

MITSUBISHI
Changes for the Better

家庭から宇宙まで、エコチェンジ



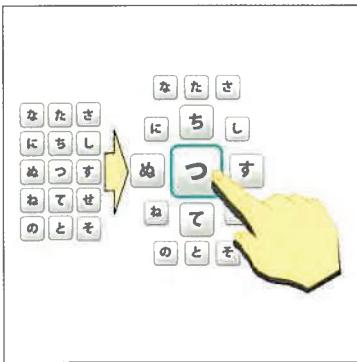
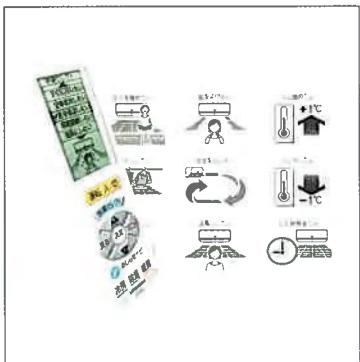
三菱電機技報

11

2011

Vol.85 No.11

特集 「ヒューマンインターフェースを
支える先端技術—新たな価値創造に向けて—」



目 次

特集「ヒューマンインターフェースを支える先端技術—新たな価値創造に向けて—」

学びの場のロボット

—ヒューマンインターフェース技術特集号によせて— 1

三宅なほみ

ヒューマンマシンインターフェース技術への

取り組みと今後の展望 2

木槻純一・石川 泰

“三菱エアコン霧ヶ峰ZWシリーズリモコン”

ヒューマンマシンインターフェースデザイン 7

小川 孝・石塚健彦

車載機ヒューマンマシンインターフェースの

ユーザー評価手法 11

若松正晴・嶋田 淳・沢田久美子

カーナビゲーション向け入力支援技術 15

松本 篤・相川勇之・平井正人・山崎 航

マルチメディアを活用した技能継承システム 19

阿倍博信・峯 慎吾・奥村誠司

音声検索技術とその応用 23

成田知宏・花沢利行・岩崎知弘

テキスト音声合成技術 27

大塚貴弘・川島啓吾・山浦 正

3Dタッチパネル技術 31

佐々木雄一・宮原景泰・川又武典

組み込み機器向け文字

ベクターグラフィックス描画回路 35

加藤義幸・鳥居 晃・濱田雅樹

図形のデフォルメ技術 39

菅沼優子・根岸博康・川又武典

画像同定技術：Visual Signature 43

工藤大樹・西川博文

モービルマッピング技術 47

前原秀明・石渡要介

特許と新案

「音声対話装置」「景観表示装置」 51

「楽曲分析装置及び楽曲検索装置」 52

Advanced Human Interface Technologies for Value Creation

Robots as Learning Partners—Forward to Special Issue on Human Interface Technologies—

Naomi Miyake

Perspective on Human Machine Interface Technologies and Future Prospects

Junichi Kitsuki, Yasushi Ishikawa

Human Machine Interface of "Mitsubishi Room Air Conditioner Kirigamine ZW series"

Takashi Ogawa, Takehiko Ishizuka

Usability Measurement Method for HMI of Automotive Infotainment Systems

Masaharu Wakamatsu, Atsushi Shimada, Kumiko Sawada

Input Support Technology for Car Navigation System

Atsushi Matsumoto, Takeyuki Aikawa, Masato Hirai, Wataru Yamazaki

Multimedia Skill Succession System

Hironobu Abe, Shingo Mine, Seiji Okumura

Voice Search Technology and Application

Tomohiro Narita, Toshiyuki Hanazawa, Tomohiro Iwasaki

Text-to-Speech Technology

Takahiro Otsuka, Keigo Kawashima, Tadashi Yamaura

3-Dimensional Touch Panel Technology

Yuichi Sasaki, Kageyasu Miyahara, Takenori Kawamata

Font and Vector Graphics Engine for Embedded Systems

Yoshiyuki Kato, Akira Torii, Masaki Hamada

Deformation Technology for Graphics

Yuko Suganuma, Hiroyasu Negishi, Takenori Kawamata

Image Identification Technique: Visual Signature

Daiki Kudo, Hirofumi Nishikawa

Mobile Mapping Technology

Hideaki Maehara, Yosuke Ishiwatari

スポットライト

カーナビゲーション “NR-MZ50”

表紙：最新のヒューマンインターフェース技術を搭載した製品

この特集号では、ヒューマンインターフェース技術を特集している。

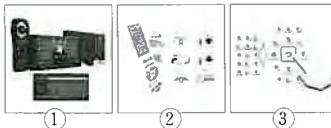
インターフェースデザイン、評価技術から、その要素技術である音声認識・合成技術、情報表示技術、データの構造化技術など、使いやすく分かりやすい製品を支える三菱電機の最新技術を紹介している。

表紙の写真・イラストはその適用製品、技術事例である。

① 少数の入力文字から施設名を高速に予測する技術を適用するとともに、携帯電話型の文字入力、分かりやすいボタンデザインなど、操作性が高いカーナビゲーションNR-HZ001。

② 業界で初めての大型フルドット液晶に、“～したい”という利用者の気持ちを表現したメニュー“快適セレクト”を搭載した、エアコン霧ヶ峰ZWシリーズのリモコン。

③ 指の近接を検知できるため、指を近づけたときにメニュー・キーが拡大するなど、使いやすく誤りの少ないインターフェースを実現する3Dタッチパネル。



巻/頭/言

学びの場のロボット—ヒューマンインターフェース技術特集号によせて—

**Robots as Learning Partners—Forward to Special Issue on
Human Interface Technologies—**

三宅なほみ
Naomi Miyake



今、市場には、自走する掃除機が出回っている。あれを“掃除機”と呼ぶ人は少ない。反対に、洗濯機を指して“ロボット”と呼ぶ人も少ない。洗濯機が留守の間に家の洗濯ものを集めて回り自分で洗剤も取りに行くとなると、これはロボットと呼んでも良さそうな気がしてくる。何がロボットかという議論の時代はもうとうに過ぎていて、今、私たちは、新しいインターフェースとして、ロボットにどんな役割を果たして欲しいかを現実的に考え始めているのではないか。

例えば、少し意外だと思われるかもしれないが、実は今、ロボットが学校の教室で子どもたちの“学び友だち”としての役割を果たせる可能性が見えてきている。

ここ2年ほど私は、小・中・高等学校の先生方と一緒に、授業の質を上げる試みに取り組んでいる。これまでどちらかと言うと先生が主役だった授業を、子どもたちが主役になる授業に少しずつ作り変えてその効果を検討するという仕事である。もっと具体的に言うと、教科書に書いてあって普通は先生が説明することを、クラスで子どもが分担して、説明し合う。次いでその分担部分を相談ずくで一緒にして、自分なりに納得できるストーリーを作る。そうして作ったストーリーを最後に発表し合って、他人の発表の気に入ったところを自分の説明に付け加えて、一人ひとりが、自分にとって最も納得のゆくわかり方を構成する。分かれで担当した知識の部品をジグソーのピースに見立て、それぞれの担当者一人ずつ集まってピースをなんとか組み合わせて答えを作ってもらうので、“知識構成型ジグソー法”といいかめしい名前まで付けている。

具体例で説明しよう。例えば、“福島で原発が壊れると、茨木のお茶農家が困るのはなぜか”という問い合わせるのに、

- ・原発が壊れるとセシウムの放射性同位体が放出され、体内被曝の可能性が増す
 - ・お茶栽培の肥料に、植物一般の成長を促進するカリウムがほぼ必ず使われる
 - ・銅山では銀や金も採れることからわかるように、元素周期表で場所が近い元素は構造や性質が似ている
- という三つの部品を分担して確認し、それぞれの担当グル

ープから一名ずつで新しいグループを作り、その3人で上の3つの観点を組み合わせて先の問い合わせに答えてもらう。どの資料を取っても、それだけでは答えは出ない。が、三つを合わせて考えてみると、実際これを試してくれたある少女が言ったように、“お茶、セシウムとカリウム間違っちゃったのね”という“その人には納得のゆく理解ストーリー”が成立する。こういう答えが出てきた時、この少女は、自分の学びの主役だったといっていいのではないか。実際この手のストーリーが作れた場合、子どもはその内容を、先生から聞いたもっと見事な説明より長く覚えていて、そこから自分でさらに調べたいことを見つけたり、学年が上がってからの次の学びにつなげてくれたりする割合が高いこともわかってきていている。

実はこの手の実践の効果は世界中あちこちで認められるようになってきているのだが、一つ困ったことがある。オトナはコドモに化けられないので、実験者がサクラとして子どもの3人組に交じってうまい支援の仕方を探るといった介入型の研究がひどくやり難い。そこで今私たちは、究極のインターフェース(！？)として、人型卓上ロボットを遠隔操作して、そっとこのジグソー活動に仲間入りさせてもらう試みを始めている。小学校5、6年生から大学生まで、試してみた限りでロボットは十分仲間として認めてもらえるようだし、ロボットがした説明も“一人前”に各自の知識構築にいかされている。おもしろいことに、このときロボットが“答えを知っているらしい”と認知されると、それだけで子どもたちの会話がぎくしゃくしてしまう。ロボットにそういう会話の記録をさりげなく取るのも便利である。そういう記録が分析できると、ぎくしゃくの原因も探しやすい。まだ始めたばかりだが、私たちの研究現場では、今あるロボットが明日の学びの仲間として、小さな一步を踏み出そうとしている。

○大学発教育支援コンソーシアム推進機構

<http://coref.u-tokyo.ac.jp/>

○文部科学省科学研究費補助金新学術領域(課題提案型)

「人とロボットの共生による協創社会の創成：ロボット共生学」

<http://www irc atr jp/human-robot-symbiosis/>

巻頭論文

ヒューマンマシンインターフェース 技術への取り組みと今後の展望



木村純一*

石川 泰**

Perspective on Human Machine Interface Technologies and Future Prospects

Junichi Kitsuki, Yasushi Ishikawa

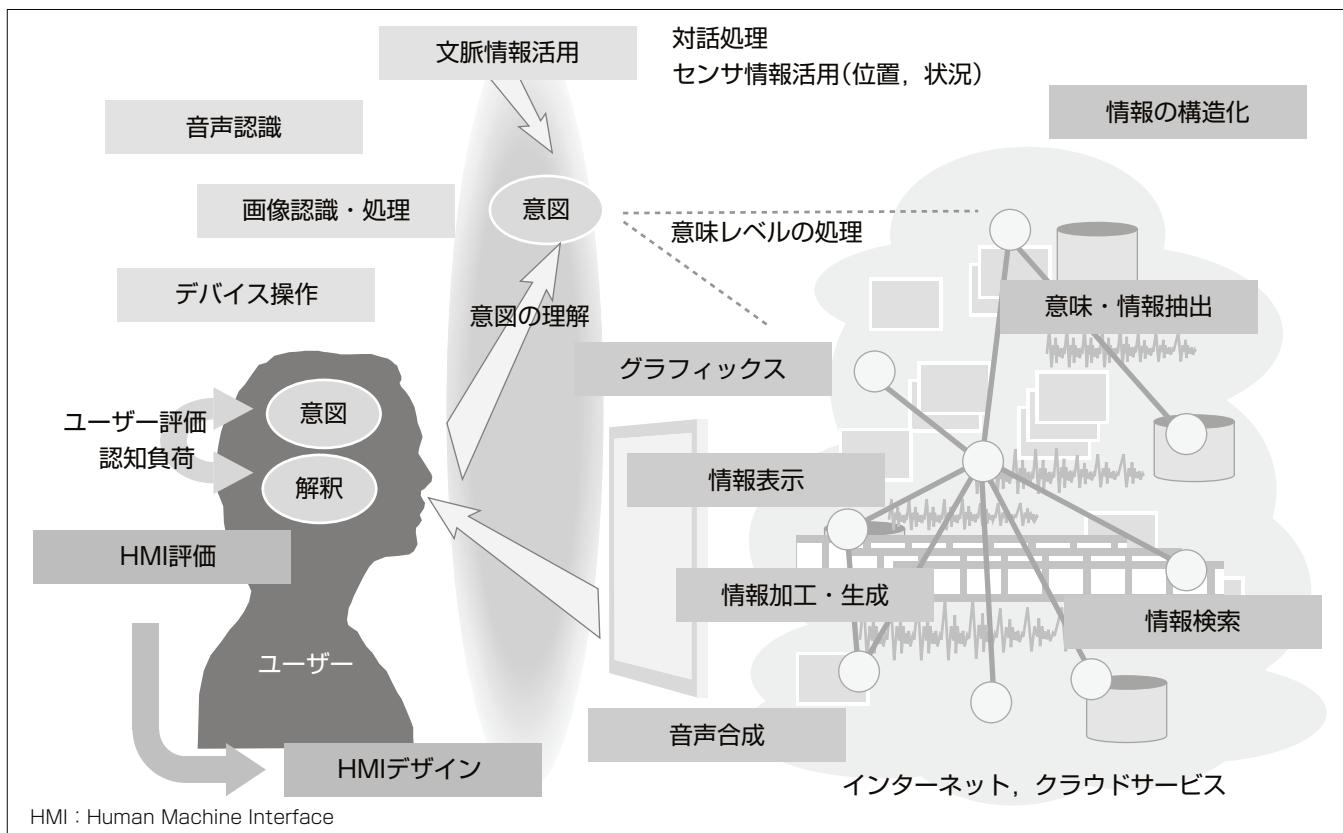
要 旨

スマートフォン、タブレット端末が普及し、多くのユーザーが高精細グラフィクスディスプレイに表示される画像、映像、アニメーションなどを含む情報を、タッチパネルで操作し、音声で情報検索をすることも珍しいことではなくなっている。ゲーム機では、加速度センサを内蔵した機器を“操作”し、カメラ映像処理によってジェスチャが認識されるようになった。かつては、“機械”的操作、その効率の視点で検討されてきたヒューマンインターフェース技術は、今、大きな変化の時代を迎えている。

インターネット上には、莫大(ばくだい)なマルチメディアコンテンツを含む情報があふれ、通信技術の進歩によって、それらの情報にだれでもがどこからでもアクセスができるようになっている。クラウドサービスは、機能やサービスがどこにあるのかをユーザーに意識させなくなった。

このような時代では、固有の機能を持った機械と人の関係ではなく、人と情報・サービス、現実世界と情報、さらには、それらを介した人と人との関係という視点でのヒューマンインターフェースを考え、製品やサービスをどのように気づき、学習し、使うようになるのかを評価し、ユーザーが得られる経験や満足を高めるインターフェース技術を提供することが、製品の価値となる。

三菱電機では、その基盤となる音声、グラフィックスなど入出力技術から、情報の構造化、利用者にとっての意味づけなどの情報処理技術までを研究開発し、これらの先進の技術を統合し、インターフェースデザイン・評価と一体となって、多様な機能や莫大な情報をユーザーにとって真に使えるものとして提供できるヒューマンインターフェースを実現していく。



ヒューマンインターフェース

人間の要求を直感的な方法で入力し、分かりやすく情報を提供するヒューマンインターフェースが、情報サービス製品の価値となる。当社は、その要素技術である、画像処理、音声処理、入出力デバイス開発から、情報の表現・加工技術、さらには、情報の関連性・意味抽出、構造化技術を開発するとともに、デザイン、ユーザービリティ評価まで一貫した研究開発を実施することで、真に“使いやすい”製品を提供している。

1. まえがき

パソコン、インターネットが家庭に広く普及し、スマートフォン、タブレット端末が急速にひろまっている。高精細グラフィックスディスプレイに表示される画像、映像、アニメーションなどを含む情報をタッチパネルで操作し、音声で情報検索をすることも珍しいことではなくなっている。ゲーム機では、加速度センサを内蔵した機器を“操作”し、カメラ映像処理によってジェスチャが認識されている。自動車の運転では、カメラ映像で確認しながらバックし、センサが障害物を教えてくれる。日々メールでコミュニケーションをとり、ネットワーク上にコミュニティを形成する。今、私たちは、情報とそれを扱う機器の中で生きていると言つていいだろう。

ヒューマンインターフェースは、古くは“機器”的操作、工場やオフィスでのその利用効率の視点で検討されてきたが、今日では、人と機能、情報、サービスとの関係、又は、現実の世界と情報との関係、それらを介した人と人との関係という、人と技術の総合的な関係性の視点で考えなければならない時代を迎えている。当社では、このような視点で、情報通信技術・情報を真に利用者にとり価値のあるものとして提供するために、インターフェース技術の開発に取り組んでいる。

この特集号では、この基盤技術について最近のトピックを解説し、インターフェース設計事例、評価技術について述べる。以降の章で、インターフェース技術を概観し、この号の論文の概要と位置付けを述べ、今後のインターフェース技術を展望する。

2. ヒューマンインターフェース

2.1 ヒューマンインターフェースの歴史

図1に、ヒューマンインターフェースの歴史を概観する。

ヒューマンインターフェースは、1970年代までは、機器の操作に関する身体的な負荷を評価し、作業における機器操作効率を考える技術であった。機器そのものの機能や効果は利用者にとって明確であり、課題は、誤操作の防止、人間工学に基づく操作効率、身体的負荷の軽減にあった。この時代コンピュータの実用化も進んだが、まだ極めて特殊な用途に用いられるだけであり、インターフェース検討の対象物にはなり得なかった。

1980年代に入ると、コンピュータが発展し、利用の中心は、オフィス業務と一部のユーザーに限られたものの、“パーソナルコンピュータ”として利用者は急速に増え始め、利用における利便性、特に利用方法の習得の観点からインターフェースの検討がなされた。1980年には、現在のマウスとGUI(Graphical User Interface)によるインターフェースが開発されている。製品でも1990年のWindows3.0^(注1)は、グラフィックス表示とポインティングデバイスによる操作を基本とするもので、現在のコンピュータ操作の原型となるものであった。それまでの、個別の機能を持つ機器とは異なり、機能そのものの理解が効率的にできるか、又は、ユーザーが目標を操作に移す場合に、誤りなく正しい操作ができるかという認知負荷の低減がインターフェース技術の解決すべき主な課題の一つとなった。これはコンピュータに限定した話ではなく、電子機器が様々な機能を持つよう

(注1) Windowsは、Microsoft Corp. の登録商標である。

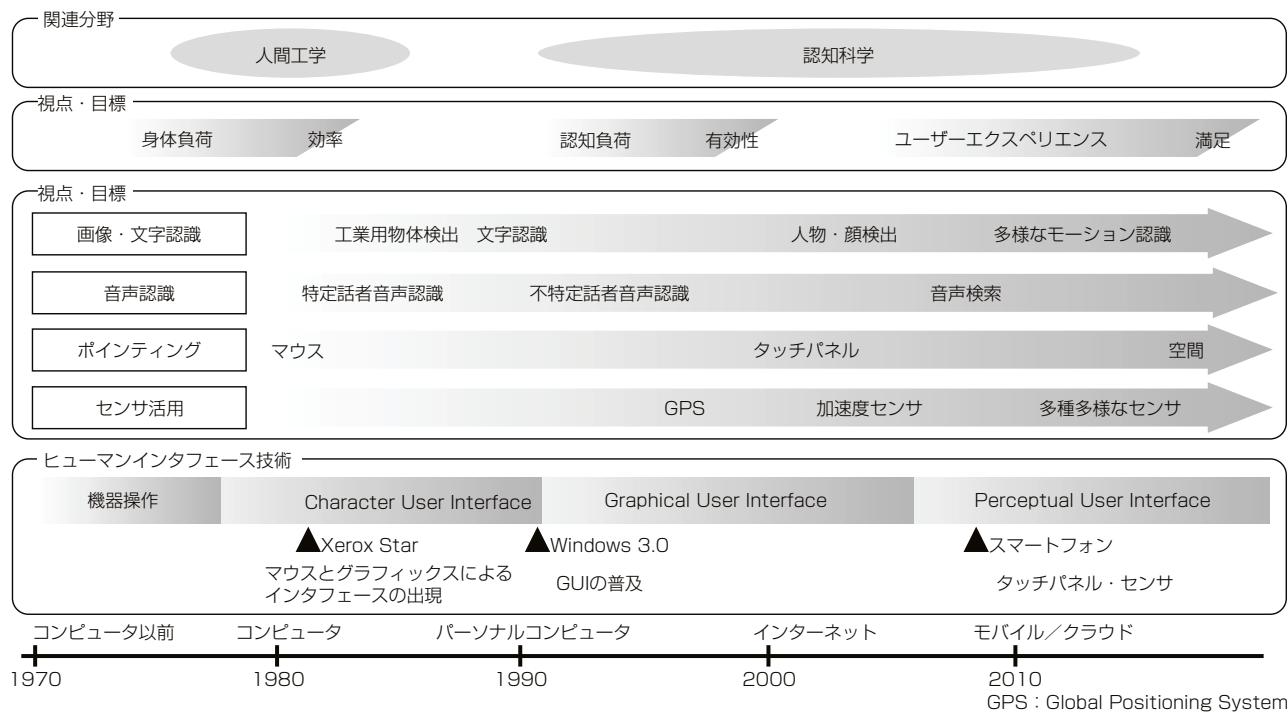


図1. ヒューマンインターフェースの歴史

になったため、これらのインターフェースデザインでも“使いやすさ”が議論されるようになった。“人間中心のデザイン”⁽¹⁾が提唱されたのは、まさにこの1980年代のことであった。インターフェース技術が人間工学視点で機械設計を考える技術分野から、認知科学を通して機能設計を考える時代に変わってきたとも言える。

1990年代後半になると、インターネットの家庭への普及が進み、パソコンはコンシューマー製品になった。直接的なインターフェース技術については大きな技術変革はないものの、情報検索技術、又は、情報のインターフェースとしてのWebデザインが議論されるようになり、インターフェース技術が単に入出力技術だけではなく、情報処理技術を含めた幅広い技術分野に拡大したと言える。

2000年代に入ると、通信の高速化、無線通信技術の発展に加え、機器の高機能化、小型化が急速に進み始めた。小さい機器での複雑な操作の問題は、携帯電話のインターフェースデザインを典型例として広く議論された。

そして現在、プロセッサの高速化・小型化はさらに進み、だれもがどこからでも大量の情報をアクセスでき、サービスを受けられる時代となった。情報通信は社会生活の重要なインフラとなっている。1990年代に議論された、PDA(Personal Digital Assistance)、ウェアラブルコンピュータ、ユビキタスコンピューティングなどのコンセプトが、形を変えて実現しつつある。インターネット上にはさらに大量のマルチメディアコンテンツが流通し、サービスのクラウド化によって、ユーザーは、どこに機能があるかを意識することがなく、今、何ができるか、どのようなサービスをいかに快適に受けられるか、コンテンツやサービスとユーザーをつなぐインターフェースに価値を求める時代になったと言える。したがってインターフェース技術は、ユーザーと情報・サービス、情報の相互関係、又は、情報や情報機器を介しての人と人の関係を広く扱う技術分野になっている。SNS(Social Networking Service)は、ネットワーク上の人ととのつながりととらえることもできるが、莫

大な情報世界を共通の興味など実世界の人のつながりを通して見るインターフェース技術ともとらえられる。

2.2 ヒューマンインターフェースの課題

新しい時代のヒューマンインターフェースについて、情報の相互関係を抽出する情報処理技術、人と機器や情報・サービスの関係としての狭義のインターフェース技術、その評価技術の3つの視点から、課題を述べる。

- (1) 個別コンテンツの流通から、相互関係・意味抽出による構造化へ

ネットワーク上には、すでに一人のユーザーが利用しきれない情報が蓄積されている。これらは、情報検索技術や、関連付けなどの技術があって初めて活用が可能となる。特に、映像情報や音声情報などのマルチメディア情報の有効活用は今後の大きな課題となる。基盤となる映像からのオブジェクト抽出や、音声画像の意味理解技術に加え、コンテンツの作成側からは、積極的なメタ情報の付与、利用履歴情報などの活用が、実用的な利用効率を考える上で重要になる。

- (2) 情報入出力から、意味・意図の入出力へ

図2はインテラクションの7段階モデルの説明図である。古くから指摘されているとおり、ユーザーは、心理的な目標の達成に向けて意図を持ち、行動を選択し、物理世界で実行をする。システムは入力によって何らかの状態変化を生じる。すなわち、表示が変わる、運転状態が変わる、音声出力される等である。ユーザーはこれを知覚、解釈し、自分の目標が達成できたかどうかを評価する。

この“意図→行動選択→実行”は実行の淵(ふち)と呼ばれ、大きな認知負荷になり得る。また、ユーザーとシステム設計者の考え方の違いによって正しい行動が選択されない事態が生じる。同様にシステムの状態が変化しても、ユーザーの“知覚→解釈→評価”に負荷や誤りが生じる可能性もある。古くからの指摘ではあるが、意図や意味の抽出は容易な技術ではない。着実な研究開発に加え、操作ではなく、目標や効果を表現する工夫、又は、ユーザーに知覚評

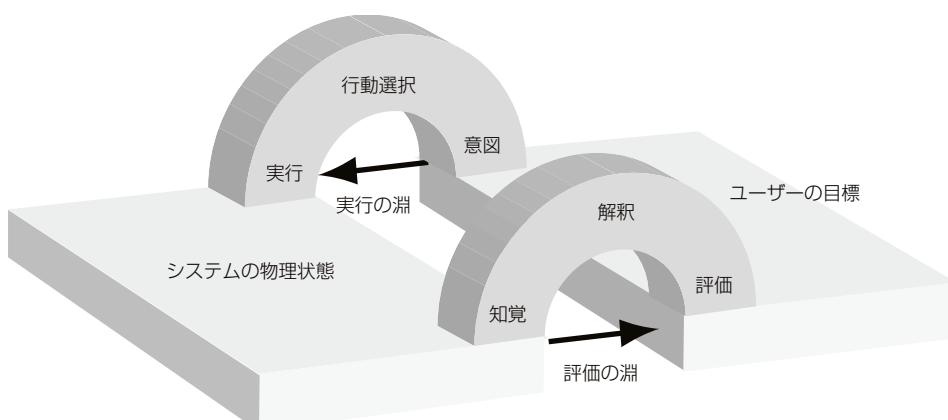


図2. 行為の7段階モデル

価しやすい情報提示を考える必要がある。現在は、メディア処理技術が進み、音声、映像、グラフィックスなどの扱いが容易になっている。これらを用いた直感的に分かりやすい表現、位置情報などのユーザーが明示しない文脈情報の活用も重要な課題である。

(3) 効率からユーザーエクスペリエンスへ

上で述べた心理側面での評価、及びそれに基づくインターフェースデザインが必要である。例えば、効率の評価であれば、タスク達成の時間を測定するなどで定量的な評価も可能であるが、心理的評価は困難で、一般にコストもかかり、限られた側面の評価しかできない場合も多い。しかし、ユーザーが日常の利用の中で、どのように気づき、学び、使えるようになるのか、その過程を把握し、得られる満足に対する評価を行うこと、すなわちユーザーエクスペリエンスの評価とそれに基づくデザインプロセスこそが重要である。評価方法に関する地道な積み重ねと、評価結果を共有し、デザインに活用する仕組みをつくりあげることが必要である。

