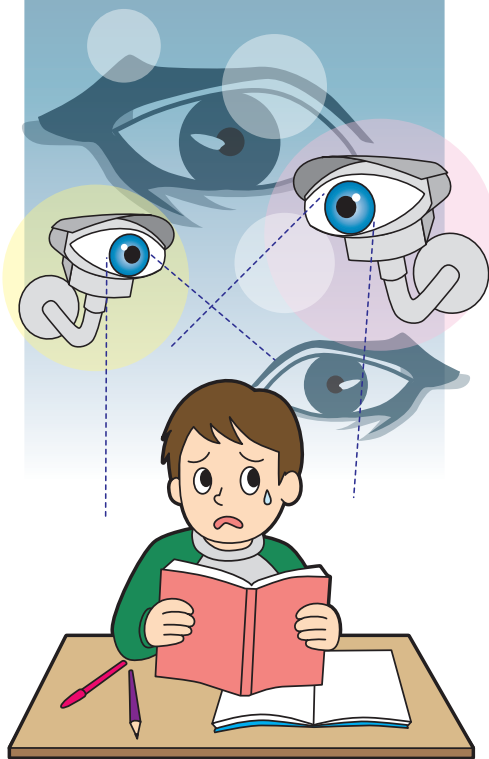
馬 以前は「プライバシー」はまさにテリトリー（個人のいる空間）の話でした。しかし、現代の法制度上のプライバシー情報とは、「自己情報のコントロール権」と考えられています。顔、健康などの生理情報、あるいは年収などの社会的な情報まで、全て自分で開示範囲をコントロールできるという権利が、このプライバシー情報だと言われています。私のつくったこの Prisure は、見ず知らずの人にはぼかして見せる、親密な人には鮮明に見せるというように、被写体としての自分の見せ方を、人間関係によりコントロールできるという意味で、現代の「プライバシー情報」を具現化したシステムになっています。

馬 それはとても重要ですね。プライバシーが保護されないことの怖さは、自分の情報を自分でコントロールできない点にあると思うからです。

馬 メンバー構成が特定される場所では、それがある程度可能ですが、これを公共空間に広げることは難しい。

アリス：物騒な事件が起こると街中にひとつでも多くカメラを付けてほしい、エレベーターでもカメラがついているだけでとホッとします。その一方で、それはそれで怖い。今回の先生の話を知っていると、フーコーの『監獄の誕生』を思い出しました。先生のこの図はまさに“公的な機関（警察）が監視する社会”そのものにみえます。

馬 アメリカではジョージ・オーウェルの Big Brother というロボットが監視する SF 作品があって、そのような社



会になることへの反発から、欧米では監視カメラについて非常に否定的でした。特にフランスでは、監視カメラを公共の場につけるなんてものほかという考えがありました。しかし、国によって監視カメラに対する反応が違います。同じ欧米でも英国は、もともと IRA 爆弾騒ぎなどがあったために、監視社会を認める雰囲気があります。たとえば千里中央から梅田まで電車に乗った時、その区間で顔写真を何枚撮られるか知っていますか？イギリスなら 200 回は撮られますよ。これは恐ろしい話ですが、カメラで撮られていなくても、実は携帯電話では個人の位置情報が吸い上げられています。その情報から、昼間と午後の人口の移り変わりを地図上に簡単にマッピングすることもできるのです。

馬 以前、携帯の中に猫を飼っていた知人が、「猫のセリフが場所によって京都弁になったり大阪弁になったりする」と言っていました。GPS 機能で知人の位置情報を取得していたのですね。

馬 もうそういう時代ですよ。携帯に細かい情報を取得する機能を追加すると、持ち主がどっちを向いているか…までわかります。これも使い道次第では便利なのです。例えば、食堂の混み具合がここから見えたり、大阪市内への道が混んでいるかどうか分かったら、とても便利ですね。このように携帯情報には正面と負の面があるのです。安全・安心社会への関心から、この種の情報もセキュリティ対策として期待されますが、それが強調され過ぎると犯罪対策のようで、せっかくの技術も夢がないように感じられ

ます。もちろん私たちの技術も監視社会を作るために開発したものではありません。情報開示による効用や、家族間の遠隔コミュニケーションツールとして便利です。技術の明るい側面にも期待したいものです。

近い将来、個人情報をも自分でコントロールする日が来るのを楽しみにしたいと思います！
馬場口先生、ありがとうございました。



先生おしえて！

馬場口先生が研究者になったきっかけは何ですか？

小学生のときにゲルマニウムラジオに衝撃を受け、電気に興味を持ちました。あまり研究者志向きではなかったのですが、大学の研究室に入って、知らない間にどんどん研究の自由さ、奥深さにはまりました。

仕事で研究しているという意識が、あまりありませんね。

馬場口 登（ばばぐちのぼる）…1979 年大阪大学工学部通信工学科卒業、1981 年同大学大学院前期課程修了、愛媛大学工学部、大阪大学工学部、産業科学研究所を経て、現在、大阪大学大学院工学研究科教授。1996-97 年カリフォルニア大学サンディエゴ校文部省在外研究員。工学博士。

