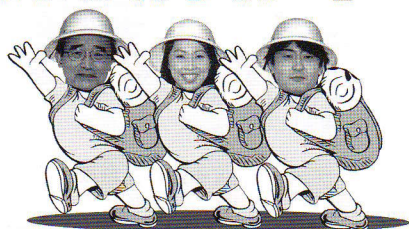


連載 **ビジョン探検隊がゆく**

第4回 顔認識高速大規模処理システム

—光学素子を使ったサーチエンジン—



映像情報インダストリアル・ビジョン探検隊

前号までは、顔画像の抽出と認識について、それぞれの過程とアルゴリズムについて探検してきた。今回は、画像入力から結果出力まで、顔認識全体のシステムについて探検してみよう。

1 女子大学にいざ出発

「隊員の諸君。たとえば、新幹線やスペースシャトルのような巨大な物体でも、実はいろいろな細かい要素技術の組み合わせによって、ひとつの大きなシステムを組み立て、所期の目的を果たそうとした形だというのはわかりますね。

顔の認識システムでは、たとえば、空港での出入国管理、あるいは、もっと一般的に、建物や部屋の入退出セキュリティなどで実用化を考えると、処理能力や処理精度を上げるためには、個々のステップの性能とシステムとしての全体像の両方を考える必要があります。

今回、幸いなことに、日本女子大学の小館研究室(写真1)を柳隊員と一緒に訪問して、大変ユニークな高速顔認識システム『FARCO (Fast Face Recognition Optical Parallel Correlator)^{1), 2)}』についてお話を伺う機会が得られました。」

「みなさんとても素敵なお方でしたね。同じ女性として、すごく刺激を受けました。」

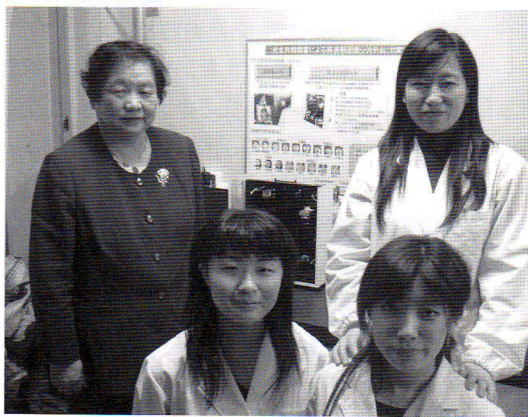


写真1 日本女子大学 小館研究室

後列左から、小館 香椎子教授、渡邊 恵理子博士、前列左から、太田 真衣子さん、石川 さゆりさん

「ええ、そうでしたね。さて、我々がFARCOシステムに大いに興味を覚えた理由は、いくつかあります。まず、エンジンの部分でいうと、フーリエ変換を行った画像間の操作を光学的に行うというアナログ過程があること。そして、その部分で並列処理が行われていること。また、顔認証サーバと携帯電話という多数の端末からなるシステムの実験が実施されていることでした。」

2 携帯電話顔認証システム —出席管理の実験

携帯電話の画面に点がふたつ現れる。それに両方の黒目を合わせて自分の顔を撮影し、サーバに送ると、サーバ上の顔データベースとの対比で認証が行われる。これが、日本女子大学 小館研究室で開発された顔認証システムの実用的な実験のひとつである。

早速、実験。男の顔ではシステムがびっくりするといけ
ないので、柳隊員に実験してもらう。

「携帯電話で手軽に撮影。すぐに出席管理ができて便利だし、代返も防止できるのね。もともとそんなズルしようとする人はいなかったかもしれないけど。」

「なるほど。面白いシステムデザインですね。学生の持っているカメラ機能付き携帯電話には、あらかじめ、Javaのアプレットがダウンロードしてあるそうです。実際の顔認証は、送られてきた画像を使ってサーバ側で行い、結果を送り返すという仕組みです。