3. ヒューマンインターフェース技術と当社の取り組み

3.1 インタフェースデザインと評価

当社では、早くから、ユーザービリティ評価に基づくインターフェースデザインを指向し、評価技術の確立を目指した研究開発を実施しながら、検討した方式を用いてユーザービリティワークショップ⁽²⁾を実施し、人間中心のインターフェースデザインを推進している。さらに、情報通信技術の研究開発部門と連携し、プロトタイピング評価や、評価結果をガイドラインとして組み込んだインターフェース設計ツールの開発などを推進し、全社の製品にユーザービリティが高く、共通したコンセプトのデザインを適用している。

3.2 音声入出力技術

音声入出力のもっとも大きな特徴は、意味レベルの直接表現にある。他の入力方式では、目標から行動の選択に淵があり、学習が必要であったり、負荷がかかるのに対して、音声は目標を直接言語表現することができる。当社では、1980年代から音声認識に取り組み、これまでに電話系音声認識やカーナビ用音声認識で評価を得ている。しかし、いまだに利用領域が限られているのは、発声できる内容が限定され、認識でも、音のパターンとしての判定をしているため、ユーザーからは、“目標→発声可能な音声コマンド”的選択→発声”という大きな心理的な淵があること、また、誤認識が生じたときのシステム挙動をユーザーが理解できない場合が生じるためである⁽³⁾。これらの問題を解決するための、意味理解、対話文脈処理、言い方に限定のない音声検索に取り組むとともに、その評価にも注力している。これらの技術によって、利用分野をさらに拡大する予定である。

3.3 グラフィックス技術

グラフィックスでの情報提示は、テキストなどに比べ、情報量が豊かで、利用者の直感的理解が容易である。これは、コンシューマー製品に限らず、プラントの運転制御、監視システム、工業システム、リモコンなどの小型デバイスにも広く求められている。パソコンなどではすでにごくあたり前に扱えるようになっているグラフィックス処理も、工業システムや組み込み機器などのハードウェア資源では、実現が困難である。そこで、小型機器に適用可能なグラフィックスエンジンを開発している。このような小規模システムから、定評のあるデジタルサイネージ、パブリックビューの大型映像システムまでの映像技術に取り組み、あらゆるシステムに高度なグラフィックス情報表現、映像コンテンツを提供している。

また、道路地図や路線図では、ユーザーが知りたい位置関係などに注目し、不要な情報を削除し、実際の図からデフォルメ図を作成する技術など、情報の表現技術にも取り組んでいる。

3.4 画像・映像処理技術

当社は映像符号化技術に高い技術を保有している。この符号化技術を基盤として、画像の同定技術と電子透かし技術⁽⁴⁾等、マルチメディアコンテンツの管理、効率的活用の研究開発に取り組んでいる。さらには、映像情報と各種のセンサ情報を統合し、高精度の三次元地図データベースを作成する取り組みは、莫大な情報の構造化、関連付けの技術として、今後も注力する予定である。

4. 特集号の概要

この特集号に掲載の論文について概要と、上に述べた要求、課題との関連を示す。

最初に、“空調リモコンのHMIデザイン”について述べる。エアコンでは、省エネルギーと快適性の両立が求められている。現在のエアコンには、各種のセンサが搭載され、詳細な制御が可能となっているが、複雑な操作インターフェースではユーザーは快適な運転をすることはできない。そこで、当社では、エアコンのメニューを、機器の操作や機能の表現ではなく、ユーザーの“～したい”という目標と、それによって得られるメリットで表現し、インターフェースを再構築した。ユーザー意図に着目し、高い操作性を実現した設計事例である。

次に“車載機HMIのユーザービリティ評価手法”における、評価方法の基盤研究について述べる。車載機では、安全な運転を確保するために、操作の認知負荷を低減する必要がある。先に述べたように、認知負荷評価は容易ではないが、ここに述べるように、種々の方法を検討しその成果を蓄積している。実際の車載機の開発ではこれらの研究に基づいた評価を行い、インターフェースデザインを実施して

いる。その一例が“入力支援技術”である。少ない操作で目的の施設名が入力できるよう、デザイナーと、情報処理技術者が共同で開発した事例である。

次に、映像情報を含むマルチメディア情報を分かりやすく提示する応用例として“技能伝承システム”について述べる。映像情報の高度な活用によって、技能という形式化しにくい対象を効率的に伝承しようという試みである。

ヒューマンインターフェースを実現するための基盤技術として“音声検索技術”“音声合成技術”について述べる。音声認識で認識可能な話しが限定されるという問題を、音声処理と、言語検索技術を有効に組み合わせることで解決している。また、音声合成は当社の家電における“使いやすさ”的基本コンセプトである“らく楽アシスト”で操作手順を声で知らせる音声ガイダンスに使われている。

“3Dタッチパネル技術”は指の近接を検知できるユニークな技術で、今後の種々の応用が期待される。

視認性が高いグラフィックスを提供するための技術の基盤となる“ベクターグラフィックス技術”と、理解しやすい表現技術である“図形のデフォルメ技術”について述べる。組み込み機器から各種の大規模映像システム、今後の期待が高いデジタルサイネージまで、高い情報表示技術が適用されている。

最後にマルチメディアコンテンツの特徴抽出、構造化技術として“画像同定技術”と“モービルマッピング技術”について述べる。ユーザーにとって大量のデータの構造化、関係性抽出を行うもので、将来のオブジェクトの抽出や、意味抽出に向けて、今後も研究開発に注力する予定である。

5. む す び

ヒューマンインターフェース技術を概説し、当社の取り組みの一部とこの特集号の掲載論文を概観した。大量コンテンツがネットワーク上にあり、クラウド上のサービスが利

用できる今日の情報通信システムでは、機能が製品の価値ではなく、情報・サービスと人をどのように関係づけるか、情報の関係性抽出・構造化によっていかに情報をユーザーの要求に応じて提供できるか、この広い意味でのヒューマンインターフェース技術こそが、製品の価値を決めると考えている。

これを考える上で、卷頭言は興味深い。設計者(先生)の理論で機能(知識)を提供するのではなく、利用者(子供)の要求を満たすピースを提供すれば(自分が納得できるストーリーを作れれば)、それはまさに、ユーザーエクスペリエンスとして(自分の知識として)定着し満足を提供できるものになるのかもしれない。子供におけるロボットの役割は、情報通信システムではヒューマンインターフェースであり、また、将来はそれが、形のあるロボットとして実現されるのかもしれない。

今後も、使いやすさの追求のため、基盤となるメディア処理から、インターフェースデザインとその評価まで一貫した研究開発を継続する。

参 考 文 献

- (1) ノーマン, D.A:誰のためのデザイン? - 認知科学者のデザイン原論, 新曜社認知科学選書 (1990)
- (2) 若松正晴, ほか:三菱電機デザイン研究所におけるユーザビリティ活動, ヒューマンインターフェース学会誌, 4(4), 207~212 (2002)
- (3) 石川 泰: UI設計とユーザビリティ:音声インターフェースの課題(音声認識実用化に向けて), 情報処理学会研究報告. SLP, 音声言語情報処理 2007(103), 35~40 (2007)
- (4) 菅井豊和, ほか:画像データ量を抑えるメタデータ電子透かし方式, 電子情報通信学会総合全国大会講演論文集, 2011年_情報・システム(2), 6 (2011)

“三菱エアコン霧ヶ峰ZWシリーズリモコン”的ヒューマンマシンインターフェースデザイン

小川 孝*
石塚健彦*

Human Machine Interface of "Mitsubishi Room Air Conditioner Kirigamine ZW series"

Takashi Ogawa, Takehiko Ishizuka

要旨

エアコンは家庭内で使われる電力量の約4分の1を占めるとと言われ、その省電力化は家庭の電力使用量低減に大きく貢献する。現在のエアコンは、センシング技術や制御技術の進化によって快適性と省電力運転のためのコントロールが可能になっているものの、機能を理解していないと操作は難しく、最適な運転ができないという問題があった。

そこで、三菱電機では、快適で環境に優しい室内空調機能の見える化と使いやすいHMI(Human Machine Interface)開発に取り組んでいる。2011年モデルでは、リモコンに業界で初めて^(注1)搭載した大型フルドット液晶によるGUI(Graphical User Interface)の導入で、“～したい”気持ちを選ぶ新しい操作方法“快適セレクト”を実現した。

“快適セレクト”では複雑な空調制御を効果でまとめ、自動で選択できるようにするとともに、選択する効果をグラフィックで表現することで理解しやすくする工夫をした。また、GUIによるメニューの体系化でリモコンのボタン数を従来の約1/3にし、ボタンと印刷文字の大型化を実現、さらに、液晶画面でもバックライトと視認性の高い大きな文字を採用することで、高齢者でも見やすい表示とした。エアコン気流やエコ貯金箱の画面でユーザーメリットを見える化し、機能を分かりやすく表現した。これらのHMIデザインによって、エアコン操作のしやすさと分かりやすさを向上させ、省エネルギーと快適性を両立させたユーザーの思い通りの運転が可能となった。

(注1) 2010年8月24日現在。当社調べ



“三菱エアコン霧ヶ峰ZWシリーズリモコン”的HMIデザイン

10素子の“3Dムーブアイ”によって、室内温度、室内湿度に加えて、床温度、天井温度、日射熱、間取り、人の居場所、人の状態をセンシングして学習するルームエアコンのリモコンである。業界で初めて搭載した大型フルドット液晶によるGUIで、使いやすさと分かりやすさを向上させ、ユーザーの思い通りの運転が可能となった。

1. まえがき

エアコンは家庭内で使われる電力量の約4分の1を占めるとされ、エアコンの省電力化は家庭の省エネルギーに大きく貢献する。より効果的な省エネルギーを実現するには、基本性能の向上に加えて外気温や部屋の状況などの使用状況に対応して最適化し、快適性を損なわない範囲での省エネルギー運転を行う必要がある。省エネルギーと快適性を両立させた理想的な運転がエアコンに強く求められている。

現在のエアコンでは、センシング技術や制御技術の進化によって機能が多様化し、省エネルギーと快適性のための細かいコントロールができるようになってきているものの、機能や運転状況を理解していないと操作が難しく、最適な運転ができないという問題があった。

当社では、この問題を解決するため、2008年モデルに業界で初めて^(注2)室内の遠くからでも運転状況を確認できる、大きく分かりやすい設定温度と、空調制御エリアを光で示す表示を室内機に搭載した。2010年モデルでは、室内機表示にエコレベル、使用電気代やCO₂排出量を追加するとともに、センシング情報から様々なエコアドバイスを液晶画面に表示するリモコンを開発し、機能の見える化と使いやすいHMIによって、快適で環境に優しい室内空調の実現に取り組んできた⁽¹⁾。

2011年モデルでは、リモコンに業界で初めて大型フルドット液晶を搭載することで、GUIによる使いやすさと分かりやすさを向上させたHMIを開発し、2010年10月に発売した。

本稿では、先に述べたGUIを搭載した“ZWシリーズ”リモコンのHMIデザインについて述べる。

(注2) 2007年10月10日現在。当社調べ

2. ZWシリーズリモコンのHMIデザイン

2.1 コンセプトの策定

ZWシリーズ2011年モデルのリモコンの開発では、高機能と操作のしやすさ・分かりやすさの両立をコンセプトにデザイン開発を進めた。開発にあたっては、フルドット液晶画面内の表現や操作フローだけでなく、操作ボタン数を削減するための方策を検討して、従来のデザインにとらわれない新形態のHMIの実現を目指した。

開発開始時に策定したHMIコンセプトを次に示す。

- (1) 複雑な制御の操作を簡単に
- (2) リモコンのボタンを減らして見やすく・使いやすく
- (3) メリットの見える化で分かりやすく

表示部は従来機種と比較し、フルドット液晶(解像度140×255ドット)によるGUIで大幅に表示能力が高められ

ている。フルドット液晶の採用は、従来のセグメント方式の液晶と違い、情報の表示量に制約が生じないため、多くの機能を分かりやすく表現できるメリットがある。

基本的な操作に関しては、通常、高機能化に伴い操作が複雑になる傾向にある。そこで、操作方法を整理することで、高機能化と、操作性の両立を実現している。すなわち、GUIによるメニュー項目は、設定内容別にグループ分けを行い、その下位階層は上位階層と関係する機能のみとしている。これによって、グループ内の項目は上位階層の名称から設定できる機能がイメージできるようにした。各項目は、階層が進むにつれて細かな情報を案内する表示構造とし、ユーザーのメンタルモデルと合わせた構成とすることで、誤解や混同なく操作が行えるようにした(図1)。

また、ボタンに関しては、“運転 入／切”“冷房”“暖房”“除湿”的基本機能をリモコン本体のボタンに割り当て、GUIとの役割を整理した。

表示される画面は、操作する機能に合わせて最適な内容を表示する。メニュー内の設定画面では、設定項目の機能の説明など理解を助ける情報を表示し、運転状態を案内する画面では運転内容と温度・湿度を大きく表示する(図2)。このように、操作に応じて分かりやすい最適な情報を表示するように工夫している。また、メニュー内の操作画面は、機能ごとの画面表現を共通化することで一貫性を持たせた。各画面での操作の一貫性を確保することは、操作方法の覚えやすさにつながる。

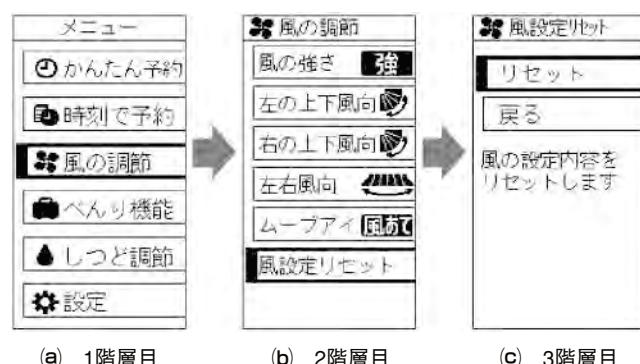


図1. 階層が進むにつれて細かな情報を案内



図2. 運転内容と温度・湿度が大きく表示される簡易表示画面

2.2 コンセプトの具体化

次に、3つのコンセプトの具体化について述べる。

2.2.1 複雑な制御の操作を簡単に

従来、ユーザーが好みの空調を行うためには、センサや風向の設定自分で組み合わせて行う必要があった。そのため、ユーザーはエアコンの機能をすべて理解していないと求める空調が実現しにくいという課題があった。そこで、ユーザーがエアコンを運転する目的に着目し、“～したい”気持ちを選ぶ新しい操作方法“快適セレクト”を創出した。

これは、運転メニューを目的行為別に整理し、操作を一本化することで、複雑な空調制御を目的に合わせて自動で選択できるようにしたものである(図3)。

選択項目は、従来の“ミスト入／切設定”のような機能を示した名称や“風速切替え”のような制御方法を示した名称から、“足元をあたためたい”のような“～したい”という名称とし、目的を示す“気持ち”を選択する操作方法とした(図4)。

これによって機能を理解していなくてもユーザーニーズに合わせた機能の選択が容易となり、ユーザーの気持ちにこたえる最適な運転が可能になった。

また、各々の空調制御の効果をグラフィックで表現することで理解しやすくする工夫をした。選択された項目の動

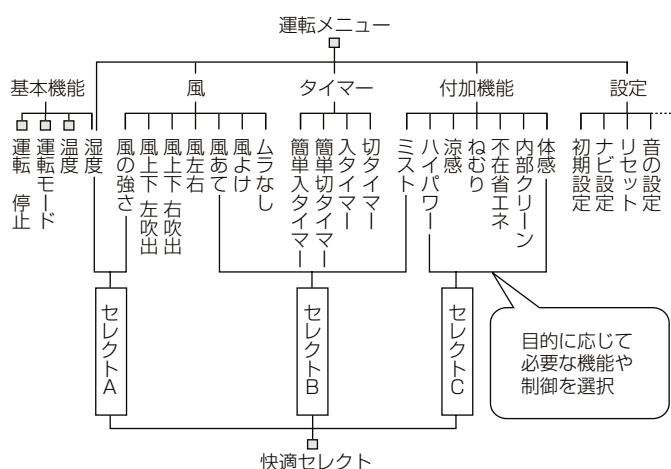


図3. 目的行為別に運転メニューを整理



図4. “～したい” 気持ちを選ぶ“快適セレクト”画面

作をエアコンと人物のイラストを使ったグラフィックで画面下段に表示することで、操作を決定する前にエアコンの運転結果をイメージできるよう工夫をした(図5)。

2.2.2 リモコンのボタンを減らして見やすく・使いやすく

使いやすいリモコンHMI実現には、リモコン本体のボタンレイアウトも重要な要素である。そこで、リモコンに対する難しい印象を和らげつつ、シンプルで分かりやすい操作性を念頭にリモコンボタンのレイアウトを検討した。

操作体系を先に述べた“快適セレクト”，各種設定の入り口である“メニュー”，エアコン情報や電気代情報確認などの入り口である“おしさせナビゲーション”的3つに整理し、操作の入り口を明確にした。詳細機能をGUIメニューに集約することで、ボタン数を従来リモコンの約1/3にし、分かりやすい印象にするとともにボタン形状と印刷文字の大形化で薄暗い場所でも見やすいボタンを実現することができた(図6)。

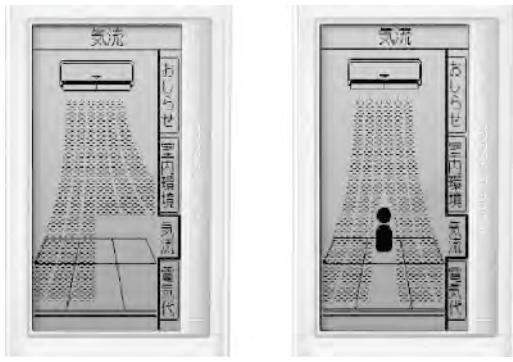
また、シンプルな操作を実現するため、カーソル操作は上下方向に限定し、複雑なボタン操作が苦手な高齢者に配慮した。その際にGUI画面に左右方向に動かしたくなる画面レイアウトを作らないようにし、リモコン側のボタンとGUIとの整合性を取った。



図5. 空調制御の効果をグラフィックで表現



図6. 従来リモコンボタンとボタン数を低減した新リモコン



(a) 異なる左右風向の運転時 (b) 風よけ運転時

図7. 気流制御の見える化



図8. 節約できた電気代を表示するエコ貯金箱画面

一方、液晶画面にも視認性の高い大きな文字を採用し、バックライトの採用と合わせて高齢者にも見やすい表示とした。また、機能をグラフィックと説明文でしっかり説明することで、従来のボタン表示だけでは分かりにくかった、機能に対する説明を充実することができた。

2.2.3 メリットの見える化で分かりやすく

従来の運転状態の確認画面は、主に設定した機能が何であるかを確認することを目的とした画面であった。そのため、複数の設定表示を見て、全体像をイメージする必要があり、直感的に理解しにくい課題があった。そこで、エアコン気流の確認画面を新たに用意し、今まで見ることができなかった運転中の風の状態を見る化することで、直感的に運転状態を理解できるように工夫した(図7)。

エアコン気流の確認画面は、上下左右風向、風速、人の風あて風よけをアニメーションで表現している。さらに、風速はドット密度の調節によって濃淡で表現し、上下風向は動作するフラップの段階数に合わせて長さと形状で表現することで、運転中の気流が直感的に理解できるよう工夫をした。

その他、節約できた電気代を表示する“エコ貯金箱”画面を設けた。エコ貯金箱画面では、楽しく省エネルギーに取り組めることをテーマに、節約できた電気代を単なる数字で表示するだけでなく、貯金箱を模した豚のキャラクターが節約した電気代をアニメーションで表示し、節約金額に応じてストーリーが変化する仕様としている(図8)。

3. むすび

リモコンはエアコンとユーザーとの接点として重要な役割を担っている。今後は、よりユーザーの気持ちにこたえた快適な空調の提供と、より一層のCO₂排出量削減及び震災の影響による節電へ貢献等、より効果的な省エネルギー運転が求められると考える。これらの実現に向け、引き続きユーザー視点の商品コンセプト発想によるHMIデザイン技術で、高度な機能がだれにでも使いやすい魅力的な商品の創出に取り組む所存である。

参考文献

- (1) 山崎友賀, ほか: 家電製品のユニバーサルデザイン, 三菱電機技報, 83, No.12, 750~752 (2009)

車載機ヒューマンマシンインタフェースのユーザビリティ評価手法

若松正晴*
嶋田 淳*
沢田久美子*

Usability Measurement Method for HMI of Automotive Infotainment Systems

Masaharu Wakamatsu, Atsushi Shimada, Kumiko Sawada

要 旨

カーナビゲーション(以下“カーナビ”という。)など車載機のヒューマンマシンインタフェース(Human Machine Interface : HMI)開発では、音楽再生など、運転中に行う機器操作が、ドライバーに大きな負荷を与えないように配慮することが重要である。

しかし、この機器操作時のドライバーの負荷を定量的に測定することは容易ではない。そこで、機器操作における認知的負荷に焦点を絞り、運転中の機器操作状況を再現し、機器開発のプロセスの中で容易にかつ、正確に測定できる指標を得ることを目標とし、評価手法の研究を行ってきた。

提案する評価方法は、ドライビングシミュレータ(以下“DS”という。)を用い、運転操作中に車載機を操作させながら他の刺激を呈示し、その刺激に対する反応を計測することで、車載機HMIの操作負荷を推定する方法である。

実験では、被験者に、光点刺激、振動刺激、音刺激を呈示して、刺激を知覚すると同時に反応ボタンを押下させる。機器操作の負荷が高い場合、刺激の知覚と反応に影響があると推測される。実験結果からは、機器操作の難易度と刺激反応パフォーマンス(反応時間、エラー率)に明らかな関連があることが確認できた。

また、同時に検討した機器操作難易度と走行時車線逸脱指標との関連も確認した。

脳波など生理計測結果との明確な関連性は認められなかったものの、この結果は、提案する方法が、機器操作の認知的負荷の評価に有効であることを示している。

今回の成果を踏まえ、測定手法の充実化を図り、今後、安全運転を阻害しない、認知的負荷の低い有効な車載機HMIの開発に活用する予定である。



DSを利用した評価実験

写真は、被験者が、画面に表示されている前方走行車両に追従しながら、左手でカーナビのリモコンを操作し、楽曲の検索・再生課題を行っている様子を示している。この課題と同時に呈示される画面上の光点刺激(画面上下に見える4本のケーブルの先端が発光部)、首筋に貼り付けた振動子の振動刺激、足元のスピーカーからの音刺激を知覚したら、右手親指付近のステアリングホイール上に固定した反応ボタンを押下する。これらの刺激への反応から機器の認知的負荷を測定する。

1. まえがき

自動車は、情報通信技術を導入しながら移動するマルチメディア空間として進化してきた。例えば、オーディオ、テレビ、ビデオ、電話、インターネット、ETC(Electric Toll Collection)、カーナビ等、様々な機能やサービスを車の中で利用することが可能となり、今後もそれらの高機能化、多機能化が加速していくものと思われる。

これらの機能やサービスを実現するための手段が車載機であり、特にドライバーがこれをHMIを介して利用する際には、高機能化や多機能化に伴う複雑な機器操作(人と機器とのインタラクション)が、メインタスクである運転操作に対して、マイナスの影響を及ぼさないように、十分に配慮することが必要不可欠である。

車載機HMIのユーザビリティを高めるためには、人間中心設計のアプローチ⁽¹⁾が重要であるが、特に、安全運転を阻害しないように、操作負荷を低く抑えることが求められる。操作負荷には、操作部の寸法、配置等に伴う身体的負荷、及び操作方法、手順等に伴う認知的負荷の二つが考えられる。今回は、後者の認知的負荷の定量的評価手法について検討してきた研究について述べる。

2. 認知的負荷の定量化

2.1 研究の目的

DS運転操作中に車載機(現行製品や開発中プロトタイプ)の機器操作を行った場合、どの程度の認知的負荷がかかっているのかを定量的に示すことができれば、HMI開発の目標設定がしやすくなる。

この研究は、開発現場で効率的に活用できる操作負荷定量化の評価指標を抽出することを目的とした。

2.2 注意資源

機器操作における認知的負荷の程度は注意資源要求量としてとらえることができる。すなわち、機器操作に多くの注意資源を要する場合、認知的負荷が高いと解釈できる。運転操作と機器操作に割かれる注意資源の量的関係を図1に示す。通常運転操作に必要な注意資源は一定量と考え、これに機器操作に必要な注意資源が加算される。注意資源総量には許容上限値が存在するが、機器操作が平易である

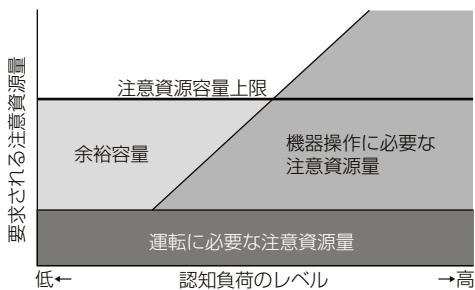


図1. 注意資源量の概念図

場合は、許容上限値までの余裕容量が十分に残っており、運転操作も機器操作も確実に実行できる。機器操作が複雑になってくると認知的負荷が増え、注意資源要求量が増大して余裕容量が減少、不足し、運転操作や機器操作のパフォーマンスに影響を及ぼすことになる⁽²⁾。

2.3 先行研究

2節の注意資源の余裕容量を定量的に推定するために、先行研究では様々な手法が検討されてきた。

主な指標としては、運転パフォーマンスとしての車線横変異量⁽³⁾⁽⁴⁾、運転操作と機器操作の課題に付加した副次課題(ランダムな刺激への反応課題)の反応時間⁽⁴⁾⁽⁵⁾⁽⁶⁾や正答率⁽⁴⁾、機器操作パフォーマンスとしての操作回数や操作時間⁽³⁾、機器操作時の機器画面注視時間や注視回数⁽⁷⁾、機器操作時の事象関連電位P300(刺激呈示後300ms前後に出現する脳波)⁽⁴⁾⁽⁸⁾、運転操作と機器操作実施後の主観評価(NASA(National Aeronautics and Space Administration)-TLX(Task Load indeX))⁽⁶⁾⁽⁹⁾⁽¹⁰⁾等がある。

しかし、これらの手法には、運転パフォーマンスに影響が現われにくい低い認知的負荷の評価が困難、特殊な機器が必要、多数の被験者が必要、等の問題があった。

3. 検討手法

3.1 指標

DSでの運転操作を前提として、次の指標を検討した。

3.1.1 刺激反応時間

認知的負荷の程度は、視覚をはじめ様々な知覚反応のパフォーマンスに現れるであろうという仮説に基づき、次の6つの刺激(光点4、振動1、音1)の一つを、1~4sごとにランダムな順序で呈示する。なお、先行研究⁽⁵⁾では光点設置位置はダッシュボード左側の1か所であるが、この研究ではより自然な前方視野内4か所とした。被験者は、各刺激を知覚したら、ステアリングホイールに設置した反応ボタンをすみやかに押下する。刺激の呈示からボタン押下までの時間を反応時間として計測記録する。

(1) 光点刺激

DSの前方視野を模擬するディスプレイ画面(幅550mm、高さ320mm)上の4か所(画面中心から上下各110mmの水平線上、中心から左右180mm位置)に設置した赤色LED(Light Emitting Diode)(発光部直径5mm、発光輝度5cd/m²)を0.3s発光させる。なお、画面までの視距離は900mmである。

