

携帯向け顔画像処理ミドルウェア FSE (Face Sensing Engine) の開発

井上 清司

本稿で取り上げるFSE (Face Sensing Engine) ^{1) 2)} は携帯端末市場を主なターゲットとして開発された顔画像処理ミドルウェアです。FSEは、2006年4月にボーダフォン株式会社 (現在のソフトバンクモバイル株式会社) の国内向け携帯電話への搭載が開始され、その後海外の携帯電話市場へも展開を加速中です。本稿では、「OKIの新規市場における新規事業創出」というミッションを負って発足し、FSEを世に送り出している当該プロジェクトの、技術開発と市場開拓の取り組みをご紹介します。

プロジェクトの歴史

FSEを開発しているプロジェクトは画像認識技術をコアとしたヒューマンインタフェース技術を研究開発するチームです。我々が顔認識技術に興味を持ち始めたのは1997年頃でした。しかし、当時社内ではアイリス認証という非常に認証精度の高いセキュリティ技術に注力していたこともあり、アイリス認証に比べて認証精度の低い顔認識技術にはあまり興味を持たれませんでした。当時はカメラとその映像を処理するキャプチャボードやPCが比較的高価で、あまり一般的に普及していなかったこと

も原因かもしれません。

そのような中で携帯電話向けに顔を中心とした画像認識技術の開発に着手したのは今から約5年前でした。きっかけとなったのは、NTTドコモによるテレビ電話機能搭載の携帯電話FOMAの発売でした。当時、日本の景気はバブル崩壊の後遺症で苦しんでおり、多くの金融機関が資金を貸し渋り、企業もまた設備投資を控え、日本企業全体が相変わらず悪循環の波に揉まれていました。そのような中で携帯電話市場だけは飛躍的に拡大し、FOMAは第3世代携帯電話への牽引役として情報・通信マーケットから高い注目を浴びていました。その中で我々のプロジェクトが目じたサービスのひとつがテレビ電話だったのです。なぜならば、当時テレビ電話機能は第3世代携帯電話の高度な通信機能を生かすサービスとして大きく期待されており、またこの機能を携帯キャリアが標準サービスとして提供するためには、テレビ電話機能に必須となる「人の顔を写すためのカメラ」が携帯電話に標準で装備されるに違いなかったからです。

当初、FSEの開発はFaceCommunicator^{®*1) 3)} というテレビ電話を安心してより楽しく利用するためのソリュー



図1 FSEを核とするさまざまなソリューション

*1) FaceCommunicatorは沖電気工業株式会社の登録商標です。

ションを実現するための技術としてスタートしました。図1に示すように、現在プロジェクトでは、FSEという顔画像処理ミドルウェアを核とし、種々のライブラリを組み合わせることで、ニーズに応じた幾つかのソリューションを提供しています。FaceCommunicatorはそのソリューション群のひとつですが、プロジェクト発足当初は、後で述べる理由によりFaceCommunicatorというソリューションを、まずは世に出すことを第一の目標に開発を進めていました。

FaceCommunicatorとはテレビ電話の際、直接顔を出したくないとか、背景を見せたくないといった思いをお持ちの方々に提供するソリューションの一つです。カメラが捕らえた話者の顔の動きをリアルタイムに3次元コンピュータグラフィックス（CG）のキャラクターの動きに反映し、話し相手があたかもCGキャラクターであるかのように置き換える機能を提供します（図2）。

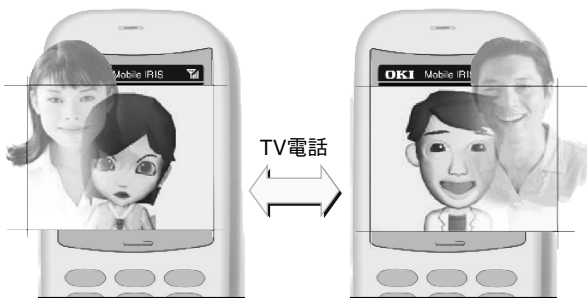


図2 FaceCommunicator

我々が、このソリューションをターゲットに技術開発を開始した理由は2つあります。1つは、勿論そこに成長すると予想されるマーケットが存在したからです。そしてもう一つは、一見矛盾するようですが、認識技術（この場合は顔認識）というものの完成度が短期間では完璧にはならないということを我々が良く理解していたからです。二つ目の理由についてもう少し補足しましょう。