サーバ側の認証プロセスには、光素子をつかったアナログ部分があるということで、そのハードウェアを見せていただきました。

四辺が30cm、厚みが20cmほどの筐体に、なにやら、光学実験のようなセットアップがしてあります(図1)。」

3 画像比較は積和演算 —パターンマッチングの基本

「隊員の二人も、画像処理の基本的な手法にパターンマッチングがあることは、よくご存知ですね。その中でも正規化相関はよく使われています。

大きな実際の画像 I の中に、小さなモデル画像 M を、原点を (u,v) だけずらして重ね合わせたときの相関係数 $r(u,v)$ を考えます。この相関係数が最大になるような (u,v) を検出することで、そこにモデルが存在する可能性があるとするわけです。正規化相関係数は次式で定義されます。

$$r(u,v) = \frac{[N \sum I_i M_i - (\sum I_i)(\sum M_i)]}{\sqrt{[N \sum I_i^2 - (\sum I_i)^2][N \sum M_i^2 - (\sum M_i)^2]}}$$

ここで、

$$I_i = I(u + x_i, v + y_i)$$

$$M_i = M(x_i, y_i)$$

N : モデル画像のピクセルの総数

ウワー、二人ともお手上げなどと言わないでください。見かけはすごいですが、一番重要なのは、 $\sum I_i M_i$ (i はモデル画像のすべてのピクセルについて)の項だけで、あとゴチャゴチャあるのは、スケールや基準点に合わせて、 $-1 \leq r \leq 1$ に正規化するためです。

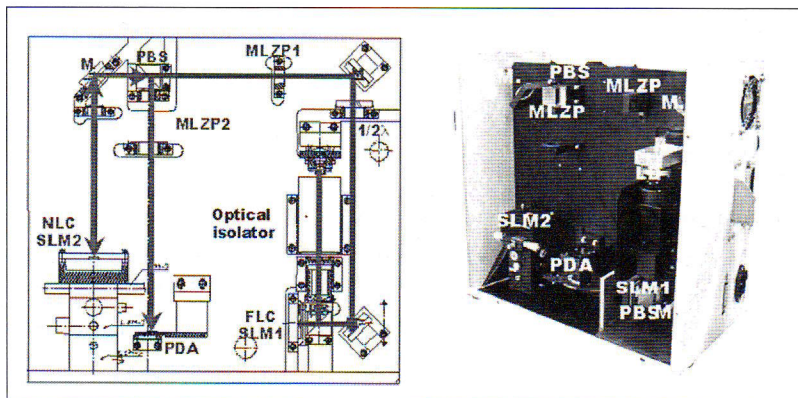


図1 光学デバイスの構成

MLZP: multi-level zone plate,
SLM: Spatial Light Modulator,
PBS: Polarized Beam Splitter,
PD: Photo diode

出典 文献²⁾

つまり、ランタイム画像とモデル画像について、対応するピクセル値の積の和をモデル画像すべてのピクセルについてとっています。積の和ですから、積和演算と呼ばれることもあります。ふたつの画像や関数の類似度(相関度)をこの積和で測るということを覚えておいてください。」

「隊長、いやあ僕にはやっぱり難しいですね」

4 コンボリューション (畳み込み)

「この際、宇野隊員のことはほっときましょう。

この積和を求めるのに、数学的に、コンボリューション(畳み込み)とよばれる演算があります。ふたつの関数や画像を平行移動させながら、重なった部分を掛けて足し合わせて行く操作です。⊗を、コンボリューションの記号とすると、離散値で定義される2次元の関数や画像では

$$(f \otimes g)(m, n) = \sum_{i=1}^M \sum_{j=1}^N f(s, t) \cdot g(m-s, n-t)$$

連続関数や画像では、和を積分で置き換えて、

$$(f \otimes g)(x, y) = \iint f(s, t) \cdot g(x-s, y-t) ds dt$$

と定義されます。要するに、このコンボリューションの最大値を与えるような (m, n) や (x, y) で、ふたつの関数や画像の類似度が一番高く、あとは、その関数値が十分高いかどうかの判断になります。

5 フーリエ変換の登場

「マシンビジョンでの一番簡単なピクセル/物理単位変換は、1対1の対応であって、何ピクセルを何ミクロンに換算するか、というだけです。もう少し変換らしいのはハフ変換ですね。これも、1対1の対応ですが、位置座標 (x, y) と曲座標 (ρ, ϕ) の間の変換です。