(2) 振動刺激

振動子(振動量1.0G、周波数50~60Hz)を被験者の右側首筋にテープで固定し、0.3s振動させる。

(3) 音刺激

ホワイトノイズを、DS走行音(A特性80dB、B特性65dB)下で十分識別知覚できる音量で、被験者前方足元のスピーカーから発生させる。

カーから0.3s出力する。

3.1.2 刺激反応エラー率

3.1.1項の(1)～(3)の刺激呈示について、全刺激数に対する無反応(ボタン押下のない)刺激数の比率をエラー率として求める。

3.1.3 生理計測

予備実験で、心拍(胸部、腹部の計3か所に電極を貼り付けて計測)、呼吸(腹部にベルト型センサを装着して計測)、皮膚電位(左手第2指、第3指の2か所に電極を貼り付けて計測)、脳波(P300の振幅を見るために計測)の各指標を検討した。

3.1.4 DS走行ログ

DS運転操作による仮想上の自車走行ログ(速度、アクセル、ブレーキ、操舵(そうだ)角、車線位置等)を記録した。

3.1.5 主観評価

実験課題終了後に、被験者の主観的印象をNASA-TLXをベースとした質問紙で定量化し記録した。

3.2 実験方法

3.2.1 被験者

運転免許を持っている成人男女12名(予備実験は26名)で実施した。

3.2.2 実験設備

模擬的運転操作を行うために、三菱プレシジョン製のDSを使用した。

DSダッシュボード左側にはA社製(予備実験はB社製)カーナビ本体を設置し、DSのステアリングホイール左側にリモコンを取り付けた。

また、課題実施中の被験者の視線を記録するために、DSのダッシュボード手前にアイトラッカー(Tobii X50)を設置した。

3.2.3 実験課題

(1) 運転操作課題

被験者は、DSに着座し、ゆるやかなS字カーブを繰り返す高速道路の左側走行車線を時速100kmで走行する仮想車両に追従しながら模擬走行する。

(2) カーナビの機器操作課題

運転操作と同時に、カーナビのハードディスクに保存された音楽の中から、実験者から口頭でランダムに指示される楽曲を検索・再生する。カーナビ操作は、リモコン操作

と音声操作を指示された順番で実行する。

(3) 携帯電話の機器操作課題

運転操作と同時に、携帯電話を左手で保持し、実験者から短い文節ごとに順次口頭で指示される文章を、メール本文として入力する。

(4) 小銭数え課題(6名のみ実施)

DSの運転席左側に設置したトレイから実験者に指示された金額の小銭を選び、運転席右側のトレイに移す。

(5) 刺激反応課題(副次課題)

(1)に加えて、(2)～(4)のいずれかを同時に実行するとともに、ランダムに呈示される刺激を知覚したら反応ボタンをすみやかに押下する(3.1.1項参照)。各実験実施時間は、15分間とした(予備実験は30分間とし、(3)と(4)は未実施)。

3.3 実験結果

3.3.1 刺激反応課題の結果

視点がDS画面上にない場合は、光点刺激を知覚できないので、無反応となる。これについては、視点のずれによるパフォーマンス低下も負荷の指標にはなるという考え方と、認知的負荷を調べるのであれば、物理的に見えない(視点が画面上にない)場合は排除し、本来、見えるはずの条件のみを負荷の指標とすべきとする考え方がある。

後者に基づき、視点ずれを除いて集計した結果が図2と図3である。DRはDS運転操作のみ、VOはDS運転+カーナビ音声操作、REはDS運転+カーナビリモコン操作、CEはDS運転+携帯電話操作、の各条件を示す。反応時間では、振動と音に比べ光点刺激が短い傾向にある。エラー率は、音声操作に比べリモコン操作と携帯電話操作が高い。

3.3.2 車線逸脱の結果

DS運転操作による走行ログから、実験条件ごとの車線逸脱時間(自車が車線中央位置から左右に25cm以上ずれていた時間)平均値を図4に示す。音声操作に比べリモコン操作は逸脱時間が明らかに増大する。

3.3.3 生理計測の結果

予備実験で計測した心拍、呼吸、皮膚電位の分析からは、カーナビ操作による覚醒水準上昇の傾向は認められたが、個人ごとのばらつきが大きい。

脳波は被験者6名のみで計測した。P300の分析では、データに含まれるノイズが多く、明確な結果は得られていないが、刺激反応時間に対応する傾向が認められた。

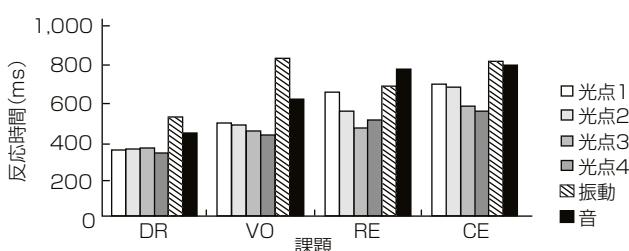


図2. 平均反応時間

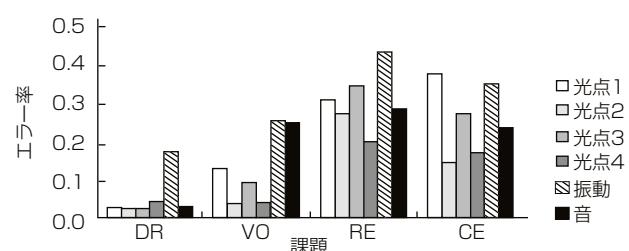


図3. エラー率

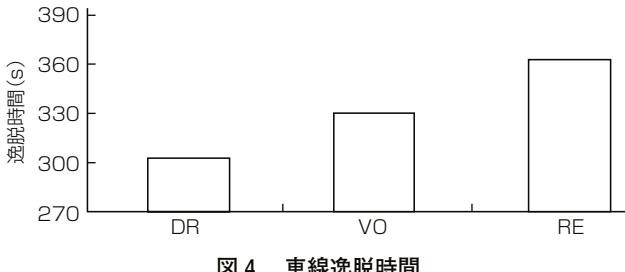


図4. 車線逸脱時間

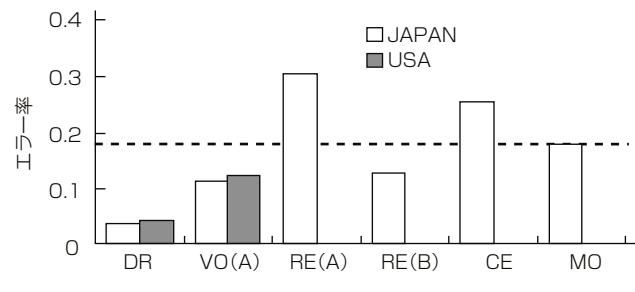


図5. 刺激反応エラー率

4. 考 察

刺激反応の時間及びエラー率は、運転のみの場合に比べ、カーナビ操作などの副次課題を加えた場合、明らかに増大傾向が認められた。すなわち、刺激反応課題の指標は、運転操作中の機器操作における認知的負荷を推定する尺度として有効であることが確認できた。

図5に、この実験(DR, VO(A=A社製カーナビ), RE(A), CE, MO=小銭数え課題)及び予備実験(RE(B=B社製カーナビ))のエラー率を重ね合わせたグラフを示す。なお、凡例のUSAは、北米で実施した追加実験(被験者4名)の結果である。A社製カーナビはリモコン操作時ナビゲーション画面視認が必要であるが、B社製はリモコン操作時に音声ガイドがあり画面視認は不要である。

小銭数え課題を基準に、法規上は禁止されている携帯電話文字入力操作など画面を視認しながら行う機器操作のエラー率を見ると、後者は明らかに率が高い。一方、カーナビ音声操作や画面視認の不要なりモコン操作では、小銭数え課題よりもエラー率は低い。小銭数え課題以外に、例えば、エアコン温度調節操作やラジオ選局操作等、運転中に通常行う操作の実験データを蓄積してエラー率の許容値を導出することも可能であろう。

DS走行ログについても、機器操作に伴う認知的負荷の増大に対して感度があり、尺度として有効であることが確認できた。今回着目した車線逸脱データと刺激反応データとの関係を更に検討していくことが課題である。

生理計測については、認知的負荷の指標としての明確な効果を確認することはできなかった。個人差が大きいこと、計測データにノイズが含まれること、計測時の被験者の負担が大きいこと等を解決することが課題である。

5. む す び

この研究では、認知的負荷を推定する指標として次のことを確認することができた。

- (1) 刺激反応時間、刺激反応エラー率は有効である。
- (2) DS走行の車線逸脱時間は有効である。
- (3) 生理計測は、可能性を確認したが、4章で述べた課題解決が必要である。

今後、これらを踏まえて、より良い製品開発に向けて評

価手法の改良に取り組んでいきたい。

今回の研究は、大阪大学人間科学研究所の篠原一光准教授、大阪工業大学工学研究科の大須賀美恵子教授それぞれとの共同研究成果をまとめたものである。

参 考 文 献

- (1) 若松正晴：車載機器のインターフェースデザイン、自動車技術, 59, No. 5, 96~101 (2005)
- (2) 篠原一光, ほか：マルチモーダル検出課題によるドライバーの認知的負荷評価に関する研究、自動車技術会学術講演会前刷集, 20105601 No. 104-10, 13~16 (2010)
- (3) 美記陽之介：音声操作が運転に及ぼす影響評価、シンポジウム ケータイ・カーナビの利用性と人間工学 2003, 59~60 (2003)
- (4) 江部和俊, ほか：ドライバーの視聴覚認知に伴う負担度評価、豊田中央研究所R&Dレビュー, 34, No. 3, 55~62 (1999)
- (5) Merat, N., et al.: The Effect of Stimulus Modality on Signal Detection, Feb. 2008-Human Factors, 145~158 (2008)
- (6) 萩原 亨, ほか：携帯電話の操作と位置がドライバーのメンタルワークロードに与える影響について、シンポジウム モバイル2007, 41~44 (2007)
- (7) 田 容旭, ほか：外国人のモビリティ向上のための車内経路誘導情報に関する研究、シンポジウム モバイル2007, 37~40 (2007)
- (8) 鎌倉快之, ほか：眼球停留関連電位のP3潜時を用いた認知負荷評価に向けて-ICAによるアーティファクト除去の適用-, ヒューマンインタフェース学会論文誌, 10, No. 2, 95~102 (2008)
- (9) 木村貴彦, ほか：運転時の心的負担課題としてのn-back課題の妥当性、シンポジウム モバイル2007, 121~122 (2007)
- (10) 三宅晋司, ほか：メンタルワークロードの主観的評価法-NASA-TLXとSWATの紹介および簡易法の提案、人間工学, 29, No. 6, 399~408 (1993)

カーナビゲーション向け入力支援技術

Input Support Technology for Car Navigation System

Atsushi Matsumoto, Takeyuki Aikawa, Masato Hirai, Wataru Yamazaki

松本 篤* 山崎 航*
相川勇之** 平井正人***

要旨

従来のカーナビゲーションにおける施設名称検索では、50音パネル上の操作で施設名称を1文字ずつ正確に入力しなくてはならず、ユーザーの利便性が低かった。効率的な文字入力の手段として携帯電話に見られる予測入力があるが、カーナビゲーションにおける施設名称検索に予測入力を適用した場合、携帯電話のような過去の入力履歴や付属語に限定した単語予測では効率的な入力が行えない。また、1文字入力するたびに予測候補を提示するためには、大規模な施設データに対する高速な予測候補検索が必要となる。

そこで三菱電機では、これらの課題を解決する入力支援技術として次の2点の開発を行い、カーナビゲーションの施設検索向けに利便性の高い入力方式を実現した。

(1) 施設名の入力に特化した階層入力方式

単語間の連接情報を用いた次単語予測によって、目的の

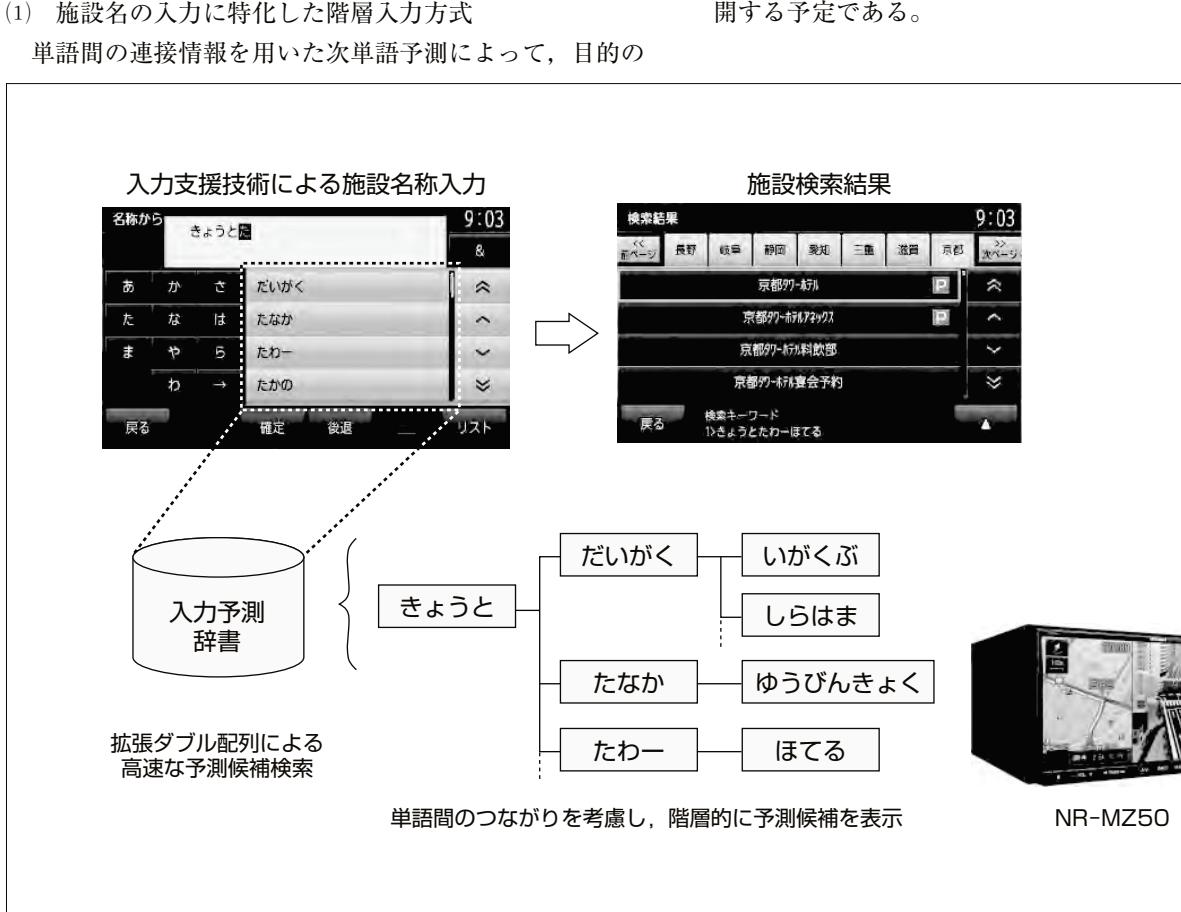
施設名を選択操作によって入力可能とし、タッチ操作数を削減して操作性を向上させた。

(2) 独自の索引方式による高速な予測候補検索

既存の検索アルゴリズムであるダブル配列方式を拡張し、予測検索用の補助情報を埋め込むことによって、大規模な施設データに対し高速な予測候補検索を実現した。

また、操作画面では、携帯電話で使い慣れているテンキー入力を採用し、各ボタンサイズを大きくするなど操作性を考慮した構成とし、入力支援技術と併せて施設名称検索の使い勝手の向上を図っている。

この技術は2011年8月に発売した市販のカーナビゲーション“NR-MZ50”に搭載しており、今後ほかの機種にも展開する予定である。



カーナビゲーション“NR-MZ50”への入力支援技術の適用

施設名称検索の名称入力で、駅や観光スポットなどよく入力される施設を中心に予測候補を高速表示する技術を搭載している。また、単語間のつながりを考慮した階層入力方式によって少ないタッチ数で施設名を入力できるようにした。さらに、文字入力ボタンは携帯電話で使い慣れているテンキー入力方式を採用し、予測候補を表示しつつも省スペースで操作しやすいユーザーインターフェースを実現している。

1. まえがき

従来のカーナビゲーションにおける施設名称検索では、50音パネル上の操作で施設名称を1文字ずつ正確に入力しなくてはならず、ユーザーの利便性が低かった。効率的な文字入力の手段として携帯電話に見られる予測入力があるが、カーナビゲーションにおける施設名称検索に予測入力を適用した場合、携帯電話のような過去の入力履歴や付属語に限定した単語予測では効率的な入力が行えない。また、1文字入力するたびに予測候補を提示するためには、大規模な施設データに対する高速な予測候補検索が必要となる。

そこで当社は、これらの課題を解決する入力支援技術として次の2点の開発を行い、施設検索向けに利便性の高い入力方式を実現した。

- (1) 施設名の入力に特化した階層入力方式
- (2) 独自の索引方式による高速な予測候補検索

また、画面構成では操作性を考慮したデザインを採用し、入力支援技術と併せて施設名称検索の使い勝手の向上を図っている。

本稿では、開発した入力支援技術及び当社市販カーナビゲーションに搭載した際のユーザーインターフェースについて述べる。

2. 入力支援技術

2.1 施設名の入力に特化した階層入力方式

携帯電話で使用されている予測変換機能では、ある候補を選択確定した直後に表示される候補は、過去の入力履歴や付属語に限定されている。初めて行く場所を入力する機会が多いカーナビゲーションでは、入力履歴を用いた予測だけでは不十分である。また、付属語を含む施設名の数は少ないため、付属語の予測によるユーザーメリットも小さい。

当社では、最初の単語を確定した後に、その単語につながる次単語候補が順次選択可能となる階層入力方式を開発した。この方式では、辞書中の全単語間の連接情報を用いた次単語予測によって、選択操作のみで入力可能な施設名を増やすことができ、操作性が向上する(図1)。

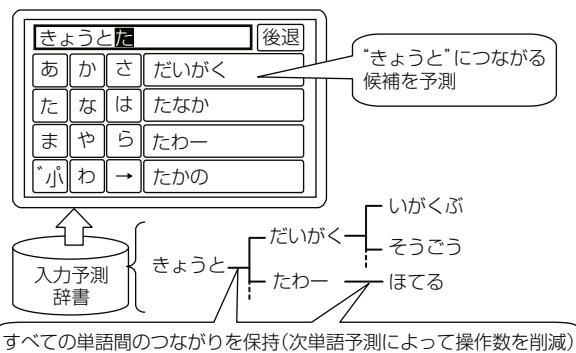


図1. 単語間の連接情報を用いた階層入力方式

階層入力方式による操作性改善を評価するため、施設名称入力にかかるタッチ操作数を従来の方式と比較した。従来方式では、施設名称の先頭から順次読みをすべて入力した場合の必要なタッチ回数を求めた。階層入力方式では、各単語の読みを先頭から順次入力し、該当単語が候補リストに出現した時点でその単語を選択して施設名称入力をを行い、読み入力と選択にかかるタッチ操作回数を加算して求めた。

ランダム抽出した約1万件の施設名を用いて評価した結果、階層入力方式では従来方式と比べてタッチ操作回数を約4割削減できることがわかった。

2.2 独自の索引方式による高速な予測候補検索

予測候補検索では文字入力するたびに提示候補を取得しなくてはならず、快適な操作性を保つためには検索エンジンの高速性が要求される。当社の予測候補検索エンジンでは、既存の検索アルゴリズムであるダブル配列方式⁽¹⁾に独自の拡張を加えることで、リソースに制限のあるカーナビゲーション上での大語彙辞書の高速な予測候補検索を可能とした。

2.2.1 従来方式

ダブル配列方式とは、木構造をBASE配列とCHECK配列という2つの配列として表現する、高速性を特長とする索引方式である。BASE配列には、木構造中の各文字の情報を格納する。文字はあらかじめ数値化され、BASE配列中の基準エントリからのオフセット値として表現される。CHECK配列には、木構造のノード間のリンク情報を格納する。照合処理の際には、入力文字列の最初から順番に各文字に対応するBASE配列の値とCHECK配列の値を参照し、これらの値の整合性チェックを繰り返すことで入力文字列が索引データ中に存在するかどうかを判定できる。

この方式では、単純な辞書検索については非常に高速である。しかし、入力された文字列に続くすべての候補を列举する予測候補検索では、図2に示すように木の途中からすべての枝をたどりながら、“みつびし”につながるすべての子ノードの検索をしなくてはならず速度性能が低下する。

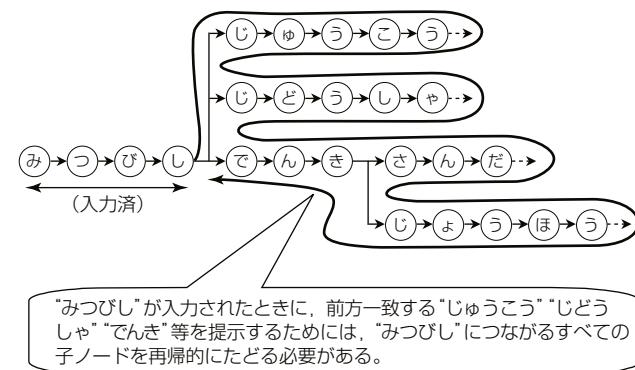


図2. ダブル配列方式(従来)

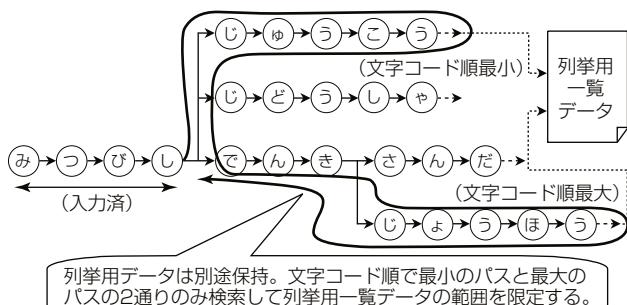


図3. 拡張ダブル配列方式

2.2.2 拡張ダブル配列方式

開発した拡張ダブル配列方式では、次に述べる索引形式の改良によって、前方一致のワイルドカード検索による高速な候補の列挙を可能とした⁽²⁾。

まず、従来方式のBASE配列とCHECK配列に加え、補助情報配列中に子ノードに対応する各文字のうち文字コード順で最小となる文字と最大となる文字をそれぞれ保持することで、木の探索を最小パスと最大パスの2回にとどめる。これだけでは途中パスに対応する情報を検索できないので、列挙用一覧データを別途保持し、木構造の末端にそのデータへのポインタを埋め込むことで、容易に全候補を列挙できるようにした(図3)。

この拡張によって、検索速度を従来の約5倍にすることことができた。これに伴い索引データサイズは従来の約1.5倍となるが、従来方式とこの方式で共通に必要な索引以外のデータが支配的なため、辞書データ全体では約2割弱の増加にとどまっている。

3. 市販カーナビゲーション“NR-MZ50”への適用

入力支援技術を製品に搭載するにあたり、視認性と操作性を考慮し、カーナビゲーション向けに最適なユーザーインターフェースを新たに開発した。この章では市販カーナビゲーション“NR-MZ50”に適用した際の画面構成と操作について述べる。

3.1 ユーザービリティに配慮した画面構成

従来の施設名称入力画面は50音入力タイプを採用し、画面全体を使って文字入力をしていた(図4)。今回、入力支援技術の搭載によって、文字入力画面を新たにデザインした(図5)。

画面の構成要素としては、①入力文字列表示エリア、②文字入力ボタン、③予測候補リスト、④ファンクションボタンの4つに分けられる。

①入力文字列表示エリア

従来と同じように、画面上部に配置した。今回、複数のキーワードが設定可能なアンド検索に対応することから入力文字数が多くなることを考慮し、2行表示ができるスペースを用意した。



図4. 従来の施設名称入力画面

①入力文字列表示エリア ③予測候補リスト

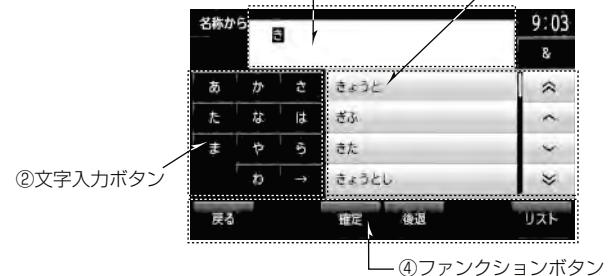


図5. NR-MZ50の施設名称入力画面

②文字入力ボタン

新たに加わった予測候補リストエリアを確保するためにも、携帯電話で使い慣れているテンキー入力タイプを採用し省スペース化を図った。また、ボタンサイズを大きくし、押しやすさを向上させた。

③予測候補リスト

予測候補を見やすく表示するため大きく配置した。また、予測候補リストは1文字入力するごとに表示内容が変化する。そこで、表示内容が変化しない文字入力ボタンやファンクションボタンと区別し、白色を採用した。

④ファンクションボタン

他の画面との整合性をとり画面下部に並べている。

全体の配置としては、画面左側から右側へ、操作が流れるように画面を構成している。左側で文字入力し、右側で予測候補を選択し、最後に画面右下の“リスト”ボタンを押下して入力施設名による検索を行う。

3.2 施設名称検索操作

3.2.1 施設名称検索操作フロー

施設名称検索の操作フローを、“きょうとたわーほてる”を検索する場合を例にして述べる(図6)。

- (1) 文字入力ボタンの操作で“き”を入力し、予測候補の“きょうと”を選択する。
- (2) “きょうと”に続く次単語予測候補が表示される。“きょうと”で始まる施設名は多数あるが、続けて“た”を入力すれば、“きょうと”に続く“た”で始まる次単語予測候補が絞り込まれる。
- (3) 所望の“たわー”選択後、“きょうとたわー”に続く次