従来こうした認識技術の精度追求には3~4年掛けていました。そして製品化は認識技術がある程度完成するのを待ってから行っていました。ところが対象市場と考えていた携帯電話市場の進歩と変化の速さはすさまじく、我々の開発完了を待ってくれません。顔認識技術を開発するにあたり、当時の我々には技術的蓄積と経験が無く、開発を加速させる要因は何もありませんでした。一方、進歩と変化の早い携帯電話市場で我々が成功するためには、早くポジションを確保し、市場をより深く理解しつつ戦略を練ることも必要です。そこで我々は、開発の途中でも、つまり認識精度が多少低く、時として結果を間違

ことがあったとしても、利用者に受け入れられるソリューションを考えなければ駄目だと判断したのです。それがあある意味で、実用的でありながらもエンターテイメント的な要素を取り入れたFaceCommunicatorというソリューションを目標としてスタートした理由でした。こう判断した背景には、やはり純粋な技術開発という断面だけでなく、新規事業創出というミッションを踏まえ、短期間で開発したものを市場へ投入し、売上げ実績を上げていく必要があったということも理由の1つでした。開発には当初予想していなかったさまざまな障害があり、苦労の連続（残念ながら今も続いています）でしたが、結果的に開発開始から半年後にはユーザーへ試作版を販売していました。その後さまざまな形で開発途上のソフトウェアを販売し、毎年収益を上げながら開始から4年が経過しています。現在は、ようやく認識精度の完成度が向上してきたことにより、FaceCommunicatorの中心部のエンジンを組み用画像処理ミドルウェアFSEとして切り出し、認識技術そのものを価値として携帯電話市場向けに販売しています。

以下では当該プロジェクトが、技術の開発とその製品化に対してどのように考え進んできたかについて説明します。

技術開発とビジネスインキュベーション

画像認識に限った話ではありませんが、通常研究開発段階のアルゴリズムを改善して認識精度を向上させようとすると、往々にしてその実行速度は遅くなるものです。しかし、このプロジェクトが取り組んできた技術開発は、短期で成果を出すという制約の中、初期の段階から、単に認識精度を上げるという目標だけでなく、携帯電話というCPU性能の低い端末上で認識精度を向上させつつ実用的な速度を達成するという、二つの課題を同時にクリアする必要がありました。我々の最初の試作プログラムは、当時の携帯のCPU上で動作させた場合、認識処理に数秒必要でした。しかし、実用に供するためには数十から数百ミリ秒で認識処理を終える必要があります。この状態を開始地点とし、認識精度を向上させながらも、10倍、20倍高速化しようということですから非常に困難な取り組みです。しかしそれができなければ売れないという現実しか待っていないのです。したがって、当初からこの顔画像処理エンジン（FSE）の開発には目標性能に到達するまでに存在する高い事業リスクを分散する必要があります。そうするために、段階的でも開発途中のソフトを製品化して販売していくというアプローチを考えていました。