さて、いよいよ、真打、フーリエ変換の登場です。2次元のフーリエ変換とその逆変換は次式で定義されます。

フーリエ変換

$$F(u, v) = \iint f(x, y) \cdot \exp[-2\pi i(xu + yv)] dx dy$$

フーリエ逆変換

$$f(x, y) = \iint F(u, v) \cdot \exp[2\pi i(xu + yv)] du dv$$

実空間の関数 $f(x, y)$ と、周波数空間の関数 $F(u, v)$ とを結びつける変換ですが、周波数空間において、ある点 (u, v) における関数の値には、実空間のすべての点 (x, y) からの寄与があります。」

6 顔認識システム

「さて、顔認識の問題に戻りましょう。半導体やエレクトロニクスの画像処理でのパターンマッチングでは、テンプレートというあらかじめ登録したモデル画像(部品などの画像であり、ひとつ、あるいは複数であっても比較的少数)と、ランタイムの画像を比較するわけです。ランタイム画像のサイズはモデル画像に比べて大きく、その中に小さなモデル画像を探すというのが典型的な例です。

対照的に、顔認識ではテンプレートが画像データベースであって大量にあります。前々回、顔の抽出のところで見たとように、ランタイム参照画像は入力画像全体ではなく顔の部分だけあらかじめ切り出してあるので、被検出画像(データベース画像)とほぼ同程度の画像サイズです。

比較の対象となる被検出画像(データベース画像 f) と参照画像(ランタイム画像 g) の類似度(相関)を測るために、それらの間のコンボリューション $h(x, y)$ を考えます。

$$h(x, y) = (f \otimes g)(x, y) = \iint f(s, t) \cdot g(x-s, y-t) ds dt$$

このとき、 h, f, g のフーリエ変換を H, F, G とすると、

$$H(u, v) = F(u, v) \cdot G(u, v)$$

という関係があることを示すことができます。いいかえると、周波数空間では、畳み込みは被検出画像と参照画像それぞれのフーリエ変換の積になっているので、乗算という簡単な演算でできることを示しています。

この乗算演算を光学的に行ってしまうのが、FARCOシ

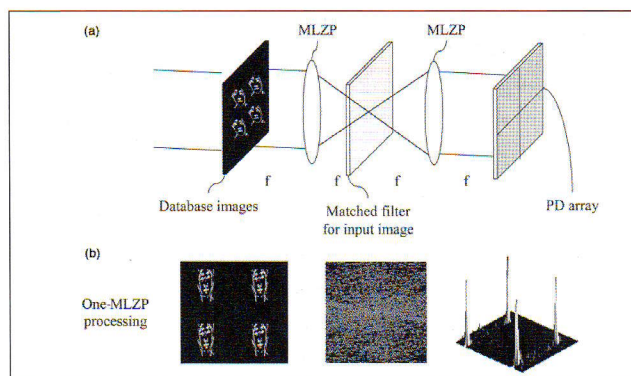


図2 光相関を使った並列演算、マッチトフィルタリング

MLZP: multi-level zone plate、PD: Photo diode

この例では、4つの並列処理をしている

出典 文献²⁾

システムのエンジン部分ということになります(図2)。実際のシステムで、このふたつの画像間のマッチングを行うには、参照画像のフーリエ変換の複素共役なもの(反転した画像、すなわち、 $-2\pi i v x$ の代わりに $+2\pi i v x$ のように位相の方向を逆転したもの)をマッチトフィルタとして使っているそうです。

7 ふたつの空間の狭間で —実空間と周波数空間

「レンズやミラー、レーザー光源も、みな実空間にあるのに、『えー、なんで周波数空間の操作ができちゃうの???』という疑問を持ったら、上出来です。個人的なことで申し訳ありませんが、私のももとの専門はX線結晶学でして、X線を結晶に当てて回折強度を測定し、結晶の原子配列を計算しようということをやっていました。ブラッグ反射とかラウエ斑点とか、高校の物理の教科書に出てくるかもしれませんね。