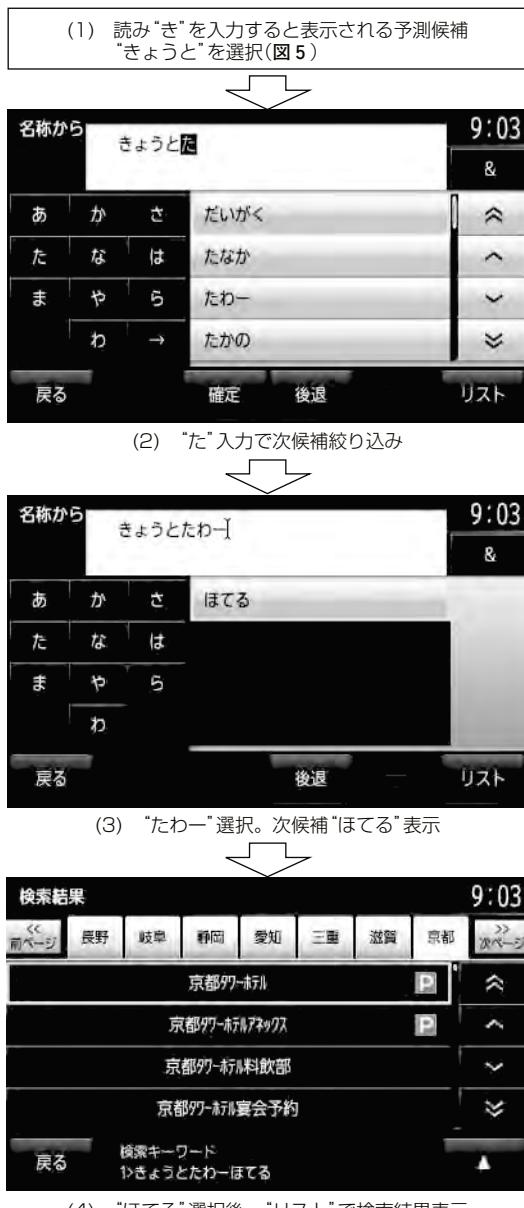


図6. 施設名称検索操作フロー

単語予測候補“ほてる”が表示される。

- (4) “きょうとたわーほてる”的入力が完了した後、ファンクションボタンの“リスト”を選択すると、“きょうとたわーほてる”的検索結果が表示される。

施設名称入力後の施設検索処理には、音声認識による施設検索向けに当社が独自開発したあいまい検索技術⁽³⁾を採用している。あいまい検索では、入力文字列に単語の脱落や語順の入替えがあった場合でも目的の施設の検索が可能であり、入力支援技術と併せて施設検索の利便性を向上させている。

3.2.2 誤入力時の再入力

一般的な携帯電話の入力予測機能は、誤って文字を確定

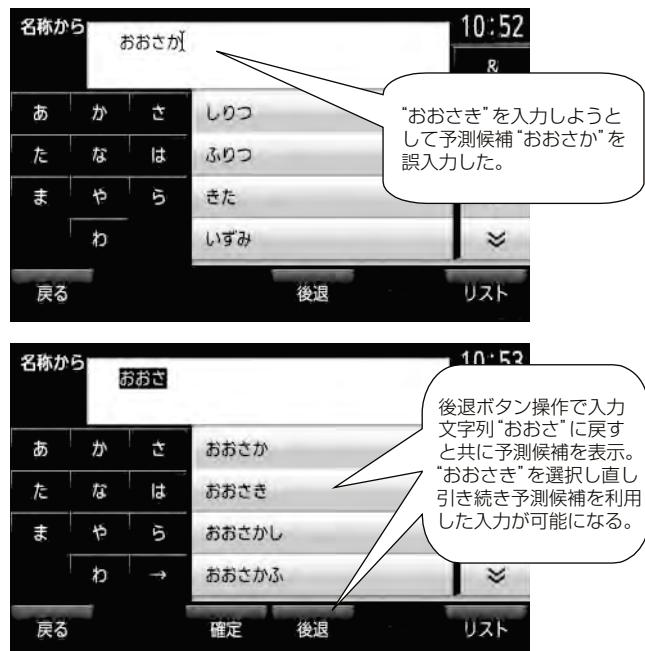


図7. 誤入力時の再入力

してしまった場合、誤入力した文字を削除し、あらためて最初から入力し直す必要がある。今回の入力予測機能では、削除中の文字列に対しても予測候補を提示するようにしたため、誤入力時でも予測候補を活用しながら少ない操作で入力をやり直せるようにしている(図7)。

4. むすび

施設名称検索の利便性向上を目的として、入力支援技術の開発を行った。次単語予測による階層的な入力方式によって操作性向上を果たし、当社独自の拡張ダブル配列によって大語彙辞書でも高速な予測候補検索を実現した。開発した技術は2011年8月に発売した市販のカーナビゲーション“NR-MZ50”に搭載しており、今後ほかの機種にも展開する予定である。

参考文献

- 青江順一：ダブル配列による高速デジタル検索アルゴリズム、電子情報通信学会論文誌、J71-D, No. 4, 1592~1600 (1988)
- 三上崇志、ほか：ダブル配列を用いた文書検索用キーワード提示の高速化、情報処理学会研究報告、2009-FI-94(9) (2009)
- 成田知宏、ほか：カーナビゲーション向け音声認識の新たな展開、三菱電機技報、84, No. 9, 495~498 (2010)

阿倍博信*
峯慎吾**
奥村誠司**

マルチメディアを活用した技能継承システム

Multimedia Skill Succession System

Hironobu Abe, Shingo Mine, Seiji Okumura

要旨

団塊の世代の熟練技術者の大量退職による労働者不足への対策や国内の生産拠点の海外移転による現地技術者への技術教育の必要性の増加等を背景に、設備保守や機器製造の現場では、熟練技術者の保有する技術やノウハウを、ほかの技術者に伝える“技能継承”が重要な課題となっている。

本稿では、動画データを中心とするマルチメディア技術を活用した技能継承システムについて、主にHMI(Human Machine Interface)の側面を中心に述べる。

このシステムでは、動画データを活用することで、形式化が困難な暗黙知を分かりやすい形で提供できる。さらに、動画データの再生に同期して関連するアノテーション(付加情報)をテキストや図形や音声で提供する。

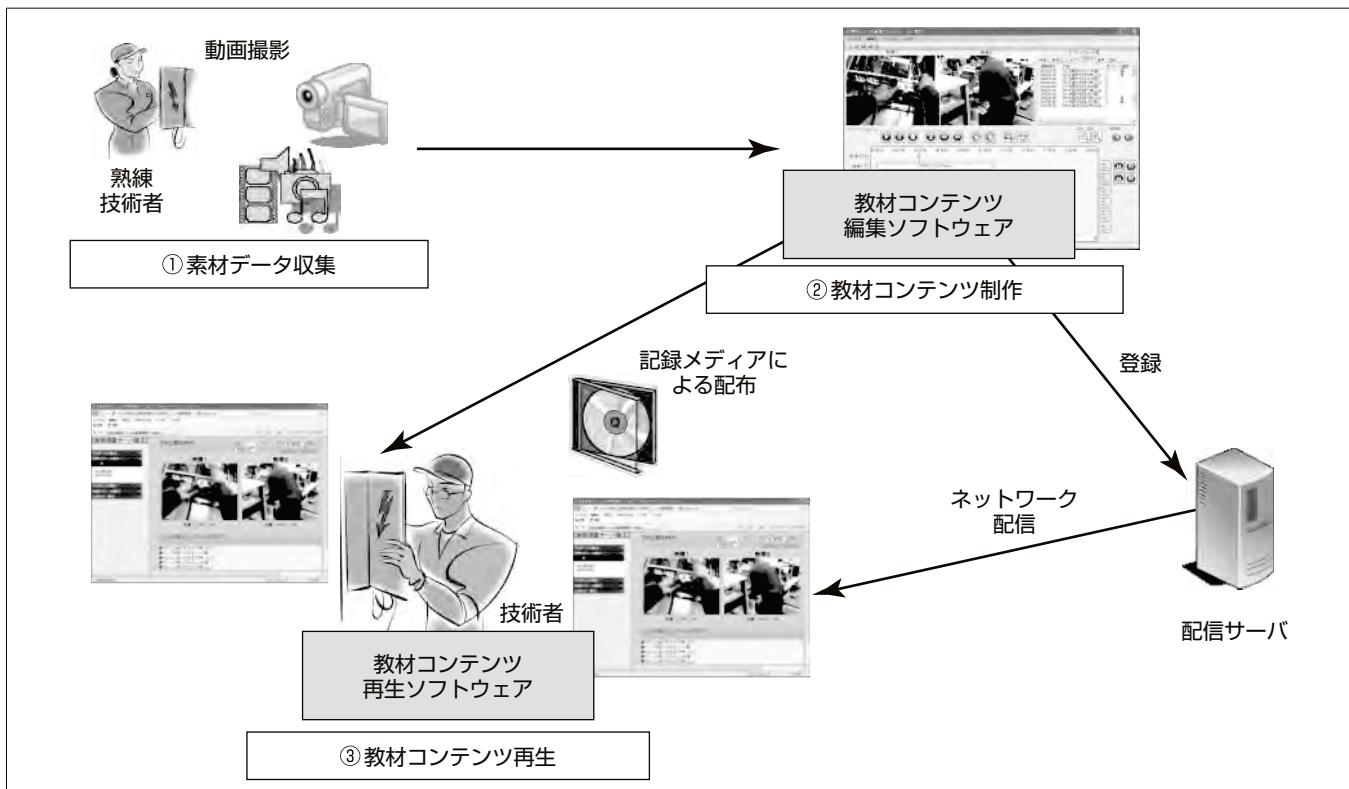
このシステムはWindows^(注1)上で動作し、教材コンテンツを制作する教材コンテンツ編集ソフトウェアと、教材コ

ンテンツを再生する教材コンテンツ再生ソフトウェアによって実現している。

このシステムの特長は、次のとおりである。

- (1) ドラッグ & ドロップなどのマウス中心の簡単操作を前提とする編集GUI(Graphical User Interface)によって、現場で教材コンテンツを簡単に制作できる。
- (2) 作業者の遠景映像と手元映像を同時撮影した動画データを同期させて再生する2画面動画同期再生機能を持つ教材コンテンツを制作できる。
- (3) 動画データと各種アノテーション(テキスト、図形、音声)の同期表示・再生機能を持つ。
- (4) 任意のテキストアノテーションの指定によって各種メディアを関連情報として表示するリンク機能を持つ。

(注1) Windowsは、Microsoft Corp. の登録商標である。



技能継承システムの構成

技能継承システムは、熟練技術者による作業現場を撮影した動画データを素材データとして入力し、動画データの配置や各種アノテーション(付加情報)やリンク等の設定をマウス操作で簡単に行う教材コンテンツ編集ソフトウェアと、制作した教材コンテンツをネットワーク経由、又は記録メディア経由で入手し、Webブラウザ上で再生する教材コンテンツ再生ソフトウェアで構成する。

1. まえがき

団塊世代の大量退職による熟練技術者の労働者不足が懸念されていた“2007年問題”は、各企業が定年延長や再雇用制度の整備を進めた結果、一旦は回避できた⁽¹⁾。しかし、2012年になると、雇用延長された団塊世代の技術者が65歳に到達するため、熟練技術者の大量退職が“2012年問題”として再び関心を集めている。

このような背景の元、設備保守の現場や機器の製造現場では、熟練技術者の保有する技術やノウハウを、ほかの技術者に伝える“技能継承”が重要な課題となっている⁽²⁾。

また、昨今の経済のグローバル化によって、国内の生産拠点(マザー工場)を中心とする製造現場の海外展開が進んでおり、海外拠点における現地技術者の教育体制整備が重要な課題となっている。

本稿では、機器製造や保守作業の現場への適用を目標として研究・開発を行っているマルチメディア技術を活用した技能継承システムについて、主にHMIの側面を中心に述べる。

2. 技能継承システム

2.1 技能継承における動画の意義

他者に技能を伝えようとする際に、伝えやすい形式知と伝え難い暗黙知に分けて説明できる。形式知を伝えるのは簡単であるが、暗黙知はそのままでは分かりづらく、伝えるためには工夫が必要である。

暗黙知を伝える方法としては映像化、数値化、言語化等が考えられるが、このシステムでは動画データを使用した映像化によって、暗黙知を分かりやすい形で再現する。

2.2 技能継承システムの要件及び設計指針

このシステムの要件について対象ユーザーの視点から整理し、次の項目を設計指針として設定した。

- (1) 教材コンテンツ制作時の要件は、制作作業の簡易化とした。設計指針として、マウス中心の簡単操作を設定した。
- (2) 教材コンテンツ再生時の要件は、映像情報の高度な活用による技能継承の効率化とした。設計指針として、動画データの2画面同期再生や動画データに同期した各種アノテーション(付加情報)の表示を設定した。

2.3 技能継承システムの特長

このシステムの特長について述べる。

(1) 現場で簡単に教材コンテンツを制作

ドラッグ＆ドロップなどのマウス中心の簡単操作をベースとしており、ワープロなどのパソコン操作ができる担当者であれば、現場で簡単に教材コンテンツを制作できる。

(2) 動画データの2画面同期再生

制作した教材コンテンツは、作業者の遠景映像と、手元

映像の2画面同期再生が可能で、作業の様子を全体と詳細の2つの視点から確認できる。

(3) アノテーションの同期表示

制作した教材コンテンツを再生する際に、動画データの再生にあわせてアノテーションを同期表示する。対応するアノテーションは、テキストアノテーション、図形アノテーション、音声アノテーションの3種類である。

テキストアノテーションは、作業手順や注意点等を説明するテキストを表示する。動画データの再生にあわせて表示されるテキストを見ながら、作業手順や注意点等を学習できる。

図形アノテーションは、動画データの画面上に、作業の注意箇所／エリア等を示す図形をオーバーレイ表示する。動画データ上に図形を表示して注意を促すことで、単に動画データを再生するだけでは見落としてしまうような技能も正確に伝えられる。

音声アノテーションは、動画データの再生にあわせてWMV(Windows Media^(注2) Video)形式やWAV(WAVE-form audio file format)形式^(注3)の音声ファイルを再生する。適切なタイミングで異常音やアナウンス音声等を再生することで、学習者の理解を深められる。

(4) リンク検索による関連情報の表示

教材コンテンツ制作時にテキストアノテーションに対してリンクを設定でき、教材コンテンツ再生時に必要に応じてテキストアノテーションに設定されたリンクを検索して関連情報を表示できる。

リンクとして設定できるメディアは、静止画ファイル(JPG^(注4)／BMP(BitMap))、音声ファイル(WMA(Windows Media Audio)／WAV)、動画ファイル(WMV)、HTML(HyperText Markup Language)ファイルである。例えば、再生されたシーンに応じて、リンク検索によって関連するWebページを表示できる。

(注2) Windows Mediaは、Microsoft Corp. の登録商標である。

(注3) Windows標準の音声ファイル形式

(注4) JPEG方式の画像ファイルにつけられる拡張子

2.4 技能継承システムの構成

このシステムは、次のソフトウェアで構成する。

(1) 教材コンテンツ編集ソフトウェア

動画データを中心とした素材データを入力し、動画データの簡単な編集、アノテーション設定等の作業を行い、制作した結果を教材コンテンツとして出力する。

(2) 教材コンテンツ再生ソフトウェア

Webブラウザ(Internet Explorer^(注5))上でスクリプトとして動作し、編集ソフトウェアで制作した教材コンテンツの再生制御を行う。

(注5) Internet Explorerは、Microsoft Corp. の登録商標である。

3. 教材コンテンツ編集ソフトウェア

編集ソフトウェアでは、アノテーション同期表示エリアに映像・音声・テキスト等の素材データを挿入し、映像／音声再生エリアに表示される動画データを確認しながら、アノテーションの同期設定を行う。

このソフトウェアでは、マウスによる簡単操作、同期設定支援機能、教材コンテンツ連結機能、パブリッシャー機能によって、複雑なマルチメディア教材の制作作業負荷を軽減できる(図1)。

3.1 マウスによる簡単操作

(1) アノテーションの容易な編集

ファイル選択ダイアログからの選択挿入、又はエクスプローラーからのドラッグ&ドロップによって、素材データをタイムライン上に挿入する。挿入された素材データは矩形(くけい)で表示され、表示時刻の調整やトリミング(動画データの再生開始位置や再生終了位置を変更して切り出す)操作もマウスで簡単に行える(図2)。

(2) 図形アノテーション設定

図形アノテーションは映像／音声再生エリアの動画データ上にオーバーレイ表示され、マウスでのドラッグ操作によって、位置や大きさを設定できる(図3)。

3.2 同期設定支援機能

アノテーション同期表示エリアでの動画データやアノテ



図1. 教材コンテンツ編集ソフトウェアの画面

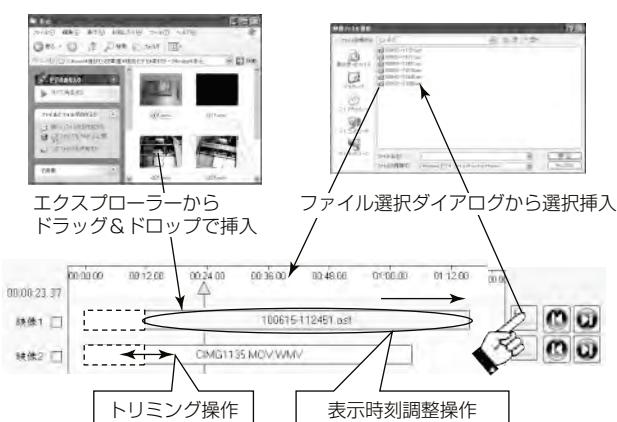


図2. アノテーションの同期設定

ーションの同期設定を支援する機能として、以下のような機能がある(図4)。

(1) アノテーション一括操作機能

複数のアノテーションのチェックボックスにチェックをつけると、関連情報を一括で操作可能となる。

(2) 吸着機能

アノテーションの移動中、スライダコントロールから表示される補助線に、アノテーションの先頭(又は最後尾)が吸着する(微調整作業の負荷を軽減)。

(3) タイムラインの拡大／縮小機能

タイムラインの拡大／縮小機能によって、細かい同期設定作業を簡単に実行できる。

(4) 微調整機能

10ms単位での高精度な時間調整が可能である。

3.3 教材コンテンツ連結機能

作業工程ごとに短い教材コンテンツを作成し、複数の教材コンテンツを連結することで、長編の教材コンテンツを作成できる。これによって、教材コンテンツの複雑化を軽減し、教材コンテンツの組み換えや、再利用が可能となる。

3.4 パブリッシャー機能

教材コンテンツ編集ソフトウェアで制作した結果を教材コンテンツ再生ソフトウェアで再生できるHTML形式に出力する。その際に複数のテンプレートを用意しておくことで、同じ制作結果から、2画面版だけでなく、1画面版等の使用目的にあわせた教材コンテンツを出力できる。

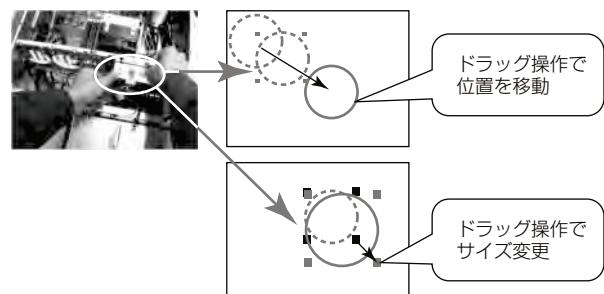


図3. 図形アノテーション設定

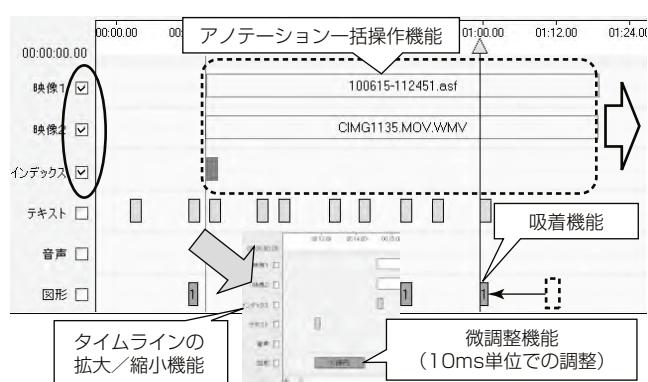


図4. 同期設定支援機能



図5. 教材コンテンツ再生ソフトウェアの画面

4. 教材コンテンツ再生ソフトウェア

再生ソフトウェアの画面は、作業映像とそれに関連するアノテーションを同期表示するメインフレームと、教材コンテンツの章構成をツリー形式で表示するインデックスフレームで構成される(図5)。

4.1 2画面動画同期再生機能

1画面の動画だけでは、作業者の姿勢や作業環境が分からぬ場合や、手元の作業が分からぬ場合がある。このソフトウェアでは、2画面の動画を同期表示させることによって、複雑な技能でも分かりやすい形で伝えられる。

4.2 アノテーション同期表示・再生機能

このソフトウェアでは、動画データと同期設定されたアノテーションを同期表示できる(図6)。

4.3 リンク検索・表示機能

教材コンテンツ制作時に、テキストアノテーションには関連情報として各種メディアへのリンクを設定できる。再生ソフトウェアで、リンクを設定したテキストアノテーションをクリックすると、リンク設定されたメディアが静止画ファイルやHTMLファイルの場合は新しいウインドウを開いて表示し、動画ファイルや音声ファイルの場合は新規にWindows Media Playerを起動して再生する教材コンテンツを学習しながら、設計図などの図面や、作業手順書等の電子マニュアルを確認できる。

4.4 インデックス/チャプタ機能

この機能は、インデックスフレームに教材コンテンツのチャプタ(教材の章立て)をツリー形状で表示する機能である。作業工程ごとに教材を分けて、長い作業工程でも見やすく管理できる。また、各チャプタ中に存在するインデックス(チャプタの中で重要な箇所を指す索引)も表示できる。インデックスをクリックすると、見たいシーンにジャンプできる(図7)。



図6. アノテーション同期表示・再生機能



図7. インデックス/チャプタ機能

5. むすび

機器製造や保守作業の現場への適用を目標として研究・開発しているマルチメディア技術を活用した技能継承システムについて、主にHMIの側面から述べた。このシステムの今後の拡張として、次の機能の開発を計画している。

(1) 画面記録ソフトウェアとの連携

機器製造や保守作業の現場では、計測機器などを接続したノートパソコンを補助ツールとして使用する場面が増えており、ノートパソコン上で動作するソフトウェアのマニュアル化のニーズが高い。そこでパソコン上の画面記録ソフトウェアと連携し、画面記録した動画データを素材データとして活用する方式について検討していく。

(2) 教材コンテンツのタブレットPC対応

ノートパソコンが使えない現場では、代替手段としてタッチパネル入力がベースのタブレットPCを適用していく必要がある。教材コンテンツは現状、マウス入力を前提としているため、新規にタッチパネル入力に対応したインターフェースを開発していく。

参考文献

- (1) 経済産業省：2007年版ものづくり白書、第2章（2007）
- (2) 森 和夫：技術・技能伝承ハンドブック、JIPMソリューション（2005）

成田知宏*
花沢利行*
岩崎知弘*

音声検索技術とその応用

Voice Search Technology and Application

Tomohiro Narita, Toshiyuki Hanazawa, Tomohiro Iwasaki

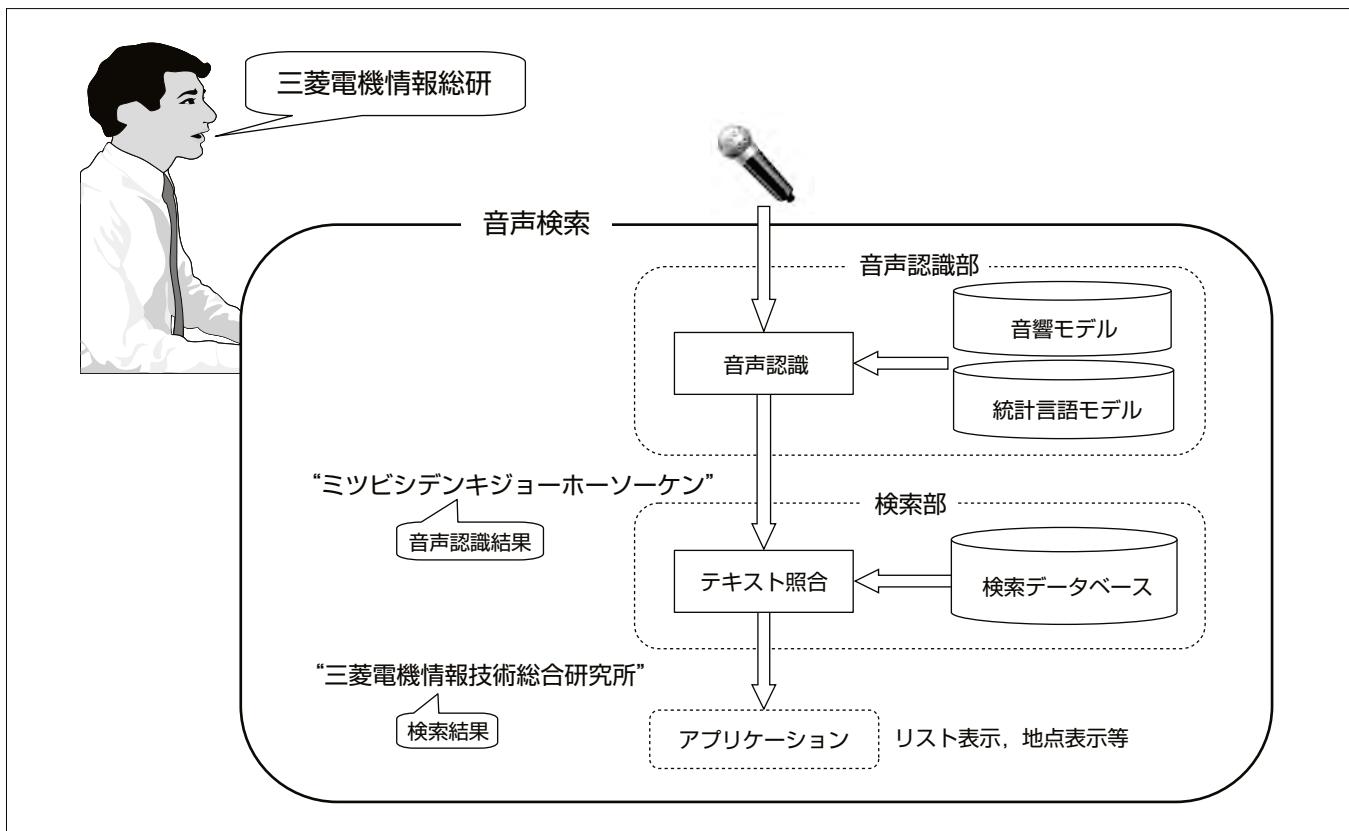
要旨

音声は人間にとて最も自然な入力手段であり、近年カーナビゲーションやスマートフォン等のコンシューマー向けの製品に音声認識機能が搭載されており身近なものになりつつある。また、製品の高機能化に伴い操作が複雑化していることも、音声による入力手段が注目される要因となっている。特に子供や高齢者等キーボード入力に慣れていないユーザーにとって音声による情報検索(以下“音声検索”という。)は利便性が高く、所望の情報を得るまでの時間を短縮できる。しかし、ユーザーは検索対象の正式名称を覚えていないことが多いため、正式名称以外の略称や通称に対応する必要がある。三菱電機ではこの課題に対し、どの音に近いかという音響的な探索と、どの言葉に近いかという言語的な探索を分離するアプローチに基づく音声検索方式を開発した。

開発した音声検索方式は下図のように、音声認識部と検索部で構成している。音声認識部は音響モデルと統計言語モデルを用いて入力音声を文字列に変換する処理を行う。検索部は音声認識結果の文字列を検索クエリとして、検索データベースからクエリに合致する情報を検索する処理を行う。この方式は次の特長を持っている。

- (1) 少少の認識誤りがあっても検索可能
- (2) 部分的な発話や倒置表現でも検索可能
- (3) 組み込み機器の限られたメモリ量、演算量で大規模なデータを検索可能