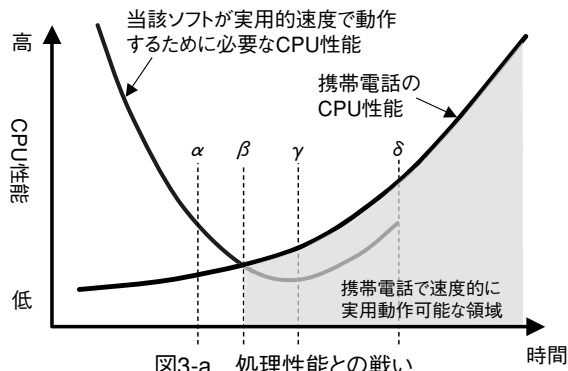


図3-a 処理性能との戦い

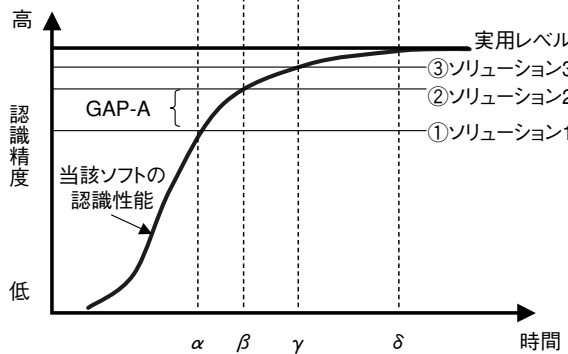


図3-b 認識性能との戦い

図3 Haphazard Approach

それが図3に示すHaphazard Approachです。これは「行き当たりばったりのアプローチ」という意味で、私が勝手にそう呼んでいるものです。なぜ我々のアプローチを「行き当たりばったり」という一見無計画にも聞こえる表現で呼んでいるのかということをお次に説明します。

技術開発チームは当然高い目標に向かって懸命に開発を進める必要がありますが、そもそも競合の多い世の中で、実用に耐え、かつ競合に打ち勝つレベルの認識精度を実現するのは大変です。その上、携帯電話という性能の低いCPU上で実用的な性能になるようにしると無理をいいますから、開発チームのモチベーションを維持し、プロジェクト全体の運営費用を継続的に確保するためには、やはり現実的にできたところまで、どうやってビジネスにするのかといった取り組みをする必要があります。どこまでの性能がいつまでに達成できるのかは、研究要素の高い開発の場合、実際にやってみなければ分かりません。一方、お客様のニーズは日々変化しており、技術レベルとニーズの最適なマッチング（すなわちソリューション）は、走りながら考えざるを得ない場合が多いと思います。それが行き当たりばったりという表現になったのです。（勿論、文字通り常に行き当たりばったりというわけではではありません。短期的な計画も中期的な計

画も立てつつ進めるのですが、環境の変化や自分達の技術の達成度に応じ、柔軟かつ大胆な計画変更を厭わないという意味です。）

図3-aは処理性能との戦いを、その期間とCPU性能との関係に着目して示した図です。図中の『当該ソフトが実用的速度で動作するために必要なCPU性能』というグラフが、我々が開発している認識技術の開発の変遷（その時々における目標でもあります）を表しています。図3-aに示したように開発を始めた当初は携帯電話のCPU性能はかなり低いものでした（クロック周波数が100MHz以下）。一方、図3-bに示すように画像認識部の認識性能も、開発当初は十分高いとは言えません。この時点での開発では、速度向上のために処理を簡略化・最適化しつつも精度が劣化しにくいアルゴリズムをいかにして見つけるかという観点で、適宜目標を変え地道なアルゴリズム開発が行われました。

先にも述べたように、ビジネス面から総合的にみると、開発の節目で段階的に製品を売っていかねば、リスク分散ができません。そこで、図3-bの破線で示したように、認識精度の成長段階でいくつかのソリューションに分けて販売することを考えたわけです。たとえばαの時点での認識精度でソリューション2を実現しようとする、認識精度の実用レベルが低く、GAP-Aのギャップを埋めて実用レベルに引き上げるための手立てを打たねばなりません。もし開発初期段階であれば、短期に精度が向上する可能性があります。開発中期あるいは後期では、小手先での精度向上は望めないため、ソリューション自身を現状の認識精度に合ったものに変えて製品化し販売することも試みなければなりません。こうした、試み自身もリスクですが、結果として何とか事業は継続しています。このアプローチで開発してきたソリューションの例を挙げます。

図3-bのα（携帯電話上では遅くて動かない時期）ではPC用のビデオチャット用ソフトウェアとして、FaceCommunicator-BBEと呼ぶ製品を販売しました。携帯電話上でようやく動作するくらいの処理性能になったβ時点では、画像合成機能を付加して画像合成ソリューション（図1右下）を発売し、韓国市場で採用されました。γ時点（CPU性能の向上に合わせて、認識精度を大きく改善するための折り返し地点）では顔認証ソリューション（図1左上）を製品化し、当時のボーダフォン株式会社（現在のソフトバンクモバイル株式会社）に採用されました。そして、現在はδ時点（完成度の高い認識精度実現時期）のポイントに向けてセキュリティソリューションを展開すべく開発を加速しています。