実際にカウンタで計測するX線の回折の位置は、周波数空間(結晶学では逆空間といいます)の幾何学的形状(逆格子と呼ばれます)で決まり、回折強度はフーリエ係数の絶対値に対応しています。実空間で測定できる量が逆空間の変数ということで、不思議に思った記憶があります。」

8 さらに高速化の試み

「FARCOでは、ひとつの比較に約1ミリ秒かかるそうで、そうすると、毎秒1,000画像程度の処理ができることになります。現在のFARCOシステムでは、この比較処理をひとつひとつのデータベース画像に対して行うのではなく、データベース中の顔画像を4つずつまとめて、並列処理をしています。この4並列処理システムでは、その4倍、約4,000顔画像/秒が可能ということでした。それならば、さらなる高速化には並列処理のチャンネル数を増やせばということになります。物理的には50~60並列は可能²⁾ だそうですが、今度はデジタル画像を光変調器に変換する時間が問題になってきます。

光演算は光速で行われるはずなのに、どうして、1ミリ秒もかかるのだろうかというのは当然の疑問です。

時間がかかるのは、データの変換(デジタルデータの画像表示)の部分とデータベース画像の入っているRAM容量の制限に原因があるのだそうです。それならいっそのこと、画像の保存もデジタルでなく光学的にしたらということで、ホログラフィック・メモリ(ホログラフィック・オプティカル・ディスク)を使ったシステムを実験³⁾しています。このシステムだと、ひとつの相関演算は10マイクロ秒程度になり、FARCOに比べ100倍の速度になります。100,000顔画像/秒の速度で処理が可能になり、顔の数が数万、数十万というような大きな顔データベースでも、ほとんどリアルタイムに近い、十分実用になる処理速度を実現することができます。」

9 デジタル処理対アナログ処理 シリアル対パラレル

「画像処理の対象となる画像は、位置も離散的、値も不連続、操作自体もその世界です。そこに、「」があらわれると、我々はハッとするわけです。隊員の二人はどうですか？ 心情的にデジタルですか、アナログですか」

「隊長、僕はどちらかというとアナログ、だって、自分、不器用ですから…」

「(ボカッ)私はデジタル思考かしら。でも、時計は針のあるアナログのほうがいいのよね。残り時間とか一瞬でわかるというか。」

「おお、面白いことをいいますね。二人は、エレクトロンもフォトンも、同じく光速で運動しているはずなのに、どうしてコンピュータだと遅くて、光演算だと高速なのだろう…、と疑問を持ちませんでしたか。」

えっ？ 持たなかった？

デジタルコンピュータといったって、エレクトロンを1個1個数えているわけではなく、電流が流れ電圧が何ボルトより上だと「1」、何ボルトより下だと「0」、その間では不定、など取り決めているわけです。時間的に考えると、コンピュータは本質的にシリアルなんですね。いくら、パラレルといったってその度合いはたかが知れています。そこへいくと、レンズなど光学的なものを考えると、レンズの周辺を通る光だろうが、中央を通る光だろうが、全部パラレルで同時進行です。

画像は面だから、本来はパラレルなものです。左の上だろうが、右の下だろうが時間的には同等なのに、コンピュータで処理するときは、スキャンしてシリアル化する、時間的に線に置き換えています。

今回、顔認識システムとして、FARCOを見せていただいたわけですが、単に顔というだけでなくマシンビジョン技術を考えてとき、『線志向』に捕われてばかりいないで、『面志向』も怠らないようにしないといけなと、自省の意味も含めて感じました。」

10 訪問を終えて

「訪問するのが女子大。勇気がいるので、柳隊員に同道してもらうことにしたわけですが、私は日本の女子大は初めてでした。アメリカにいた頃は、いまはハーバード大学と一緒にいるラッドクリフ(Radcliffe)、ペンシルベニア大学の近くには、津田塾創立者、津田 梅子がかつて留学していたプリンマー (Bryn Mawr)、ボストン郊外には、クリントン前大統領夫人ヒラリーの母校、ウエルズレー (Wellesley) などのカレッジが、身近なところにありました。