この方式は、カーナビゲーションにおける施設名検索に適用している。今後マニュアル検索、テレビの番組表検索等に広く応用が期待されている。



音声検索方式

音声認識部では音響モデルと統計的言語モデルを用いた音声認識を行い、入力音声を文字列に変換する。検索部では音声認識部で得られた文字列を検索クエリとして検索データベースから所望の情報を検索する。この図のように“三菱電機情報総研”という略称から“三菱電機情報技術総合研究所”を検索することができる。

1. まえがき

モバイル端末や車載端末のストレージ容量の増加、通信インフラのブロードバンド化が進むにつれて、大規模な情報を音声で検索したいという要求が高まっている。特に子供や高齢者等キーボード入力に慣れていないユーザーにとって音声検索は利便性が高く、所望の情報を得るまでの時間を短縮できる。

本稿では音声検索を構成する音声認識部及び検索部について述べ、応用例と今後の展望を述べる。

2. 音声検索の課題とアプローチ

音声検索の課題として、ユーザーが検索対象の正式名称を覚えていない、覚えていたとしても広く普及している略称や別名が用いられるという問題がある。例えば、“三菱電機情報技術総合研究所”は略称の“三菱電機情報総研”と呼ばれることが多い。この場合、従来方式では正式名称の“三菱電機情報技術総合研究所”のみが音声認識語彙として登録されるため、正しい音声認識結果を得ることができない。この問題に対しては、事前に様々な言い換え表現を作成しておくアプローチ⁽¹⁾があるが、認識語彙数の増加に伴い言い換え表現の追加・整備コスト及び演算量が増加するという問題があった。

これらの問題に対して当社では、どの音に近いかという音響的な探索と、どの言葉に近いかという言語的な探索を分離するアプローチを採用している。音響モデルと統計言語モデルを用いて音声から文字列への変換を行う音声認識処理と、音声認識のポスト処理としての検索処理に分離することで、語彙数が大幅に増加した場合でも演算量の増加を抑制することが可能になった。また、認識処理と検索処理を分離したこと、多少の音声認識誤りに対しても頑健で、部分的な発話や倒置表現でも検索可能な機能を実現している。

3. 音声検索方式

図1に開発した音声検索方式の構成を示す。音声認識部では音響モデルと統計言語モデルを用いて音声認識を行い、入力音声を文字列に変換する。検索部では音声認識部で得られた文字列を検索クエリとしてデータベースの検索を行う。検索部は音声認識部とは独立な構成となっており、タッチパネルなどの音声以外の入力結果に対しても検索が可能である。これより先、音声検索対象が施設名の場合を例に音声認識部と検索部の詳細について述べる。

3.1 音声認識方式

音声認識部で用いる音響モデルは音響特徴量の統計量を表したもので隠れマルコフモデルと呼ばれる確率モデルで表現される。統計言語モデルは認識対象語彙のデータであり、単語間の接続関係を表現したものである。

3.1.1 音声認識の基本方式

まず、入力信号に対してパワーやその他の情報に基づき音声区間の検出を行う。次に検出した音声区間にて一定時間ごとに特徴量の抽出を行う。特徴量には音声スペクトルの概形を表現するパラメータが用いられる。次に音響モデルの尤度(ゆうど)計算を行う。尤度計算は入力信号に対するモデルの尤(もっと)もらしさを計算する処理である。一般的に音響モデルの最小単位は前後の音の違いを考慮し音素の特徴を表現するトライфонと呼ばれるモデルが使用されるが、必要なメモリ量と演算量が大きいという問題がある。そこで、音素片と呼ばれる独自のコンパクトなモデルを採用することで、省メモリ化と少演算量化を実現している。最後に統計言語モデルを利用したデコード処理が行われ、入力音声に対する尤もらしさ(認識スコア)が一番高い単語列を認識結果として出力する。

3.1.2 統計言語モデルを用いた音声認識

2章で述べたように言語モデルに“東京国立近代美術館”

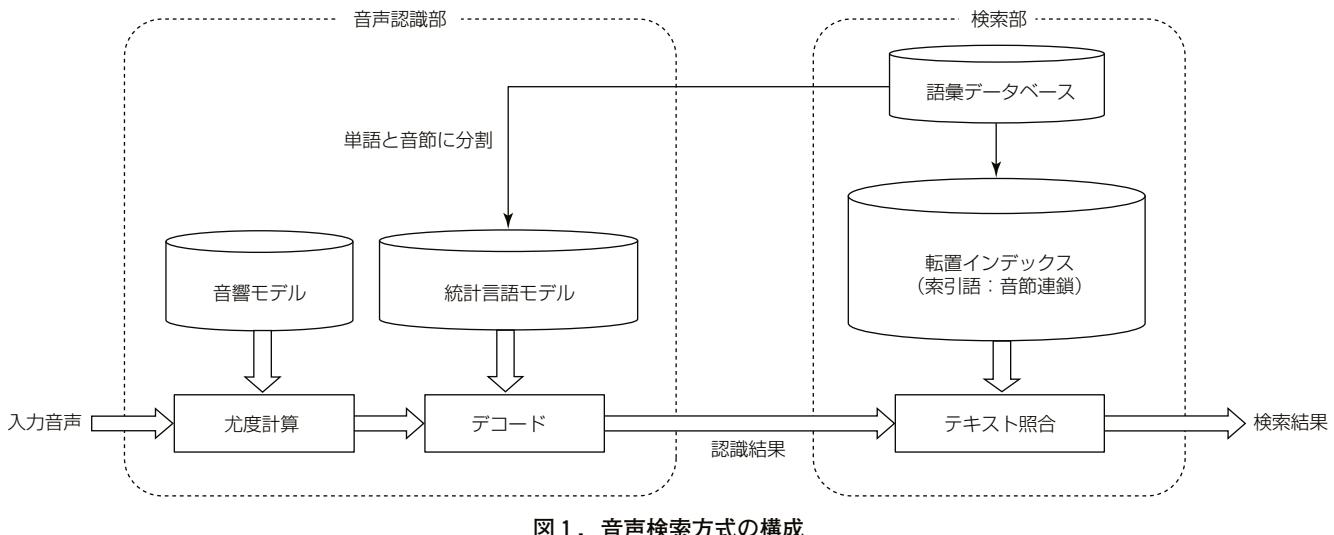


図1. 音声検索方式の構成

だけを登録した場合，“東京近代美術館”のように“国立”を省略した発話は認識できない。対策として“東京国立近代美術館”を“東京”“国立”“近代”“美術館”というように単語に分割し、各単語とともに単語間の接続規則を言語モデルに登録する方式がある。例えば、“国立”は省略可能という規則を登録すれば、“東京近代美術館”という言い換え表現も認識可能となる。しかし、多様な呼称をカバーする規則を人手で整備するのは困難である。

これに対し統計言語モデルでは単語間の接続のしやすさを統計量に基づく確率で表す。例えば、先に述べた例の場合，“東京”という単語の次に“国立”や“近代”が続く確率が登録されており、次に続く可能性のある単語を予測しながら認識を行う。単語間の確率はデータベースから自動的に算出が可能である。また、データベースに出現しない単語接続に対しても、近似的な確率を与えることで多様な呼称に対応することが可能である。

ただし、検索対象が数十万語の規模になると、検索対象に含まれる単語数も数万語のオーダーとなるため、音声認識時のメモリ量と演算量が問題となる。そこで、各単語をより小さな単語に分割して表現する方式⁽²⁾を採用している。例えば、“美術館”は、“びじゅつ”と“かん”という単語で表現する。単語の総数は施設名中での出現頻度の高い上位N個に制限する。単語の連鎖で表現できない場合は、音節の連鎖で表現する。例えば、“近代”は、“き”“ん”“だ”“い”という音節の連鎖で表現する。このように音節を単位とすることで、検索対象の施設数が増加しても認識単語数を一定数に保ち、メモリ量及び演算量に制限のある組み込み機器でも大規模な音声検索を可能にしている。

3.2 検索方式

3.2.1 テキスト照合に基づく音声検索

3.1.2項で述べたように音声認識部では施設名を単語と音節に分割して認識する。認識結果は登録施設名と完全一致しない場合が生じるので、認識結果の音節列と最も良くマッチする候補名を検索する方式⁽²⁾を採用している。検索部では図2に示す転置インデックスを用いて、各施設に対する検索スコアを算出し、スコア最大の施設を検索結果とする。索引語としては図に示すとおり2音節の連鎖を用い、索引語ごとに、その索引語を含む施設のID番号を保持する。

例えば、認識結果の音節列が“しいぱら”的な場合，“しい”“いぱ”“ぱら”的な3種類の音節連鎖が存在する。次に転置インデックスを参照して、これらの音節連鎖を含む施設名に対し、音節連鎖1個ごとにスコアを1点加算する。図2の例では、施設ID=2“しいぱらだいす”的なスコアが3となり、検索結果の第一位として出力される。このように転置インデックスを参照することで、高速にスコアを計算することができる。なお、音節連鎖を索引語とする利点は、単語や

語彙データベース

施設名ID	施設名の音節列
0	よこはま すう ら しあ
1	よこはま ぎん こ う
2	し い ぱ ら だ い す



自動生成

転置インデックス

索引語：音節連鎖	施設名ID
よこ	0, 1, ...
こは	0, 1, ...
はま	0, 1, ...
し い	2, ...
い ぱ	2, ...
ぱ ら	2, ...
:	...

図2. 転置インデックス

単語連鎖と比較して音声認識誤りに頑健なことである。例えば，“東京近代美術館（とーきょーきんだいびじゅつかん）”という発話に対し、音声認識部の出力結果が“東京近代武術館（とーきょーきんだいぶじゅつかん）”となった場合、単語単位では誤認識した“武術館”が“美術館”とはマッチせずスコアに寄与しないが、音節単位なら“じゅつ”“つか”“かん”的3個の音節連鎖が“びじゅつかん”とマッチするため、正解の“東京近代美術館”的スコアに寄与できる。

3.2.2 必須(ひっす)形態素の検証

3.2.1節で述べた検索方式の問題点の一つとして不自然な検索結果が湧き出す場合がある。例えば，“清水砂丘（しみずさきゅー）”という発話の音声認識結果が“しみずさき”のように誤認識を含む場合，“A書房弘前清水3丁目店（/えー／しょ／ぼー／ひ／ろ／さ／き／し／み／ず／さ／ん／ちょー／め／て／ん／）”のほうが，“清水砂丘（/し／み／ず／さ／き／ゅ／ー／）”よりもマッチする箇所（下線部）が多い（/は音節の区切りを表す）。このため、前者が正解よりも上位候補として誤検索されるという問題点があった。

“A書房”などのブランド名を含む施設を検索する場合、ブランド名は省略しないと考えられる。そこで、施設名ごとにその施設を検索する際に省略しないと考えられる文字列（以下“必須形態素”という。）を抽出した必須形態素辞書を事前に作成しておく、検索結果のうち必須形態素を含む施設名に対し、入力音節列中にその施設の必須形態素が含まれるか否かを検証する方式⁽³⁾を開発した。そして、必須形態素が含まれない場合、検索スコアに事前に定めたペナルティを付与する。なお、必須形態素の検証には置換、挿入及び脱落を許す文字列マッチングを用い、誤認識で音節表記が完全一致しない場合でも検証可能とする。

このような必須形態素の検証処理によって、認識結果の音節列“しみずさき”は“A書房弘前清水3丁目店”的必須形態“えーしょぼー”を含んでいないため、検索スコアにペナ



図3. HDDカーナビゲーションシステムNR-HZ001

ルティが付与されることになり、正解の“清水砂丘”より上位に検索されることを抑制できる。

4. 音声検索の応用例

ここまで述べた音声検索技術の応用例と今後の展開について述べる。

4.1 カーナビゲーションの音声50音検索

音声50音検索とは、音声で施設名を入力し検索を行う機能であり、当社のHDD(Hard Disk Drive)カーナビゲーションシステム“NR-HZ001(図3)シリーズ”に搭載されている⁽⁴⁾。従来のカーナビゲーションシステムでは50音検索の入力画面(図4)上でタッチパネルから1文字ずつ施設名の読みを入力する必要があり入力に時間が掛かっていた。さらに、走行中はタッチパネル操作が禁止されるため、停車中以外は入力することができなかった。NR-HZ001では音声50音検索を導入することで、所望の施設を検索するまでの時間を短縮するとともに、走行中の50音施設検索を可能にしている。この製品の音声50音検索には次の特長がある。

- (1) 都道府県単位で最大100万件規模の超大語彙の施設検索が可能
- (2) 部分表現や倒置表現等、正式名称以外の表現での施設検索が可能
- (3) 追加発話による施設の絞り込みが可能

追加発話による絞り込みの例を次に述べる。まず、“三菱電機”と発話すると“ミツビシデンキ”を検索クエリとして検索が行われ、検索結果がリスト表示される。次に絞り込みのために“研究所”を発話すると、“ミツビシデンキ”と“ケンキュウジョ”を検索クエリとして再検索が行われ、施設名のリストが更新される。このように複数の発話によって逐次的に施設を検索することができる。

4.2 今後の展開

本稿で述べた音声検索技術は次のような様々な製品や情報システムに適用可能である。



図4. NR-HZ001の50音検索画面

- (1) 製品マニュアルの検索
- (2) 番組表の検索
- (3) コールセンターでの顧客情報の検索
- (4) 点検業務における業務データの検索

ただし、これらの製品・システムは各々機能要件が異なるため適用の際には別途検討が必要である。例えば、使用時の騒音環境、使用可能なメモリ量及び演算量、要求性能、応答速度等が要検討項目に該当する。今後はこれらの製品・システムに対する音声検索のニーズを探るとともに、必要な要素技術の検討を進める予定である。

5. む　す　び

音声検索を構成する技術要素と応用例について述べた。大規模データの検索では、入力時間及び入力自由度の面から音声は最も適した入力手段であると考えられる。今後は様々な製品やシステムへの展開を検討するとともに、意図理解技術や音声対話技術を統合することで、より直感的で自由な音声検索の実現を目指す所存である。

参考文献

- (1) 岩崎知弘, ほか: カーナビゲーションシステムの音声インターフェース－現状と将来展望－, 三菱電機技報, 78, No. 9, 613~617 (2004)
- (2) 花沢利行, ほか: 統計言語モデルによる音声認識とテキスト照合に基づく大語彙音声検索, 日本音響学会研究発表会講演論文集, 61~62 (2009)
- (3) 花沢利行, ほか: 必須形態素の検証によるテキスト検索の改善, 電子情報通信学会総合大会論文集D-14-7 (2011)
- (4) 成田知宏, ほか: カーナビゲーション向け音声認識の新たな展開, 三菱電機技報, 84, No. 9, 495~498 (2010)
- (5) 岡登洋平, ほか: 統計言語モデルによる音声認識とテキスト照合に基づく大語彙音声検索の精度改善検討, 日本音響学会研究発表会講演論文集, 99~100 (2010)

大塚貴弘*
川島啓吾**
山浦 正**

テキスト音声合成技術

Text-to-Speech Technology

Takahiro Otsuka, Keigo Kawashima, Tadashi Yamaura

要旨

三菱電機は家電製品の高度な機能をらくに使いこなして暮らしを楽しくするアシスト機能を“らく楽アシスト”と名づけて展開しており、その中核機能の一つが音声による操作ガイドである。また、何らかの作業中の人が、同時に多くの人に情報を提供するには、音声は最も有効な情報伝達手段である。この音声を、あらかじめ録音することなしに、テキストから合成する技術が、テキスト音声合成技術であり、高い期待を受けている。しかし、当初は、機械的に単調な音声しか合成できず、応用分野は限定されていた。そこで、当社では、幅広い分野に適用できるよう、音声合成技術の高品位化と、利用しやすい合成ソフトウェア、ツールの開発に取り組んでいる。

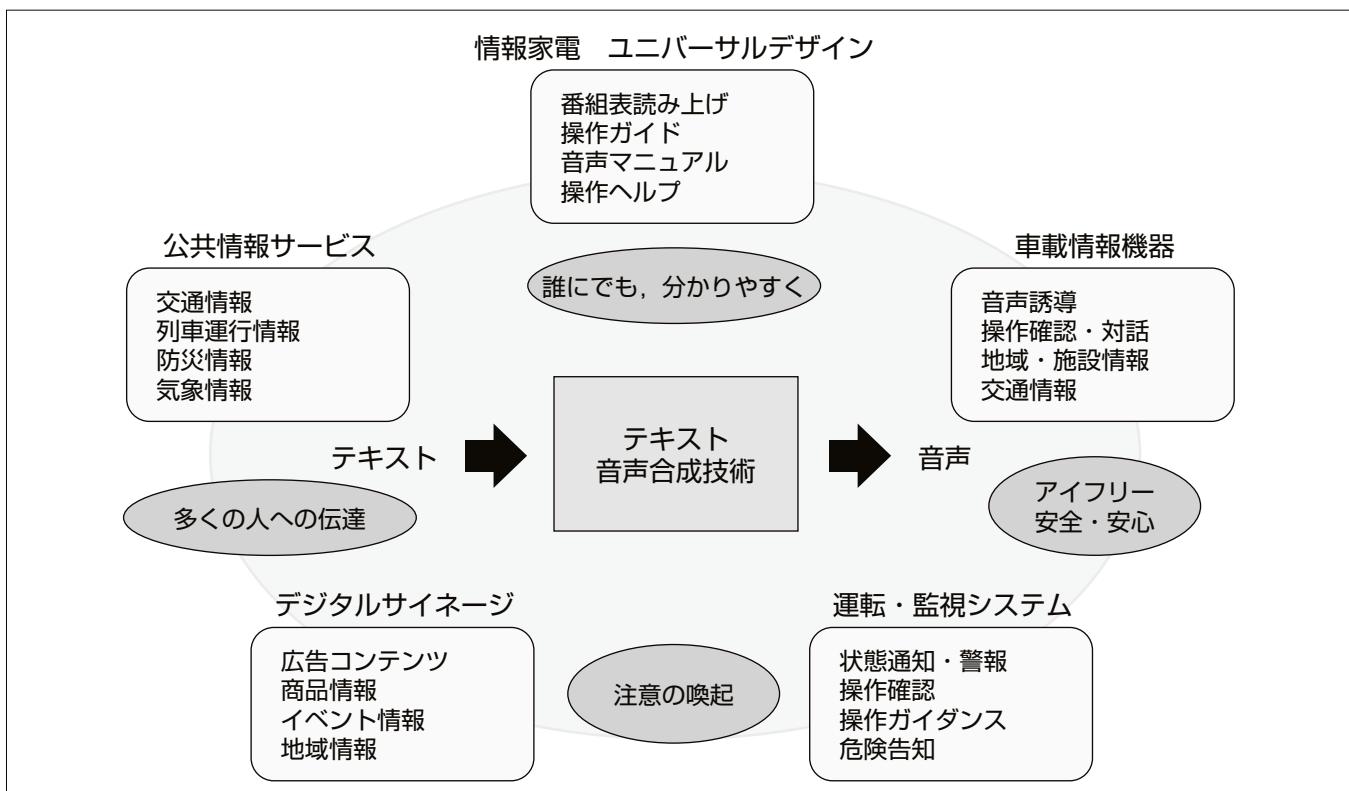
高品位化に対しては、大規模な音声データベースから、イントネーション・リズムの制御規則を学習する方式と、合成に用いる波形辞書をデータ収録時の変動要因を取り除

いて作成する音質安定化手法によって、明瞭(めいりょう)で自然な合成音声を得ている。

さらに、広い応用分野に適用できるよう、次の特長を備えている。

- (1) 漢字仮名交じり文のほか、ITS(Intelligent Transport Systems)車載器用音声合成記号(JEITA((社)電子情報技術産業協会) TT-6004)などの中間言語からも合成可能
- (2) 適用するハードウェアスペックに応じて、適切な処理量・メモリ量で構成することが可能
- (3) 音声作成補助ツールによって、簡単な操作で所望の抑揚、発話速度に調整可能

このテキスト音声合成技術は、AV(Audio Visual)家電やカーナビゲーションシステムをはじめとして、様々な応用分野への適用が進んでいる。



テキスト音声合成技術の応用分野

テキスト音声合成は、入力されたテキストから音声を自動合成する技術であり、音声を収録する必要がないのでコストがかからず、伝達した内容を即座に分かりやすく、作業中の人が多数の人にも提供ができる。高機能化する家電の使い方のガイドや、運転中の人にに対する情報提供等に適用されており、今後さらに、運転監視、公共情報サービス、デジタルサイネージ等へ適用が期待されている。

1. まえがき

音声による情報伝達は、分かりやすいだけではなく、作業中の人や多数の人への情報伝達が可能であり、他に代替しえない情報伝達手段である。任意のテキストから音声を合成するテキスト音声合成技術⁽¹⁾は、録音することなく、自由に音声を合成できる技術として古くから大きな期待を受けてきた。しかし、当初の合成音声は、内容を理解できる了解性はあったものの、肉声感のない機械的音質で、韻律(イントネーション、リズム)も単調なため、応用分野は限定されていた。

当社では、合成音声の高品位化に加え、種々の用途に適用しやすい合成ソフトウェアとツールの開発に取り組んできた。その結果、家電製品をらくに使いこなす当社の“らく楽アシスト”の中核機能の一つである“音声ガイド”に適用されるなど、現在、幅広く応用が進みつつある。本稿では、テキスト音声合成方式の概要を解説し、その特長を述べる。

2. テキスト音声合成方式

2.1 概 要

図1に、このテキスト音声合成方式の構成を示す。音声合成するテキスト(漢字仮名交じり文)が入力されると、言語辞書を参照し、テキストの解析を行い、読み、アクセント位置、品詞等を決定し、中間言語と呼ばれる表現に変換する。

イントネーション・リズム制御では、音声のイントネーションを表すピッチ周期(声の高さ)系列と、リズムを表す各音素の継続時間長を生成する。それらの情報をもとに、波形辞書中の音片(短い合成単位)データを読み出し信号処理によって要求されたピッチ周期、継続時間になるように加工し、接続することで音声波形を生成、出力する。

この方式では、中間言語も直接入力することができる構成となっており、ITS(高度道路交通システム)で規定され

ているJEITA TT-6004 ITS車載器用音声合成記号にも対応している。また、中間言語のみを扱う、テキスト解析部を持たない簡易型の合成ソフトウェアを端末機器で動作させ、テキスト解析をサーバで動作させる分散構成も可能である。さらに、読み、アクセント型、ポーズ位置、ポーズ長等を中間言語上で直接指定することで、所望の合成音声を得ることもできる。

2.2 テキストの解析

テキスト解析部では最初にプリプロセッサを適用する。これは、略語の処理などを行うもので、例えば、“10/9”を“10月9日”又は“9分の10”などへ変換する。種々の規則が用意されており、用途に応じた変換規則が適用できる。

次に、読み・アクセント解析を行う。合成対象であるテキストが入力されると、言語辞書を参照し、候補となる形態素を抽出したのち、接続コスト最小法に基づいて最適形態素列の選択を行う。形態素の接続しやすさを表す接続コストを最小となるように形態素を選択することで、文法の誤りや、通常、解析が困難な話し言葉表現に対しても適切な解析が可能となっている。

次に、選択した形態素列に対し、複合化処理(形態素の結合とアクセントの移動)、前後の音韻環境を考慮した無声化処理を行い、読みやアクセント位置を示す中間言語を出力する。

言語辞書の基本部分は、大量のテキストデータから抽出した分野に依存しない出現頻度の高い単語で構成している。さらに、適用製品で使用するテキストを用いて分野特有の単語を追加することもできる(例えば、交通情報を想定した地名データなど)。言語辞書には、表記・読み・品詞・アクセント位置等が含まれる。

2.3 イントネーション・リズム制御

イントネーション・リズム制御では、統計的予測モデルによって、音声のイントネーション(ピッチ周期系列)とリズム(音素の継続時間長)を決定する。まず、テキスト解析結果又は中間言語から、予測を行うための要因(音節の文

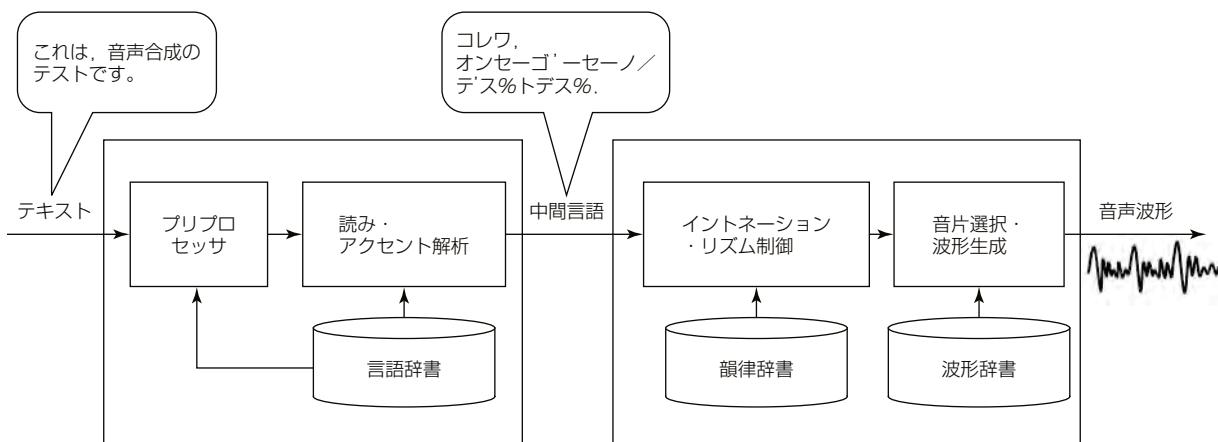


図1. テキスト音声合成方式の全体構成

内位置、フレーズの音節数、アクセント型、音素名等)を抽出する。これらの要因を入力とし、予測モデルを用いてイントネーションとリズムを決定する。

予測モデルで用いる要因に対するパラメータセットをここでは韻律辞書と呼び、学習用音声コーパス(音声と言語情報等を体系化したデータベース)から機械学習によって求める(図2)。一般には予測値とコーパスの観測値の平均誤差を最小とするように学習を行うが、このとき、全観測値に対する平均誤差最小を基準とすると、考慮していない要因による変動を平均化してしまい、合成音声の自然性が損なわれるという問題があった。そこで、学習過程で、要因に対する固有の変動を持つ観測値を選択し、それに対する誤差を最小化する学習方式を開発した。この方法によって大量コーパスの学習でも自然性が維持できる。

2.4 音片選択・波形生成

波形辞書には、音声コーパスから抽出した種々の音声素片が記憶されており、音片選択では、合成する音素の継続時間長やピッチ周期系列に近く、前後の音片データとのつながりの良い音片データを選択する。

波形生成では、合成する音素の継続時間長やピッチ周期系列に合わせるよう、選択された音片データを信号処理によって変形・接続して音声波形を作成する。

波形辞書は、ナレーターが発声した大量の音声コーパスから、学習によって最適な区間を選択し、音素波形を切り出して構築する。一般にデータが大量であれば、前後の音の種類、声の高さ、継続時間による変化のパターンが増え、有効な学習によって品質が向上する。しかし、収録ごとの声質が変化してしまうことも多く、必要のない変動データを学習してしまうなどの問題が生じ合成音声が不安定なものになるという問題があった。そこで、同じ特徴を持つと