今後の展開

当初FSE-Ver1のターゲット市場は携帯電話のインカメラでした。インカメラというのは冒頭で説明したテレビ電話のためにつけられた2つ目のカメラ（ディスプレイ側に埋め込まれ、携帯電話の操作者側を向いたカメラ）です。しかし、テレビ電話可能な端末が広く普及してきた現在でも、残念ながらテレビ電話の利用者が増加しないため、FSE-Ver2として顔認識機能を追加開発し、インカメラの利用価値向上を模索するキャリアに向けた端末ロックソリューションを展開してきました。

今後、我々はFSEをグローバルに展開していこうと考えています。海外の携帯電話のカメラ搭載状況を見ると、インカメラの搭載例は少なく、アウトカメラ（いわゆるデジタルカメラ的な用途のカメラ）が搭載されている携帯電話の方が圧倒的に多いようです。そこで、アウトカメラの撮影特性に対応したFSE-Ver3を開発し2006年11月より販売開始しています。FSE-Ver3では、従来の顔の検出機能を強化し、小さく撮影された顔を検出できる他、複数人の顔を同時に検出し、且つ予め登録されている人物の顔と瞬時に照合を行い、撮影中の人の中から特定の人物だけを探し出すことが可能になります。また、FSE-Ver2まではキャリアを中心としたソリューションを展開して来ましたが、FSE-Ver3からは携帯電話メーカーを対象としたソリューションを展開しています。

図4にFSE-Ver3のマルチ顔検出機能を利用したソリューション例を示します。予め友達や、自分の子供の顔を登録しておくことで、撮影時に画像の中から登録してある人物の顔位置を検出し、その人物にオートフォーカスすることが可能となる他、写真を保存する際、ファイル名に被写体の名前を挿入するといったことも可能になります。このような機能が搭載されれば、エンドユーザーの撮影画像に対する検索やファイル整理作業が非常に楽になると考えています。

日本の携帯電話は、単なる「電話」という枠組みを超



図4 FSE-Ver3のマルチ顔検出機能

え、ユビキタスサービス時代のキーデバイスとして、また個人の生活を便利かつ豊かにしていくプラットフォームとして発展（イノベーション）しています。そのイノベーションのひとつにカメラ機能があると言えるでしょう。携帯電話内蔵カメラはものすごい速度で高性能化し、最近では通常のデジタルカメラ並みの解像度が内蔵された携帯電話も販売されています。これは、いつでもどこでも写真を撮影し、記録として残したいというユーザーニーズがあることの現れであると言えるでしょう。我々はFSEをこうした携帯電話のカメラのイノベーションを通して、個人の利便性や、安心・安全を向上させるコアとして供給していきたいと考えています。

またカメラ自身の成長は携帯電話だけでなく、汎用のデジタルカメラ・ビデオカメラ・監視カメラにもあり、リアルタイム処理を得意とするFSE-Ver3は他の組み込み機器にも搭載が可能となっています。さらに、地上波デジタル放送の普及やHDDの小型化が組み込み機器への動画の蓄積保存を促進し、さまざまなシーンで人の顔の検索が必要になってくると考えています。

最後に、このプロジェクトが目指すものはソフトウェアライセンスビジネスによる高収益事業の創出です。世の中で誰もがソフトウェアを容易に作れる時代に、いかに他にまねのできないソフトウェア（商品）を作るかという課題と、そしていかに継続して収益を獲得するかという課題の両方を克服しなければなりません。そのため我々のイノベーション開発は今後も止めることなく続くに違いありません。 ◆◆

参考文献

- 1) 天本, 増田, 杉岡: 顔認証で携帯電話アクセスをコントロール—組込向顔画像処理ミドルウェアFSEとその携帯電話への応用事例—, 画像ラボ, Vol.17, no.10, 2006年
- 2) Mobile phone security in your face, Embedded Computing Design, Vol.4, No.5, 2006年
- 3) 宮崎, 井上, 天本, 石川: ビジュアルコミュニケーションに安心と楽しさを付加するソフトウェア - FaceCommunicator -, 沖テクニカルレビュー197号, Vol.71 No.1, pp.114-117, 2004年1月

筆者紹介

井上清司: Seiji Inoue. 情報通信事業グループ インキューベーション本部 センシングソリューション開発部