日本人の女性科学者というと、湯浅 年子、柳澤 桂子、米澤 富美子などが思い浮かびます。このような日本人の女性科学者に共通していることは、今回お目にかかった小館先生もそうですが、ご自身の優れた研究能力に加え、周りを引っ張っていく組織能力(手腕という言葉が適切かも知れませんが)に優れておられることが印象的です。究極の研究室というのは、先生と生徒という間柄だけでなく、指導者、卒業生、上級生、下級生と、レベルの違う仲間がお互いに影響しあいながら、成長していく“塾”のようなものであろうというのが私の信念です。とても気持ちのいい見学訪問でした。」

謝辞：取材させていただいた日本女子大学 小館 香椎子教授、渡邊 恵理子博士、そして研究室の学生の皆さんに感謝いたします。

参考文献

- 1) 渡邊 恵理子、小館 香椎子、光相関に基づく高速顔認識システム、画像ラボ、2006年、3月号、p19-23.
- 2) Eiko Watanabe and Kashiko Kodate, Implementation of a high-speed face recognition system that uses an optical parallel correlator, Appl. Optics, vol. 44, No. 5, p666-676, 2005
- 3) Eiko Watanabe and Kashiko Kodate, High speed image search engine using collinear holography, Optical Pattern Recognition XVII, David P. Casasent, Tien-Hsin Chao, Editors, 2006 Proc. of SPIE, vol. 6245, 62450E1-7, April 2006

ビジョン探検隊・今月の一言

大橋隊長

最近、物事をデジタルに考えるようになって(カメラまでそうです)、光相関演算などというアナログの話を聞くとたいへん新鮮に感じます。私の学位論文は、高温における結晶のX線構造解析だったので、フーリエ変換と聞くと血が騒ぎます。ただ、結晶構造は3次元なので、画像の平面への投影ではなく、ホログラフィーのようなものでないかとだめでしょうね。小館先生と私は同い年。渡邊さんは私の娘(米国国立研究所で研究生生活を送っています)と同い年。何かの因縁かなあと感じました。

柳隊員

今回お邪魔した日本女子大学の小館研究室、研究に打ち込んでいるみなさんがとても素敵に思えました。私もこの探検にもっともっとのめり込みたいと思います。

(おっ、いいましたね(大橋))

宇野隊員

うまく言えないけど、アナログっていうのは瞬間的に全体を把握する、っていう性質があるのかなと感じました。デジタルは道を辿っていくというか。人間の視覚はどうなっているんでしょうね。



大橋 慶一 Ph.D.

yoshi@eizojoho.co.jp

米コグネックス社で長年、製品企画、製品開発、販売戦略などに携わる。2004年、帰国。現在の関心は知的財産のビジネス化。

Topics

コグネックス、自動車運転支援システムのリーディングデベロッパAssistWareを買収

コグネックスは、新たな分野の取り組みとして車載用マシンビジョンセンサ市場に参入することを発表し、車線逸脱警報システムの開発を手がける、米AssistWare Technology社を買収した。

AssistWareのSafeTRAC車線逸脱警報(LDW: Lane Departure Warning)システムは、マシンビジョン技術を応用して前方の道路を監視し、自動車が車線を逸脱したり、運転パターンが不規則になったりとすると警告を発する技術で、レーダー探知器と同じくらいの小さなサイズで、ダッシュボード上に設置される。LDWシステムは、小型のビデオカメラ、デジタル信号プロセッサ、そしてAssistWare独自のマシンビ

ジョンアルゴリズムを使用しており、道路上の車両の相対位置を判断して、車両が車線外にはみ出したり、蛇行していたりした場合に警告する。

同社ではまず、米国だけでも年間30万台販売されているという大型トラック市場への取り組みに力を入れ、この車載ビジョンセンサを同社の新しいビジネスへと成長させていくとしている。

☆コグネックス株式会社

ビジネス・デベロップメント本部

TEL.03-5977-5446 FAX.03-5977-5401

<http://www.cognex.co.jp/>