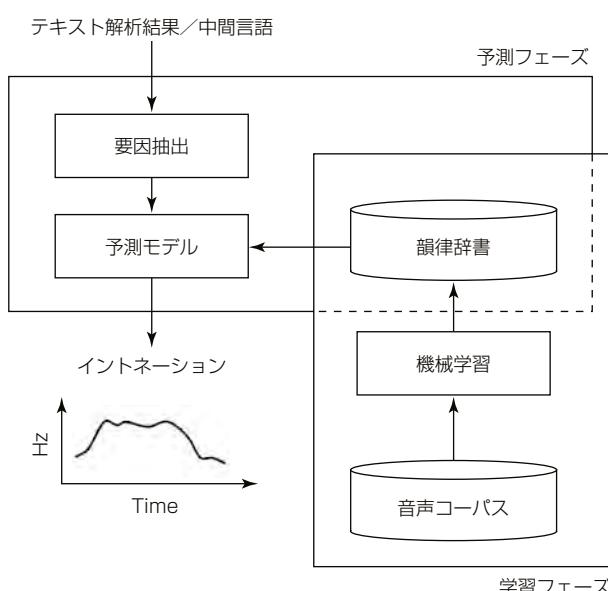


図2. イントネーション制御

考えられる音から収録ごとの変動成分を予測し、変動成分を引き去ることで、音質の均一化を行う音質安定化技術を開発した。

2.5 動作ハードウェアのスケーラビリティ

処理速度及びメモリサイズは、使用するハードウェアに応じて、スケーラブルな構成をとることができる。処理能力の低いハードウェアで使用する場合は、波形辞書からの音片データの選択時に、予備選択方法を導入することで音質劣化を最小限に抑えながら演算量の低減が可能である。(200MHzのCPU(Central Processing Unit)で使用率25%未満)

組み込み向け機器で省メモリサイズが求められる場合には、音片データ数を絞り込むことで波形辞書のサイズを削減でき、最小構成では約7MBのメモリで実装可能である。さらに、入力を中間言語に限定し、かつ波形辞書を外部ディスクに格納することができれば約3MBまで低減できる。

一方、サーバ向け機器でメモリや処理能力の制約が緩い場合には、音片データのバリエーションを増やすとともに、波形辞書からの音片データ選択を高精度に行うことで、最高の音質を得ることができる。

2.6 評価

開発技術(統計的予測モデルによるイントネーション・リズム制御、及び独自の音質安定化技術)の効果を確認するために、5段階MOS(Mean Opinion Score:複数の評定者による5段階主観評価値の平均)による主観評価試験を行った。評定者8名で、提示された合成音声に対して、情報案内音声として5段階(5:非常に良い、4:良い、3:普通、2:悪い、1:非常に悪い)で評価した。合成音声は、単語及び文の24サンプルを用いた。

組み込み向け機器を想定し、使用メモリ7MBという条件で方式を構成した。開発した技術を適用前の合成音声(旧版)と適用後の合成音声(新版)とを比較した。図3に示す結果から、旧版と比較して1点以上の大きな改善効果が分かる。

3. 音声作成補助ツール

高い肉声感と了解性が求められる公共放送、カーナビゲーション音声、情報家電の操作ヘルプ音声等では、編集音

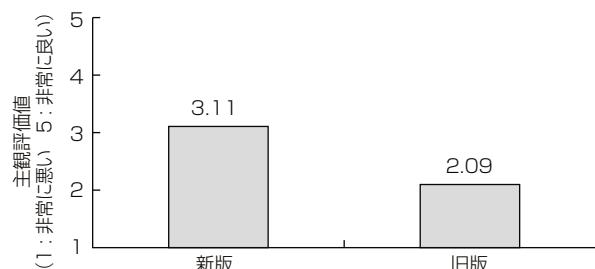


図3. 主観評価結果



図4. 音声作成補助ツールのメイン画面



図5. 音声作成補助ツールのサブ画面

声合成方式や録音再生方式が用いられることが多い。しかし、これらの方法は、音声収録・整備にコストがかかること、一貫した音色で多数の音声を収録することが困難であること、後日の音声データの追加の際に同一ナレータを確保できないリスク等の課題があり、任意のテキストから音声を即時作成できるテキスト音声合成方式への期待が高い。

そこで、製品開発者が、ナレータ音声の代わりとなる、より高品位な合成音声を作成するために、次の機能を持つ音声作成補助ツール(図4、図5)を試作した。

- ①任意のテキストから合成音声を生成できる。
- ②合成音声のアクセント位置などを中間言語で調整できる。
- ③合成音声全体の、声の高さ、話速、声の大きさ、抑揚の大きさを調整できる。
- ④合成音声の部分的な声の高さや話速を調整できる。

このツールによって、操作に慣れれば、録音音声に匹敵する韻律特徴を持った合成音声が実現できることを確認した。今後、さらなる操作性と音質の改良を行い、より高い合成音声品質が要求される分野に適用を図る予定である。

4. む す び

当社で開発したテキスト音声合成方式と音声作成補助ツールについて述べた。テキスト音声合成方式は、様々な製品への適用に高い期待を受けている。車載機器・情報機器の応答音声⁽²⁾、プラント運転・監視システムにおける機器の誤操作を防ぐためのガイダンス音声や状態通知音声、視覚障がい者のサポート⁽³⁾、複雑になる家電製品の“らく楽アシスト”のさらなる展開等である。今後も、音質の向上に加え、音声作成補助ツールによる音声コンテンツ作成など、利用用途の拡大に努める。

参 考 文 献

- (1) 藤井洋一, ほか: テキスト音声合成技術, 三菱電機技報, **76**, No. 8, 507~510 (2002)
- (2) 成田知宏, ほか: カーナビゲーション向け音声認識の新たな展開, 三菱電機技報, **84**, No. 9, 495~498 (2010)
- (3) 川勝かがり: デジタルテレビのユニバーサルデザイン, 三菱電機技報, **85**, No. 3, 167~170 (2011)

佐々木雄一*
宮原景泰*
川又武典*

3Dタッチパネル技術

3-Dimensional Touch Panel Technology

Yuichi Sasaki, Kageyasu Miyahara, Takenori Kawamata

要旨

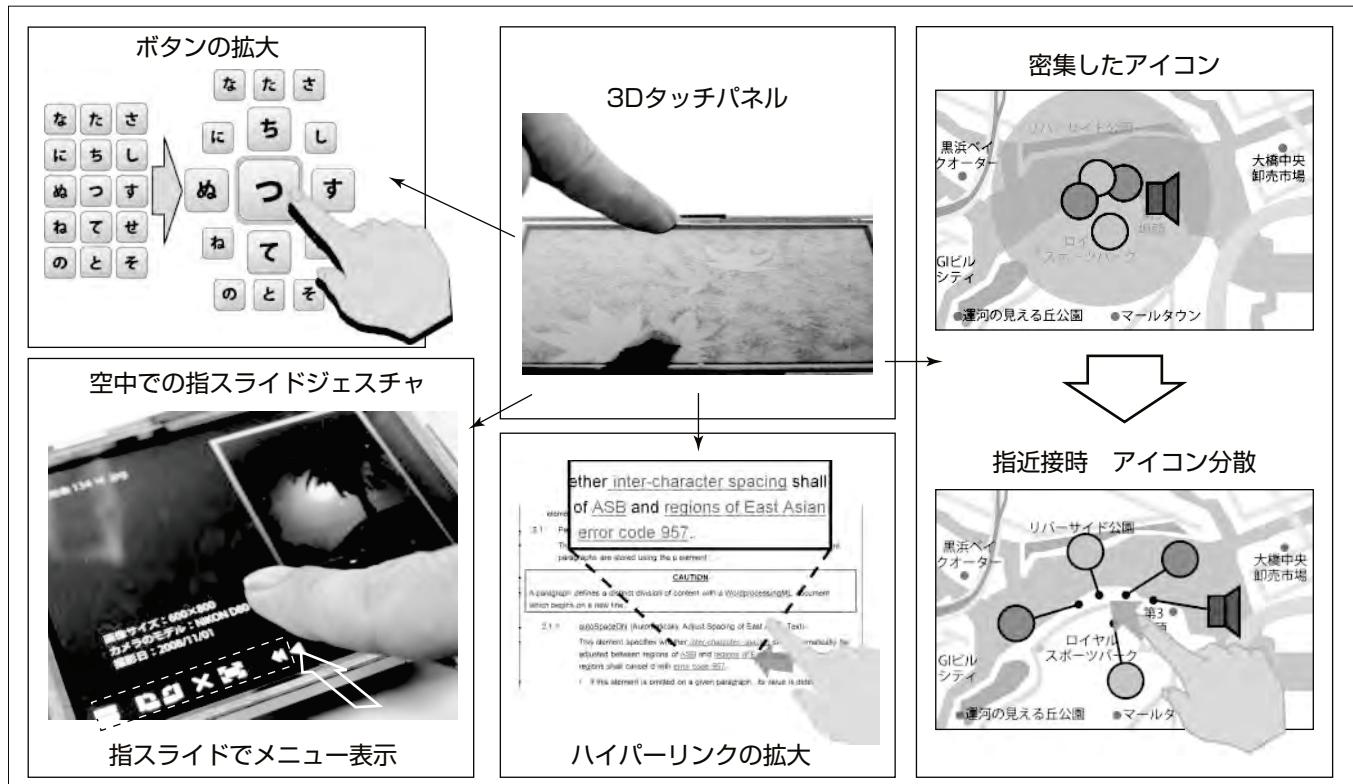
近年、スマートフォンやタブレットパソコン等、静電容量方式のタッチパネルを搭載した情報端末が増加している。このタッチパネルは、軽いタッチで反応し、耐久性に優れているという理由から、今後も更に広く普及していくものと考える。

タッチパネルで指による操作をする場合、ジェスチャ操作で直感的に操作することができる反面、ボタンやハイパーアリンク等が狭い領域に密集しているために、選択ミスが発生しやすいという課題があった。これは、操作領域を大きくして指操作に適したコンテンツにすることで解決可能であるが、一覧性が低下する。

そこで、三菱電機ではタッチパネル面から離れた位置に存在する指を検知することで、ユーザーが操作しようとしている領域を動的に拡大表示し、一覧性を保ったまままで選択ミスを抑制する3Dタッチパネルを試作した。

3Dタッチパネルは、ITO(Indium Tin Oxide:酸化インジウムスズ)センサ電極で構成される静電容量(投影容量)方式のタッチパネルを用いている。指の近接を検知する場合は、検出感度を高めるために複数電極を結合して用い、指の接触を検知する場合は、高い分解能で位置検出するために電極を分離する。この動的な電極結合制御方式を開発した。さらに、電極間の寄生容量を低減させるシールド電極制御や、3次スプライン補間による座標計算を適用した結果、タッチパネル面から20mmの高さにある指の近接が検知可能となり、指の近接、接触位置を実用的な精度で検出できることを確認した。

今後は、より遠方の指を検知できるように感度を向上させるとともに、3Dタッチパネルで指の近接を活用したユーザーインターフェースについても検討を行う予定である。



3Dタッチパネル技術の応用例

3Dタッチパネル技術は、一般的なITOセンサ電極で構成される静電容量方式のタッチパネルを用いて、タッチパネル上20mmに近接している指の位置を検出することを可能とする。指の近接を検知してユーザーの操作予測を行うことで、例えば、近接位置周辺のボタンやハイパーアリンクの拡大、空中で指をスライドするジェスチャによるメニュー表示、密集したアイコンの動的な分散等、タッチパネルの操作性向上に幅広く役立てることができる。

1. まえがき

近年、スマートフォンやタブレットパソコン等タッチパネル付きの情報端末が増えている。タッチパネルの指検出方式には、抵抗膜方式や静電容量方式、赤外線方式、超音波方式等様々な種類があり、用途に応じて最適な方式が採用されている。この中で、静電容量方式は、タッチパネルと指との間に生じる静電容量を検出するもので、非常に軽いタッチで入力操作が可能であり、耐久性に優れているため、今後も広く普及していくものと考える。

タッチパネルで指による操作をする場合、ジェスチャ操作で直感的に操作することができる反面、ボタンやハイパーリンク等が狭い領域に密集しているために、選択ミスが発生しやすいという課題があった。これは、操作領域を大きくして指操作に適したコンテンツにすることで、選択ミスを少なくすることができる。しかし、この方法では、操作領域が大きくなる分、他の情報を表示できなくなり、一覧性が大きく低下する。

そこで、当社では指の近接を検知し、ユーザーが操作しようとしている領域のみを動的に拡大表示することで、一覧性を保ったまままで選択ミスを抑制する3Dタッチパネルを試作した。

2. 投影容量型タッチパネルの指検出原理

静電容量方式のタッチパネルには表面容量型と投影容量型の2種類の指検出方式がある⁽¹⁾。3Dタッチパネルは、投影容量型の指検出方式に基づいており、この章では、この投影容量型タッチパネルの指検出方法について述べる。図1が投影容量型タッチパネルのセンサ電極の構成である。一般的にセンサ電極は、抵抗値が均一で透明な導電体であるITOが用いられる。各電極はXYの重なり部分を最小にするため、ダイヤモンド型を連ねた形状となっている。本稿では、この形状を一本のセンサ電極として簡易化して図示する。

検出回路は、指の静電容量を検出するために、まず、タ

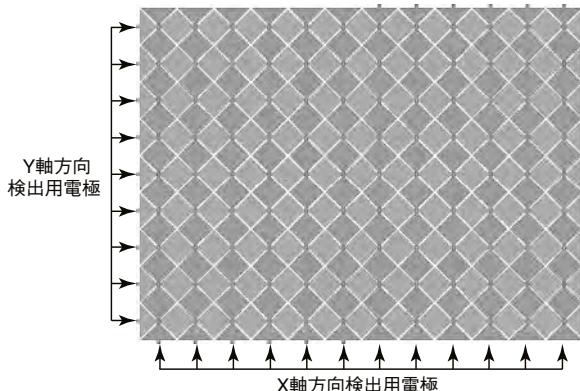


図1. 投影容量型タッチパネルのセンサ電極構成

ッチパネル面に配置された電極に信号を加える。ここで、タッチパネル面に指が近づくと、電極と指との間に静電容量結合が発生し、電荷が蓄積される。次に、電極と検出回路を接続する。このとき、指が存在する電極には、蓄積された電荷に応じて電圧が発生する。このため、電極からの信号を測定すると、指がない状態に比べて出力レベルが変化し(図2)，指の有無と接触位置を検知することができる。

検出回路は指がない状態の出力信号をベースライン値として保存しておき、毎回のセンサ出力値とベースライン値の差分を検出値とする。指が存在しない場合は小さな検出値が、存在する場合は指の静電容量に対応した大きな検出値が得られる。

この投影容量型の原理を用いることで、指の接触だけではなく、指の近接も検知することができる。ただし、ITO透明電極は抵抗成分が大きく静電容量検出の感度が低いため、得られる検出値が非常に小さいという課題がある。そこで、当社ではITO透明電極の感度を向上させる電極の結合制御とシールド制御を行い、他の近接検知用のデバイスを使わずに、指の近接と接触の両者を検知できる3Dタッチパネルを実現した。

3. 3Dタッチパネル

今回試作した3Dタッチパネルの構成を図3に示す。このシステムは、近接した指の検出を実現するために、高感

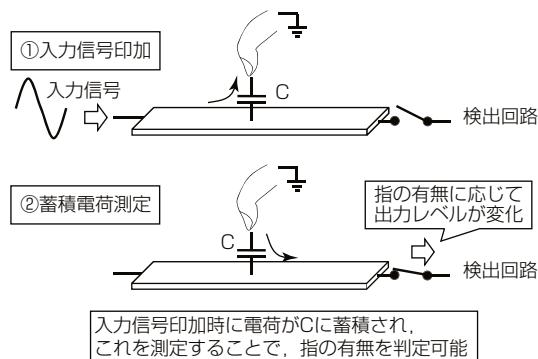


図2. 指の静電容量の検出方法

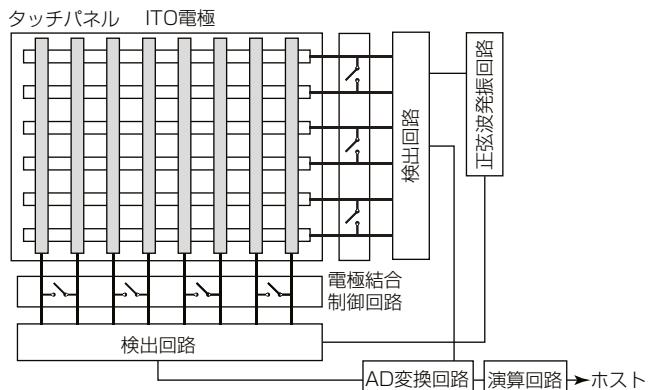


図3. 3Dタッチパネルの構成

度化処理(電極結合制御), ノイズ低減処理(シールド電極制御・デジタルフィルタ)を行う。さらに、近接した指の位置精度を向上させるために、3次スプライン補間による座標算出処理を行う。

3.1 電極結合制御

電極と指との間の静電容量の大きさは、二つの間の面積に比例し、距離に反比例する。したがって、電極の面積が大きいほど指との間に生じる静電容量は大きくなり、指を近接させた場合の静電容量変化の検出感度を向上させることができる。そこで、指の近接を検知する場合は、図3の電極結合制御回路部分で、隣り合う2本の電極を結合して1つの大きな電極として動作させる。これによって、タッチパネル面と指との間に生じる静電容量を感度良く検出することが可能となる。

さらに、タッチパネル面に指が接触した場合には、電極を分離することで、通常のタッチパネルの指検出と同等な分解能を実現する。これによって、近接時の高感度化と接触時の高分解能を両立できる。

図4に電極の動的な結合制御の概念図を示す。

3.2 シールド電極制御

今回試作したタッチパネルのように複数の電極がある場合、指近接を検知している電極(検出電極)と、周囲の電極の間に電位差があると寄生容量が生じる(図5のCf1, Cf2)。この場合、検出電極から寄生容量Cf1, Cf2を介して周囲の電極に信号が流れ込むため、本来計測したい、指近接による静電容量Cs分に起因する検出値は低下する。指と電極の間の静電容量を感度良く検出するためには、この寄生容量を抑制することが重要である。そこで、検出電極と周囲の電極が同電位となるようシールド電極制御を行う。図5に示すように、X軸方向検出用電極を検出電極とした場合、それと直交するY軸方向検出用の電極には、検出電極に流

れている検出信号と同じ振幅・位相を持つ信号を印加する。これによって、検出電極と周囲の電極の電位差が低下し、寄生容量による感度低下を抑制できる。

3.3 デジタルフィルタによるノイズ除去

検出された出力信号には、図6で示した直交する電極の寄生容量以外のノイズも含まれる。特に指の近接を検知する場合は、微弱な静電容量変化をとらえる必要があるため、信号中のノイズ成分の除去が必要である。図6に予備実験で得られた指近接時の検出電極の出力信号を示す。指近接の信号には高周波ノイズが重畠していることが分かる。このため、ローパスフィルタによる高周波ノイズ除去を行う。フィルタには色々な種類があるが、タッチパネルの応答性能をできるだけ落とさないよう、遅延の少ないIIR (Infinite Impulse Response: 無限インパルス応答) フィルタを採用した。

3.4 座標算出

ノイズ除去後の出力信号から座標を算出する。指が近接している場合は電極結合によって座標分解能が低下するため、各電極の出力信号の分布から座標を補間することで座標分解能の向上を図る。図7のように、各電極の出力値を滑らかに結ぶ三次元スプライン曲線を求め、その曲線が極大値となる位置を補間座標値とする手法⁽²⁾を用いた。

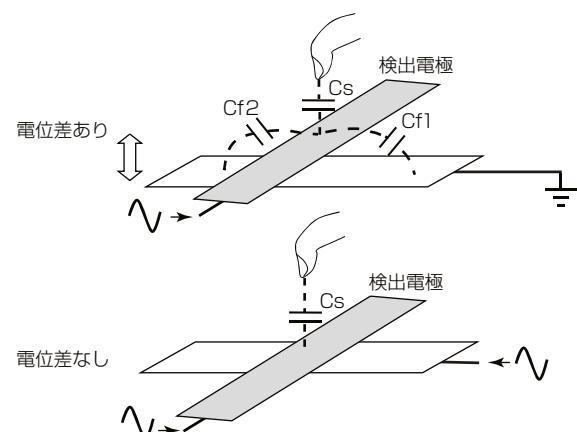


図5. シールド電極制御

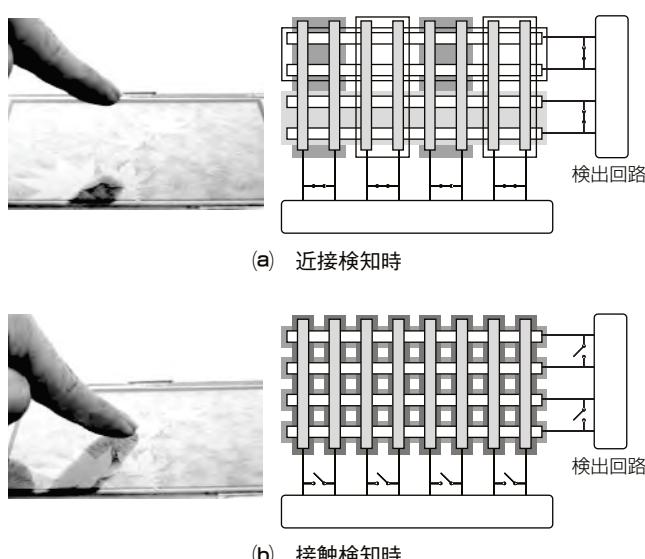


図4. 電極の動的な結合制御

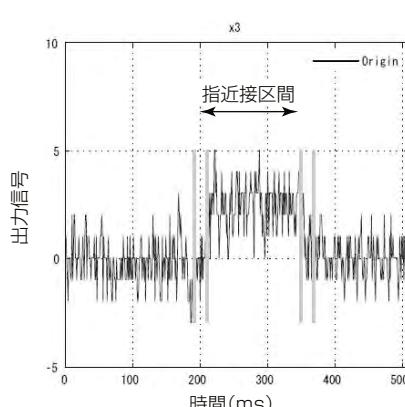


図6. 指近接時の出力信号

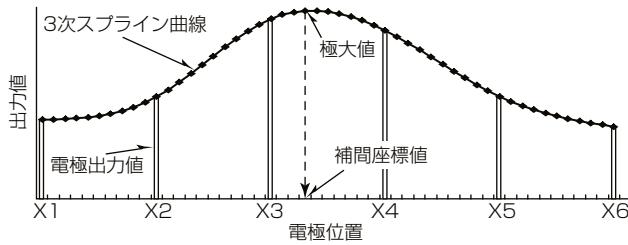


図7. 3次スプライン曲線による座標補間

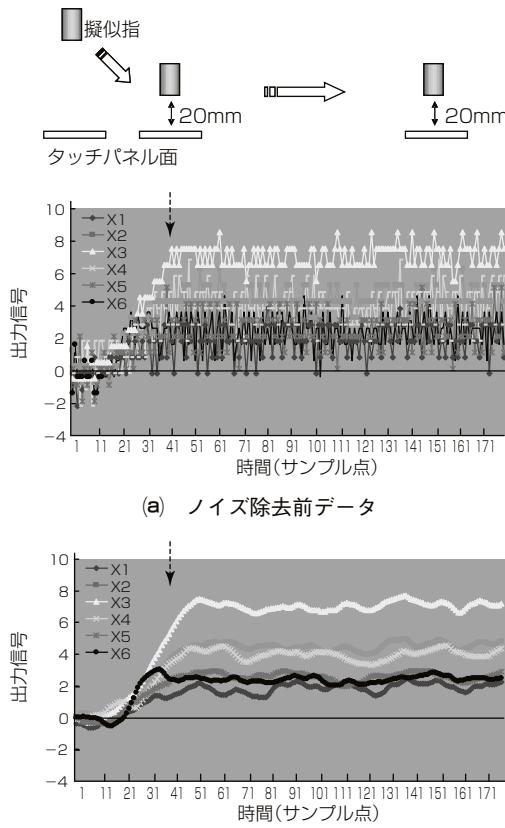


図8. 指20mm近接時の電極出力値

4. 実験

まず、タッチパネル上20mmの高さにある指の近接を識別することができるか確認するため、出力信号を評価した。図8は、タッチパネル上のX軸方向電極(X3)に、タッチパネル面上50mmの位置から20mmの位置まで擬似指(人間の指に近い特性を持つ金属性の円柱)を近接させた場合の、X1からX6までの電極の出力信号について、ノイズ除去前と、ノイズ除去後のデータを示したものである。図中の点線矢印は、擬似指が20mmの位置に達した時刻を示す。ノイズ除去後のデータでは20mm近接時に、周囲の電極(X1～X2, X4～X6)に比べて、指を近接させた電極(X3)の出力値が明らかに大きいことが分かる。この結果、閾値(しきいち)処理によって20mmの指の近接が検知可能と判断できる。

次に、近接・接触時の検出信号から算出した座標値の精

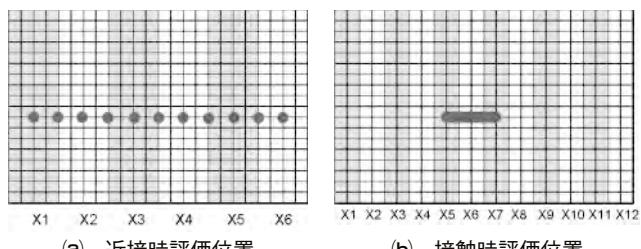


図9. 座標精度評価位置

表1. 座標誤差評価結果

	平均誤差	最大誤差
近接	18pixel(3.2mm)	37pixel(6.6mm)
接触	9.4pixel(1.7mm)	17pixel(3.1mm)

度の評価を行った。図9に座標算出精度を評価するために擬似指を近接・接触させたタッチパネル面上の位置を示す。近接時は、電極幅(結合電極の幅)の1/2ピッチ(約10mmピッチ)ごとにタッチパネル面上20mmの高さに擬似指を近接させて座標算出結果を取得する。また、接触時は、図に示すように2本の電極間を電極幅の約1/54ピッチ(約0.2mmピッチ)ごとに擬似指を接触させて座標算出結果を取得した。

表1に、取得した座標と正解座標との差の平均値(平均誤差)と最大値(最大誤差)を示す。表のpixel単位はVGA (Video Graphics Array)解像度(640×480pixel)で求めたものである。表から、平均誤差は、近接時で3.2mm、接触時で1.7mm程度であり、指入力としては実用的な座標精度が得られていると考える。

5. むすび

ITO透明電極で構成したタッチパネルを用い、静電容量検出の感度を向上させる仕組みを導入することで、指の近接・接触を検知可能な3Dタッチパネルの試作を行った。評価の結果、タッチパネル面上20mmでの指の近接を検知可能であること、近接・接触時に実用的な精度で座標検出が可能であることが分かった。今後は、より遠方の指を検知できるように感度を向上させるとともに、3Dタッチパネルを用いて、例えば、指の近接によって操作ボタンを拡大表示し、指の接触によって機能実行するなどのユーザーインターフェースについても検討を行う予定である。

参考文献

- (1) 岡野祐一, ほか:タッチパネルの最新技術動向, 映像情報メディア学会誌, 63, No.8, 1101~1106 (2009)
- (2) Rekimoto, J.: Smart Skin, An Infrastructure for Freehand Manipulation on Interactive Surfaces, CHI2002, 113~120 (2002)

組み込み機器向け文字／ベクターグラフィックス描画回路

加藤義幸*
鳥居晃*
濱田雅樹*

Font and Vector Graphics Engine for Embedded Systems

Yoshiyuki Kato, Akira Torii, Masaki Hamada

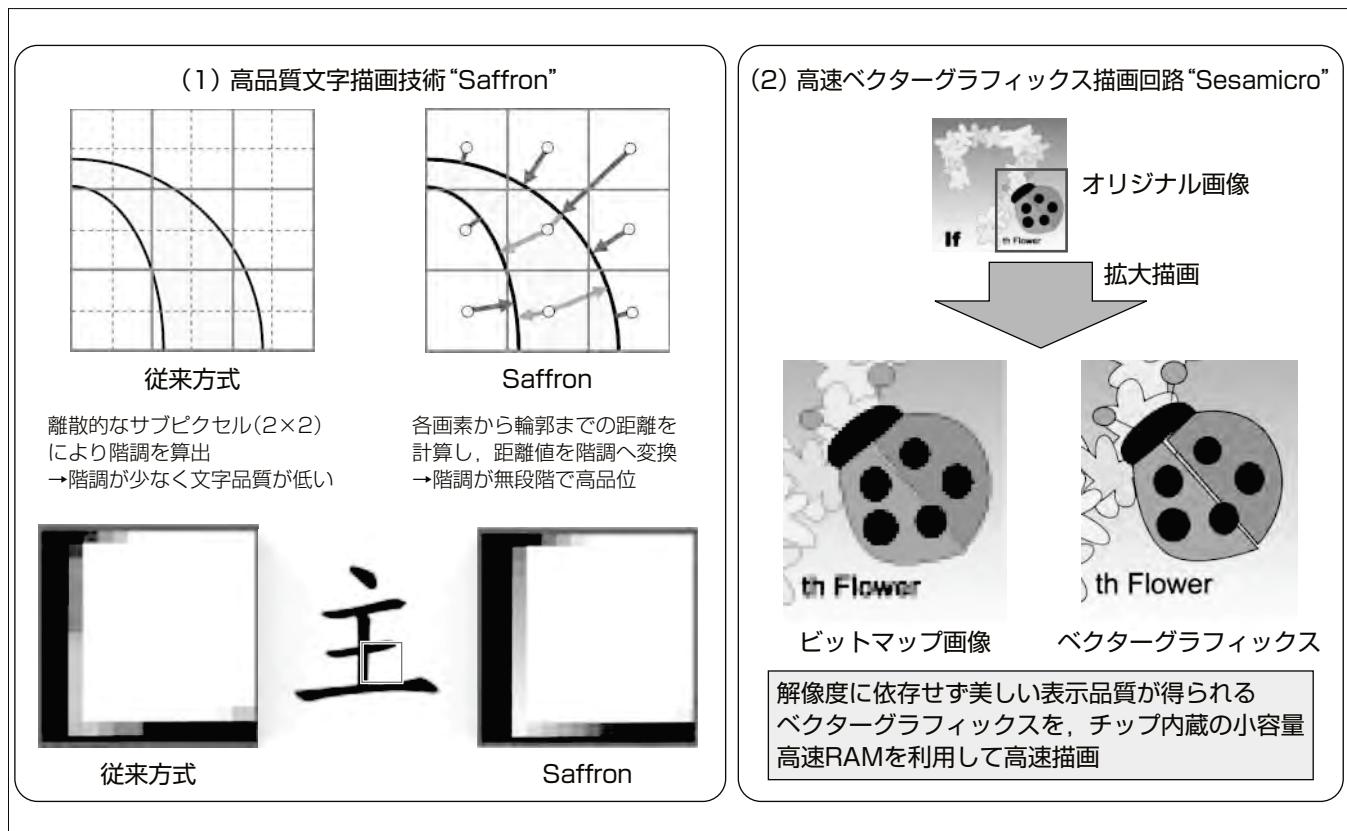
要旨

液晶表示パネルなどによって操作案内を行う機器が様々な分野で増え、高性能なパソコン画面と同等の滑らかで視認性の高い表示や、素早い反応が求められている。

これら組み込み機器でパソコンと同等な表示画面を実現するには、CPU(Central Processing Unit)やGPU(Graphics Processing Unit)にもパソコン並みの高い処理能力が要求される。ところが組み込み機器ではこれらの処理能力は装置の大きさや消費電力によって制限されるため、文字や画像は、能力の低いCPUでも処理可能なビットマップの単純なコピー処理によって表示していた。このため、様々なデザインの図形や文字を、綺麗(きれい)に任意のサ

イズで表示する場合、反応速度が遅くなる問題があった。

今回、組み込み機器への適用が可能なGUI(Graphical User Interface)描画処理用のグラフィックス描画回路を開発した。この回路は、距離場(画素と文字輪郭間の距離情報)を用いた文字描画技術Saffronによって、スケーラブルで高品質な文字を高速に描画する。また、解像度に依存しない美しい表示が得られるベクターグラフィックスを、チップ内蔵の小容量RAM(Random Access Memory)を用いた描画技術によって高速に描画する。これらはCPUに負荷をかけることなく描画を実行できるため、処理能力に制限のある組み込み機器でも高速表示を実現した。



高品質文字描画技術“Saffron(サフロン)”と高速ベクターグラフィックス描画回路“Sesamicro(セサミクロ)”

CPU能力に制限のある組み込み機器でも高性能パソコンと同等の見やすい表示画面が求められていることから、高品質文字描画と高速ベクターグラフィックス描画の回路を開発した。文字描画には距離場を用いたアルゴリズムを適用し、高品質文字の高速描画回路を実現した。ベクターグラフィックス描画は、チップ内蔵の高速RAMを利用するコンパクトな回路による高速描画を実現した。

1. まえがき

近年、液晶表示パネルなどによって操作案内を行う機器が増え、最新のパソコン画面同等の滑らかで視認性の高い表示や、素早い反応が求められている。

しかし、組み込み機器では装置の大きさや消費電力の制約等によってCPUやGPUの処理能力が制限されるため、文字や画像は能力の低いCPUでも処理可能なビットマップの単純なコピー処理によって表示していた。このため、様々なデザインの図形や文字を、綺麗に任意のサイズで表示する場合、反応速度が遅くなる問題があった。

今回、三菱電機は組み込み機器への適用が可能なGUI描画処理用の小型高速回路を開発した。本稿では、この小型高速回路の開発の中で実現した文字／ベクターグラフィックス描画方式について述べる。

2. 文字描画技術

2.1 距離場を用いた描画アルゴリズム

スケーラブルフォントを描画する際には、アンチエイリアス処理を行うために、輪郭部分の画素ごとに描画する割合(カバレッジ値)を求める必要がある⁽¹⁾。今回、カバレッジ値の算出には、距離場を用いた高品質文字描画アルゴリズム“Saffron⁽²⁾⁽³⁾”を使用した。距離場とは、各画素と文字形状との最短距離を表すイメージのこと、距離値から各画素のカバレッジ値を算出する。距離場は連続した数値情報であるため、一般的なサブピクセル分割による離散的なカバレッジ値算出と比較し、文字の輪郭部分を多階調で表現可能である。このため、高品質な文字表示が可能で、視認性を向上できる。

また、図1に示すように、距離値からカバレッジ値を算出する際のパラメータの変更によって、1種類のフォントデータから太らせた文字やぼかした文字の描画ができる。この機能の利用によって、1種類のフォントデータから様々な修飾文字の描画が可能となる。また、Saffronでは、アウトラインフォントに加えて、骨組みフォントであるストロークフォントも描画可能なため、フォントのデータサイズをアウトラインフォントの約1/4へ削減できる。これによって、多国語環境におけるメモリ容量を小さくしてチップコストを抑えることが可能となる。



図1. 距離場を用いた文字描画処理

2.2 文字描画方式

文字描画回路の構成を図2に示す。まず、座標変換処理では、指定されたフォントデータを読み出して、拡大・縮小・回転等の変換処理を行い、文字形状を任意のサイズへ変換する。曲線分割処理では、曲線を直線の集合体として描画を行うために、曲線形状に合わせて微小線分へ分割する。分割処理は、図3に示すとおり、頂点間の距離や曲率によって評価値を計算し、最適な分割数を決定する。

次に、ラスタライズ処理では、線分データから画素を生成し、距離バッファへ描画を行う。距離バッファは、距離場情報を格納するメモリで、これによって文字輪郭を無段階の階調で表示できる。

通常、直線とスキャンラインの交点座標の算出には、除算処理が必要である。しかし、一般の反復法を用いた除算回路では、回路規模が大きくレイテンシ(遅延)も長い課題があった。そこで、交点を算出する特別な除算回路(シフトアップ型除算回路)の新たな開発によって、表1に示すとおり回路規模を約1/5に削減し、レイテンシを約1/3に短縮した。

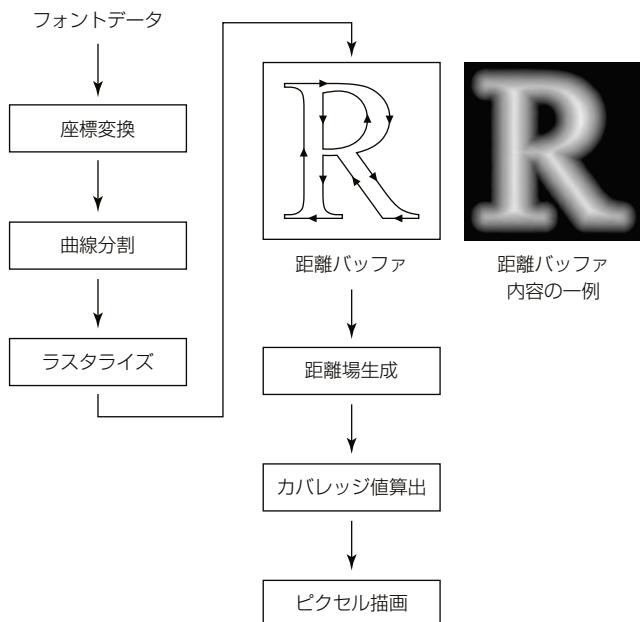


図2. 文字描画回路の構成

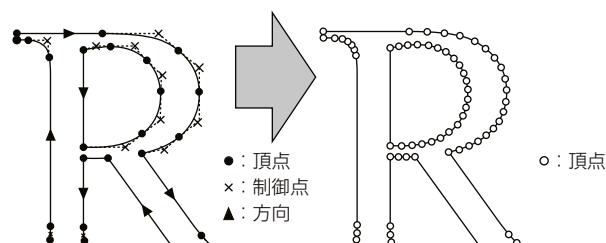


図3. 適応型曲線分割

表1. 除算回路比較

	従来方式	新方式
除算法	反復法	回復法
構成要素	乗算器	加算器
補間処理	なし	あり
回路規模	120Kゲート	25Kゲート
レイテンシ	10クロック	3クロック

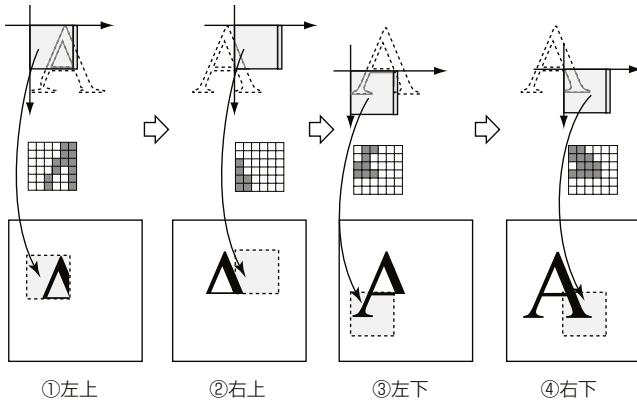


図4. 文字のタイル形状分割描画

また、図4に示すように、ラスタライズをタイル上に分割して処理することによって、交点座標のリスト作成とソーティング処理を不要とした。さらに、タイル処理に必要なメモリに、チップ内蔵された高速で小容量なRAMを用いて、小回路規模での高速処理を実現した。内蔵RAMは、2ポートメモリ構成として、描画処理と外部メモリへのライトバック処理を並列に動作させる。これらによって、外部メモリへのアクセスを最小限に抑え、消費電力の低減が可能となった。

3. ベクターグラフィックス描画技術

3.1 ベクターグラフィックス処理

今回開発した高速ベクターグラフィックス描画回路“Sesamicro”では、図5に示すように、文字描画回路と同様、座標変換、曲線分割、ラスタライズ、カバレッジ値算出の基本処理を実行する。その後、カラー演算処理であるグラデーション、テクスチャマップ、アルファブレンドの処理を実行する。

グラデーションは、色の制御点情報に基づきカラー値の補間処理を行って、なだらかな色を生成する。テクスチャマップは、図形に対してテクスチャイメージの貼り付け処理を行う。テクスチャイメージはビットマップ画像で指定し、 3×3 の変換行列を用いて変形し、奥行き方向の透視補正を施す。また、色をカラー変換行列で変換するカラートランスマスクも実行する。

3.2 アニメーション方式

ベクターグラフィックスをアニメーション表示させるためには、フレームごとに異なる画面を描画して表示を切り

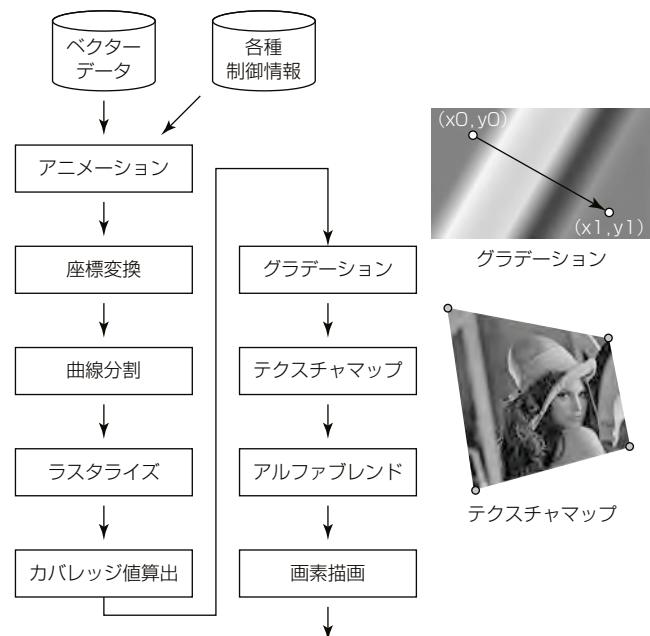


図5. Sesamicroの処理フロー

替える必要がある。一般的なベクターグラフィックスのGPUでは、イミーディエイト方式が採用されるため、文字や図形描画に必要なデータをCPUが毎フレームGPUへ転送しなければならない。このため、CPUの処理負荷が高くなり、本来のCPU処理に悪影響が生じる。

一方Sesamicroでは、アニメーション自動実行機能によって、CPUに処理負荷をかけずにアニメーション表示を実行する。Sesamicroは、外部メモリに格納された一連のベクターデータやアニメーションデータ等を読み込み、描画処理を実行する。さらに、描画終了後、垂直同期信号に合わせてフレームの切替えを自動的に行う。

アニメーションデータは、前フレームと後続フレームの差分情報で定義し、画面のパーティデータに対する座標変換行列の指定によって、移動・拡大・縮小・回転のアニメーション表示を行う。画面制御情報には、パーティデータの画面配置情報などを記述し、複数の独立したパーティを一つの画面上に柔軟に組み合わせて表示できる。

アニメーションデータは、図6に示すように、アニメーションコンバータによって、SVG(Scalable Vector Graphics)データや作画ツールで作成した出力データを、Sesamicro用のディスプレイリストデータへ変換して作成する。

これらによって、画面デザインを一般のツールを用いて開発でき、CPUに処理負荷をかけずにアニメーションの実行ができる。また、解像度に依存しないベクターグラフィックス方式を用いるため、表示機器の画面解像度が変更されても、その都度、画面デザインを作り直す必要がなく、開発コストを低減できる。

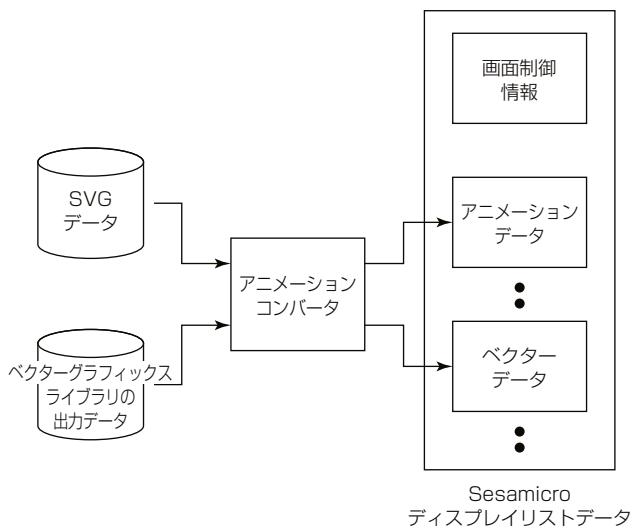


図6. アニメーションデータ作成の流れ

表2. FPGA評価基板の仕様

CPU	PowerPC ^(注1) 66MHz
FPGA	Xilinx Spartan3 ^(注2) 66MHz
外部メモリ	SDRAM 64ビット 66MHz
ピクセル形式	16ビットカラー
表示モニタ	1024×768(文字描画) 1280×768(アニメーション)

(注1) PowerPCは、IBM Corp. の登録商標である。

(注2) Spartanは、Xilinx, Inc. の登録商標である。

SDRAM : Synchronous Dynamic Random Access Memory

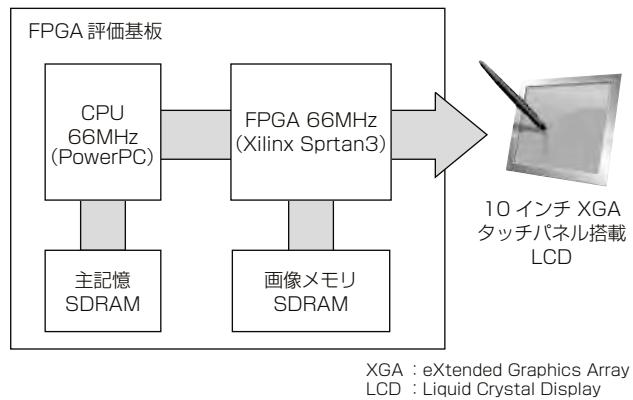


図7. FPGA評価基板の構成

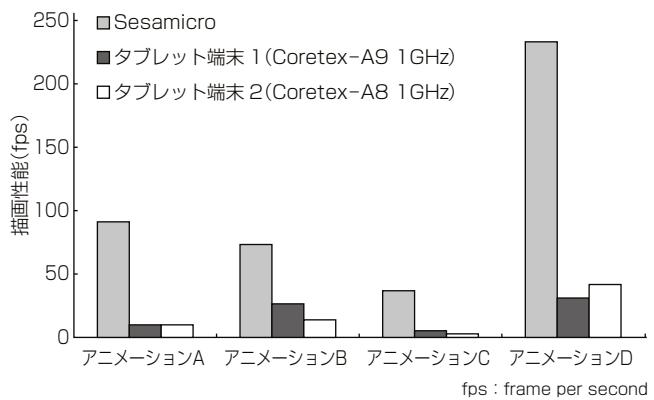


図8. ベクターグラフィックスのアニメーション描画性能比較

4. 性能評価

今回開発した文字描画回路とベクターグラフィックス描画回路をFPGA(Field Programmable Gate Array)に実装し、性能評価を実施した。FPGA評価基板の主な仕様を表2に、構成を図7に示す。

文字描画性能は、8万文字/秒(日本語フォント、48PPEM(Pentium Pro Equivalent MHz))となり、画面全体に小さな文字を表示した場合でも、毎秒60フレームの描画性能を実現した。

一方、ベクターグラフィックスによるアニメーション性能評価は、Flash^(注3)のSWF(Small Web Format)をコンバートしたデータによるSesamicroのアニメーション描画性能と、ARM製Coretex^(注4)-A8及びCoretex-A9と汎用GPUを搭載したタブレット端末のFlash描画性能との比較によって行った。この結果、図8に示すように、Sesamicroはベクターグラフィックスを効率良くハードウェア処理することによって、タブレット端末に比較し、約3～9倍高速であることを確認した。

(注3) Flashは、Adobe System, Inc. の登録商標である。

(注4) Coretexは、ARM Ltd. の登録商標である。

5. むすび

本稿では、GUIを高速に描画する文字描画及びベクターグラフィックス技術について述べた。CPUに負荷をかけずに高速描画が可能で、回路規模も小さいため、組み込み向け表示装置に、低コストで搭載が可能である。

この技術は、鉄道表示器や車載表示器等への適用を行っており、さらに、今後は他の組み込み機器への適用も予定している。

参考文献

- 加藤義幸, ほか:組込み向けスケーラブルフォント描画回路, 電子情報通信学会総合大会講演論文集, C-12-28 (2009)
- Friskin, S. F., et al.: Adaptively Sampled Distance Fields: A General Representation of Shape for Computer Graphics, ACM SIGGRAPH, 249～254 (2000)
- Friskin, S. F., et al.: Designing with Distance Fields, ACM SIGGRAPH, 60～66 (2006)

菅沼優子*
根岸博康*
川又武典*

図形のデフォルメ技術

Deformation Technology for Graphics

Yuko Saganuma, Hiroyasu Negishi, Takenori Kawamata

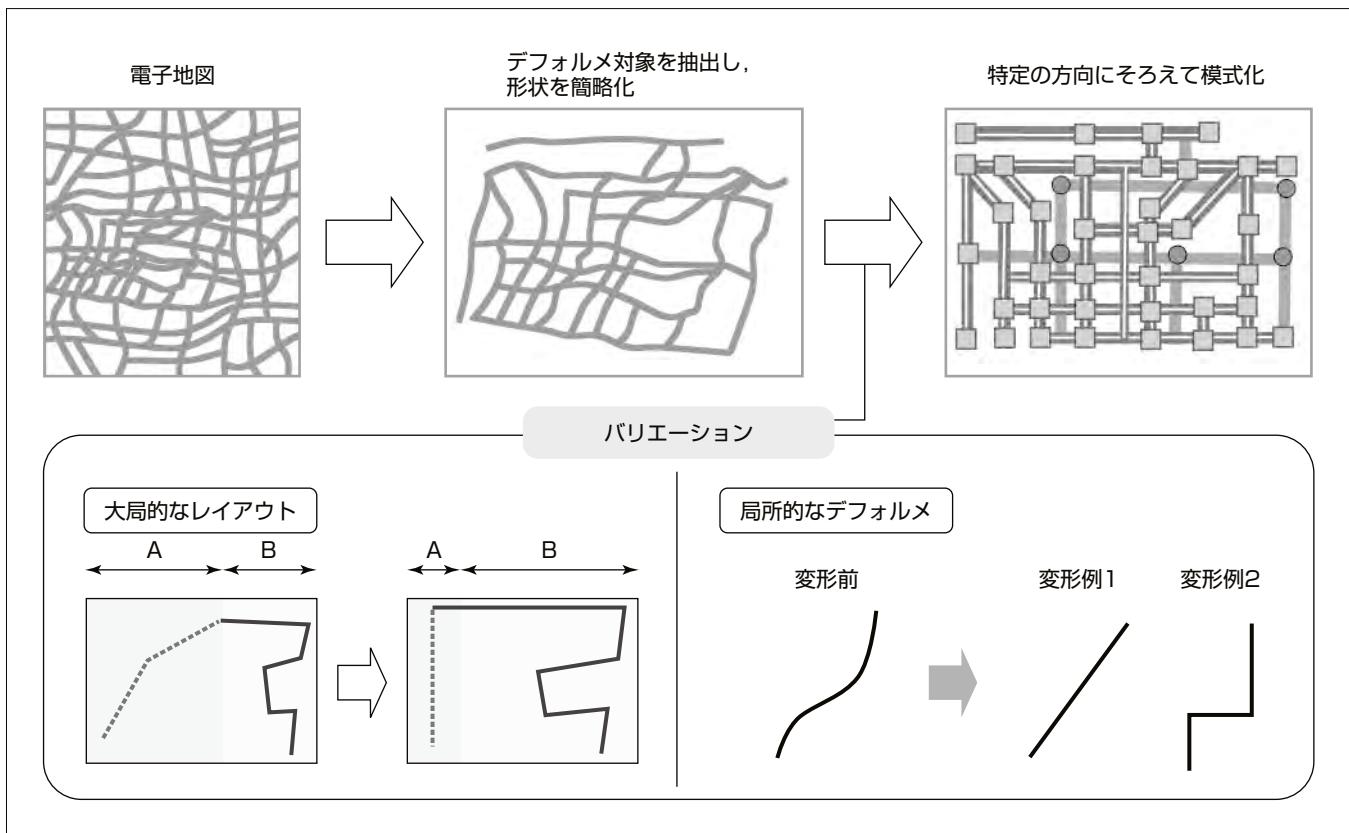
要旨

情報の多様化が進み、情報量が増加の一途をたどっている昨今では、膨大な情報の中から必要な情報を抽出し、利用者にとって分かりやすい表現で提供することが求められている。図形のデフォルメ技術は、情報を取捨選択し、分かりやすく表示する技術として注目されている。店舗への案内図や鉄道路線図等、日常でも広く活用されているが、現状、デフォルメ地図の多くはデザイナーによって作成されており、製作に多くの時間と費用を要するという課題がある。また、デフォルメ地図を継続的に使用するシステムでは、使用しているデフォルメ地図に情報の追加や削除を行うことが考えられるが、デザイナーが行う場合は手作業での改訂となり、短期間に実施することが困難である。

一方、デフォルメ地図を自動生成する試みもなされているが、デフォルメ対象の増加に伴い、膨大な処理時間を要するという課題がある。このため、現状では大規模なデフォルメ地図を容易に生成することが困難である。

このような課題に対して、電子地図から広域道路網のデフォルメ地図を容易に生成する方式を開発した。道路の形状を大きく変形させる大局的な座標変換を備え、形状の簡略化度合いや道路を配置する方向等を局所的に制御可能な幾何学的アプローチをとることで、処理時間の問題を解決した。

本稿では、開発したデフォルメ地図自動生成方式について述べる。



デフォルメ地図自動生成方式

電子地図からデフォルメ対象の道路を抽出し、形状の簡略化を経て、模式化したデフォルメ地図を自動生成する。大局部的レイアウトと局所的デフォルメによって、デフォルメのバリエーションを広げる。大局部的レイアウトでは、デフォルメ地図全体における領域の配置を制御し、局所的デフォルメでは、個々の線分の配置方向を制御する。両者を組み合わせることで、雰囲気の異なるデフォルメ地図を容易に生成することができる。

1. まえがき

情報の多様化が進み、情報量が増加の一途をたどっている昨今では、膨大な情報の中から必要な情報を抽出し、利用者にとって分かりやすい表現で提供することが求められている。図形のデフォルメ技術は、電子地図などの詳細な情報の中から情報を取捨選択し、図形の形状を簡略化するなどして見やすくする技術であり、必要な情報を分かりやすく表示する技術として注目されている。鉄道路線図など、日常でも広く活用されているが、現状、デフォルメ地図の多くはデザイナーによって作成されており、製作に多くの時間と費用を要するという課題がある。また、使用しているデフォルメ地図に対して、道路などの追加削除を行う場合、又は、利用者の要望に応じて図形の形状を変更する場合には、現在使用しているデフォルメ地図を改訂する必要がある。デザイナーが行う場合は手作業となるため、短期工程で実施することが難しい。こうした背景の中、電子地図からデフォルメ地図を自動生成する技術を開発した。

本稿では、広域道路網に対応したデフォルメ地図自動生成方式について述べる。

2. デフォルメ地図自動生成における技術課題

今回開発したデフォルメ地図自動生成方式では、道路をデフォルメ対象とした。そこで、ここでは、地図の例として道路を用いて述べる。

デフォルメ地図の自動生成は、図1に示すように、交差点をノード、道路をエッジとするグラフデータに対して、道路の接続関係が分かりやすくなるようにエッジを変形させるグラフレイアウト問題である。グラフレイアウト問題では、一般的に、エッジを表示させる方向(以下、“エッジの配置方向”といふ)が水平、垂直、斜め45度のいずれかの方向となるようにエッジを変形する。しかし、従来の主な手法である最適化の手法⁽¹⁾では、ノード数の増加に伴い、変形後の新しい座標を求めるのに膨大な計算時間を要するという課題があり、大規模なグラフデータへの適用が困難である。また、従来手法では、一般的に、電子地図を変形時の元地図とし、元地図に対してエッジの配置方向が45度

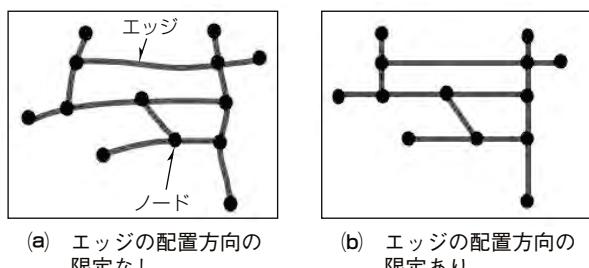


図1. ノードとエッジ

の整数倍の方向となるようにエッジを変形する。しかし、デフォルメ地図を利用する用途によっては、エッジの配置方向を特定の方向に限定して模式化するだけでなく、デフォルメ地図を全体としてとらえた際に、利用者にとって重要な部分を分かりやすくしたいという要求がある。

このような要求に対して、従来手法では、エッジの配置方向を特定の方向に限定させることはできるものの、利用者の思い描いている形状に近いデフォルメ地図を自動生成することが困難である。

3. 目的

2章で述べた課題に対して、この開発では、広域で多数の道路を含む場合でも比較的短時間でデフォルメ地図を生成できること、利用者がほしいと思う形状に近いデフォルメ結果を得られるように、雰囲気の異なるデフォルメ地図を容易に生成できることを目的とした。

具体的には、数十キロメートルといった広範囲を対象とし、多数の道路をデフォルメできること、パラメータを変更することによって、印象の異なるデフォルメ地図を容易に生成できることを目的とした。

4. 広域道路網のデフォルメ地図自動生成方式

4.1 利用者の望む内容に近づけるための指標

2章で述べたように、用途によっては、デフォルメ地図全体で重要な部分を見やすく強調したデフォルメ地図を生成する必要がある。このように特定の部分を見やすくするデフォルメでは、デフォルメ結果の良し悪(あ)しを所定の方向に配置できたエッジの個数のみで評価するのでは不十分であり、利用者が想定するデフォルメ地図に近い結果が得られたかどうか、利用者側から評価する必要がある。

こうした主観評価に依存する部分を工学的な指標として自動生成方式に組み入れるために、今回開発した方式では、利用者が意図する形状のデフォルメ地図に近づけるため2つの指標を設定した。“大局的なレイアウト”と“局所的なデフォルメ”である。

“大局的なレイアウト”は、デフォルメ対象とする領域を、デフォルメ結果を出力する領域に対してどのようにマッピングさせるかを規定する。具体例を図2に示す。デフォルメ結果を出力する領域は、ディスプレイの解像度や紙面の縦横サイズから定まる。図では、デフォルメ地図を出力する紙面のサイズに合わせて縦横比を変更し、重要視したい領域であるAの領域がデフォルメ地図全体の中で大きな領域を占めるように領域Aに含まれる道路の配置位置を変更している。重要視したい領域がデフォルメ地図全体でどの辺りにどの程度の大きさで表示されるのかを利用者の要望に一致させることによって、限りなく存在するデフォルメのバリエーションに対して利用者の意図に近い内容を得る

ことができると思った。

もう一つの指標である“局所的なデフォルメ”では、デフォルメ度合いを規定する。電子地図を入力として道路の方向が水平、垂直、斜め45度の方向となるように変形する場合、元の地図から道路の方向を大きく変えずに変形する場合と、道路の接続関係をより一層見やすくするために、元の地図の道路方向から変えた方向に配置してよいとする場合とでは、得られるデフォルメ地図の雰囲気が異なる。例を図3に示す。階段状の道路に対して、例1は元の道路方向に近い方向で変形した例であり、例2は元地図で右斜め方向の道路を垂直とみなして変形した例である。このように、個々の道路の方向をどのような方向に配置させるかによっても、デフォルメ地図の印象は変化するため、デフォルメ地図の雰囲気を局所的に制御する指標として、道路形状のデフォルメ度合いを定め、パラメータとして変更できるようにした。

このようにして、“大局的なレイアウト”と“局所的なデフォルメ”的2つの指標を軸としたデフォルメ地図生成方式とすることによって、利用者が思い描く形状のデフォルメ地図に近い結果を得られるようにした。

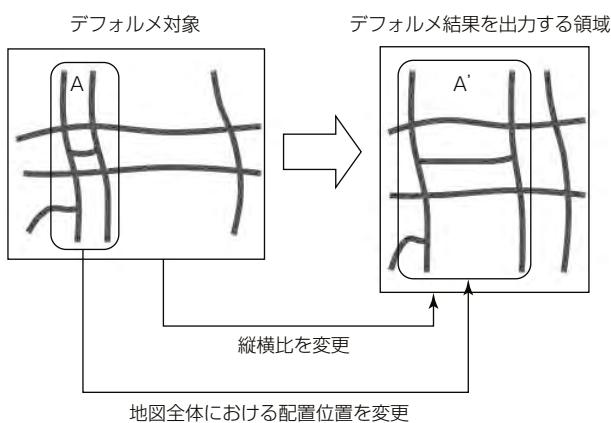


図2. 大局的なレイアウトの変更

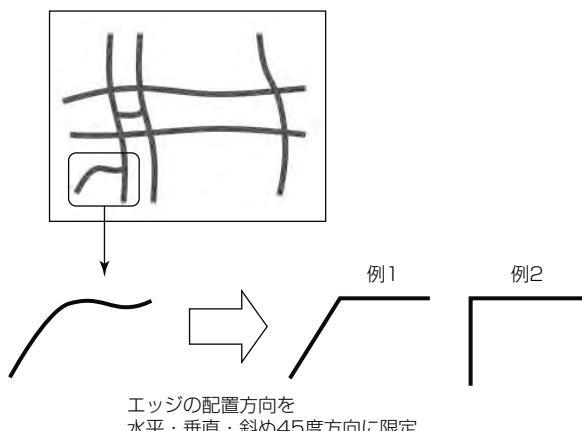


図3. 局所的なデフォルメ

4.2 大規模データに対応可能な自動生成

2章で述べたように、最適化の手法を用いる従来の自動生成手法では、多数の道路を含む広域のデフォルメ地図を生成することが難しい。今回の開発では、道路の方向を45度の整数倍の方向にする理論上の最適解を得ることを目的とせず、利用者の意図に近いデフォルメを実現することを目的とするため、4.1節で述べた2つの指標に基づく幾何学的なアプローチを採用することで、大規模データに対する処理時間の問題を解決した。

デフォルメ地図生成の処理の流れを図4に示す。まず、道路の電子地図データを入力し、デフォルメ対象とする道路を抽出する。次に、抽出した道路の形状を直線化して簡略化する。この処理は、局所的なデフォルメに関する処理であり、簡略化の度合いをパラメータで変更できる。続いて、簡略化したデータに対して大局的な座標変換を施し、出力するデフォルメ地図の縦横比やどのようなレイアウトでデフォルメ地図を出力したいかに応じて、道路の形状を大きく変形させる。この処理は、大局的なレイアウトに関する処理であり、対話的な操作によってレイアウトを変更することができる。最後に、個々の道路の配置方向が45度の整数倍の方向となるように、個々の道路を構成するノードについて移動先の座標を算出し、エッジを変形する。この処理も、4.1節で述べた2つの指標のうち、局所的なデフォルメに関する処理である。なお、移動先の座標を算出する際は、移動の前後で付近の道路との交差関係や交差点に接続している道路との位置関係が変わらないような座標を算出する。

5. 適用例

開発した方式を用いて、電子地図データからデフォルメ地図を自動生成するシステムを試作し、比較的広い範囲のデフォルメ地図を自動生成する実験を行った。

道路の電子地図から任意の道路をデフォルメ対象として抽出した後、道路の形状を簡略化した結果を図5に示す。東西方向に約34km、南北方向に約36kmの範囲で、エッジ数は3,441である。電子地図が持つ属性情報を活用するため、属性情報を持つノードをすべて保持した。このため、直線で表現されている道路が多数のエッジから構成されて

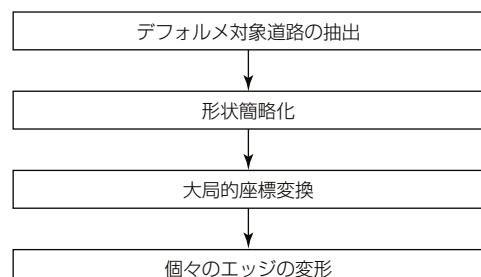


図4. デフォルメ地図の生成手順



図5. 形状を簡略化した地図

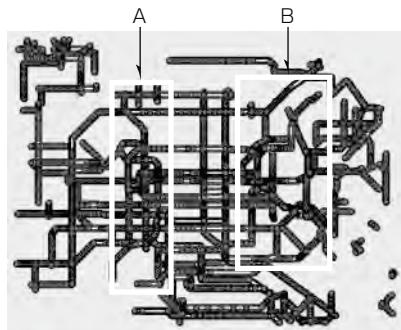


図6. デフォルメ例(大局的座標変換なし)

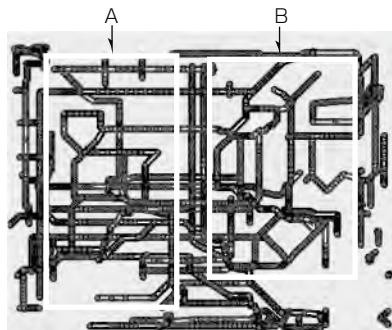


図7. デフォルメ例(大局的座標変換あり)

いる。この簡略化地図に対して、縦横比を3:4とするデフォルメ地図を生成した。図6は大局的な座標変換を行わない結果であり、図7は大局的な座標変換を施した結果である。図7では、AとBの領域がデフォルメ地図全体の中で大きく表示されるように変形した結果である。どちらの場合も、生成に要した時間は約20秒(CPU(Central Processing Unit) 3 GHz, Core2Duo^(注1))であった。

最適化問題として解く従来手法⁽¹⁾では、データ規模の増加に伴い計算時間が増大し、数百のエッジからなる路線図のデフォルメに約20分の時間を要している。今回、2つの指標に基づき利用者の意図に近い解を得る方式とすることで、データ規模が多い場合にも短時間で生成できることを確かめた。

また、図8の簡略化地図に対して、異なるパラメータでデフォルメした例を図9に示す。図9(b)では、水平・垂直方向の道路が多くなるようにパラメータを設定した。図9(a)と図9(b)の比較から、局所的なデフォルメでデフォルメ



図8. 形状を簡略化した地図

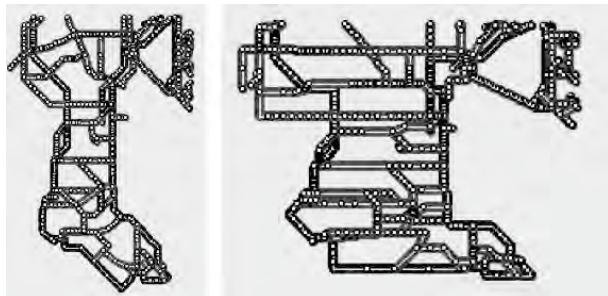


図9. 局所的なデフォルメの効果

度合いを変えることによって、異なる雰囲気のデフォルメ結果を生成できることを確かめた。

このようにして、大局的なレイアウトと局所的なデフォルメ度合いの各々を変化させることによって、デフォルメ地図のバリエーションを広げることができる。

(注1) Core2Duoは、Intel Corp. の登録商標である。

6. む す び

電子地図から広域道路網のデフォルメ地図を容易に自動生成する方式について述べた。

今回開発した方式では、道路の配置を大きく変える大局的な座標変換を備え、形状の簡略化度合いや道路の配置方向等を個々の道路について局所的に制御することで、利用者が意図する内容に近いデフォルメ地図を容易に生成できるようにした。

今後は、利用者が思い描いていたデフォルメ地図に対して、どの程度満足できるデフォルメ地図が得られたかどうかについて被験者評価を行い、自動生成方式に反映していく予定である。

参 考 文 献

- (1) Nöllenburg, M., et al.: A mixed-integer program for drawing high-quality metro maps, In : Patrick Healy and Nikola S. Nikolov, editors, Proc. 13th Internat. Sympos. Graph Drawing(GD'05), volume 3843 of Lecture Notes in Computer Science, 321~333, Springer-Verlag (2006)

画像同定技術：Visual Signature

工藤大樹*
西川博文*

Image Identification Technique: Visual Signature

Daiki Kudo, Hirofumi Nishikawa

要旨

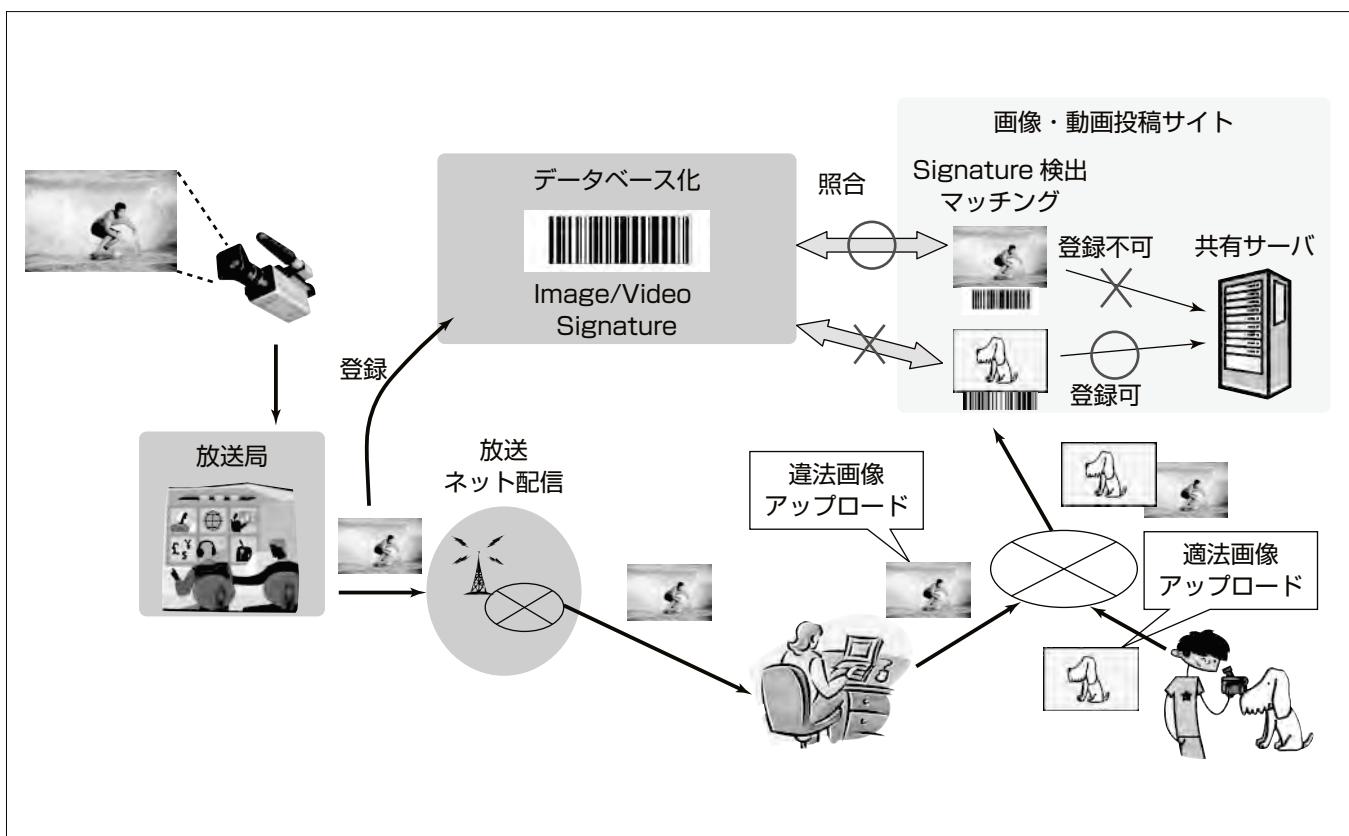
インターネットは、回線の高速化に加え、デジタルカメラ、パソコン、携帯電話の普及によって、画像映像の膨大な流通インフラとなっている。さらに、スマートフォンの普及によって、画像の発信・活用がだれにでも可能な世界が実現されていくだろう。しかし、所望の画像映像コンテンツの効率的な検索は困難であり、有効な検索・管理手法への要求は、さらに高まっている。

そこで、タイトルなどのテキスト情報に加え、画像そのものの特徴量から、画像コンテンツの検索管理を行う方法が検討されてきている。ISO(International Organization for Standardization)/MPEG(Moving Picture Experts Group)では、この画像特徴量の記述方法の標準化を行い、2002年にMPEG-7標準(ISO/IEC 15938シリーズ)を発行し

た。このMPEG-7の拡張として画像の同一性の検出を目的に定義されたのが、三菱電機が標準化に大きく寄与したVisual Signatureである。

Visual Signatureは、画像の空間的な特徴を効率的に表現するもので、回転・拡大・縮小等の画像編集後も同一性的判断が可能である。部分切り出しへの耐性を持つ静止画向けのImage Signatureと、時間軸上での編集への耐性を持つ動画像向けのVideo Signatureがある。

この技術によって、違法画像コンテンツの検出など、画像コンテンツ管理のほか、実世界の画像からインターネット上の同一画像を検索する画像検索型のネットショッピングや、ユーザーに必要な情報を提供するサービス等、幅広い画像利用サービスが実現できる。



Visual Signatureによる違法コンテンツフィルタリング

あらかじめ、著作権のある画像に対してVisual Signature 特徴量を抽出し、データベース化しておく。ユーザーが画像をアップロードした場合、まずマッチングサーバに送られ、データベース内の画像とマッチングが行われる。その結果、マッチングしなかった、すなわち、著作権ありとして登録されていなかった画像のみ蓄積サーバに登録される。

Visual Signature の高速マッチングによって、実用的な時間で処理可能となる。

1. まえがき

現在、インターネットには多くの画像コンテンツがあふれている。回線の高速化に加え、デジタルカメラ、パソコン、携帯電話等、画像が簡単に扱える機器の普及によって、コンテンツ共有サービスは拡大し、インターネットはマルチメディアコンテンツの流通インフラとなっている。今後、クラウドベースのストレージサービスも普及し、個人のコンテンツもSNS(Social Networking Service)などを通じて共有することが当然となるだろう。スマートフォンやタブレット端末にも、十分なローカルストレージと無線ネットワーク環境が備わり、大量のコンテンツに対して“いつでも”“どこでも”アクセス可能な環境が整ってきている。環境がサービスを普及させ、サービスの普及がさらにサービスを魅力的なものとし、インターネット上のマルチメディアコンテンツは膨大なものになると考えられる。

このような環境では、ユーザーが所望のコンテンツを素早く簡単に見つけること、また、権利侵害など違法利用を防止することは極めて重要であり、これを実現するための画像検索・管理技術への期待はますます高まっている。

画像情報管理では、作成者などのテキストデータに加え、画像そのものの特徴をメタデータとすることが検討されてきた。MPEG-7(ISO/IEC 15938シリーズ)は、MPEG-2/4などの画像圧縮技術と異なり、マルチメディアコンテンツの検索を行うためのメタデータの国際標準であり、検索に用いる特微量及び記述子の標準化が行われてきた。

近年、このMPEG-7で画像の同一性の判定を行うVisual Signature(画像同定)技術の標準化が行われた。この標準化で、当社は、その提案技術が標準方式に採用されるなど、中心的な役割を果たしてきた⁽¹⁾。本稿では、このVisual Signature技術について述べ、利用分野を展望する。

2. Visual Signature

2.1 画像同定

画像共有サイトには、権利者の許可なく画像が投稿、共有される場合がある。これらの画像を容易に検出するのが画像同定技術である。

一般に、デジタルデータの同一性を判断する方法として、SHA(Secure Hash Algorithm)に代表されるハッシュ関数がある。ハッシュ関数とは、デジタルデータのビット列を要約する関数である。内容が1ビットでも異なると、全く異なる値になるため、効率的にデジタルデータの同一性を確認できる。画像データの場合、図1上部に示すように、回転やモノクロ化などの処理を施しても、人間の目には同一画像として認識される。しかし、これらの画像はすべてビット列が異なるため、ハッシュ値が異なってしまう。このように、画像同定では、編集・加工に対する頑健性が要

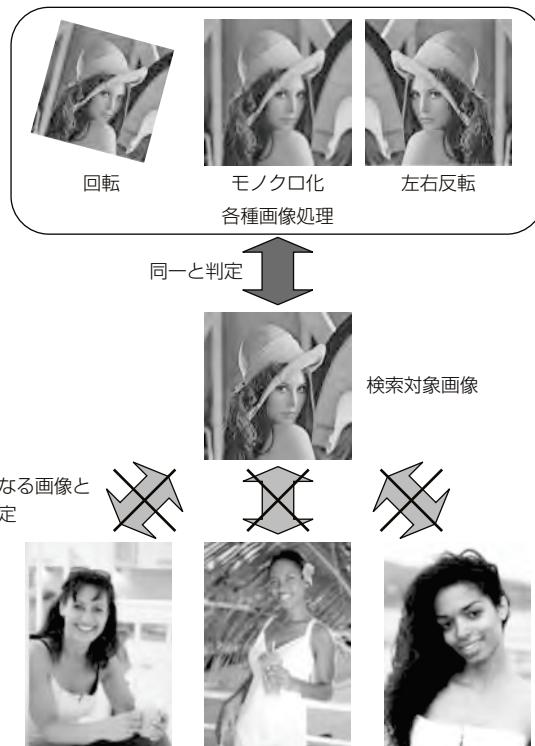


図1. 同一の画像

求され、ハッシュ関数では対応できない。

Visual Signatureは、これらの処理を行っても同一の画像と判断できることを目的として開発された。“女性の画像”を検索するための類似画像検索技術とは異なり、Visual Signatureでは図1下部の画像は別画像と判定されなくてはならない。

2.2 Image SignatureとVideo Signature

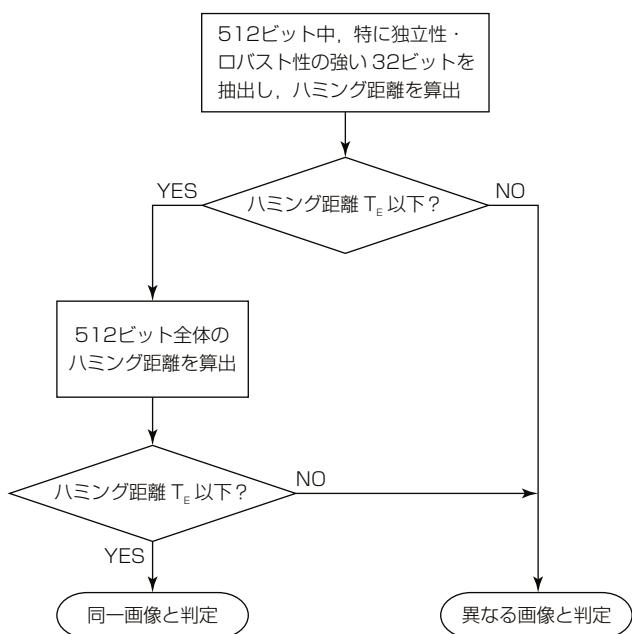
Visual Signatureには、静止画向けのImage Signature⁽²⁾と、動画像向けのVideo Signature⁽³⁾の2種類が存在する。

Image Signatureはトレース変換⁽⁴⁾を基にした特微量であり、2009年に標準化された。輝度情報を空間的周波数で表現しており、特微量間の独立性が極めて高く、ほかの画像を誤って同一と判断するのは100万ペア中で0.05ペアである。また、画像特徴を空間的な分布で表現するため、画像のクロッピング(一部の切り取り)など、大幅な改変処理が施された後でも94%以上を同一と判定する高い頑健性を持つ⁽¹⁾。

一方、動画像データに対する同定を目的としたVideo Signatureは、2010年に標準化された。Image Signature同様、高い独立性・頑健性を持つ。

3. Image Signature

静止画の同定では、画像全体の特徴を高速に比較するGlobal Signatureと、特徴点を抽出し比較するLocal Signatureの2種類のマッチング手法が用意されており、用途に応じて利用可能である。



3.1 Global Signature高速マッチング

Global Signatureは画像全体を512ビットの特微量で表現する。通常、2特微量間のハミング距離(2つのビット列中、対応する位置にある異なるビットの個数)を特微量距離とし、高速なマッチングが可能である。

しかし、比較対象画像数が膨大な場合など、より優れたユーザーインターフェースを実現するために、さらに高速なマッチング方法が考案されている。

その手順を次に示す(図2)。

- ①512ビットの特微量から32ビットを選択する。
- ②選択した32ビット間のハミング距離を算出、閾値(しきいち)判定を行う。閾値よりハミング距離が小さい場合のみ次の③に進む。
- ③従来の512ビットのハミング距離を算出し、閾値判定処理を行う。

すなわち、部分特微量で予備判断をするマッチング手法である。利用する32ビットは、独立性・頑健性の高いものが標準化の過程で実験的に選択、規格化されている⁽⁵⁾。

この高速化手法によって、Global Signatureマッチング速度は7,000万枚/秒(Intel Core2Duo^(注1) 3 GHz使用時)に達し、マッチング精度の低下は1%程度であり、実用上問題ないレベルである。

(注1) Core2Duoは、Intel Corp. の登録商標である。

3.2 Local Signatureマッチング

Local Signatureは、画像の特徴点ごとに局所的なトレース変換を行い、特微量とするものである。図3に特徴点の例を示す。このマッチングでは、特微量距離の近い特徴点が幾何学的に類似しているかどうかを次の手順で判定する。



図3. Local Signatureの特徴点

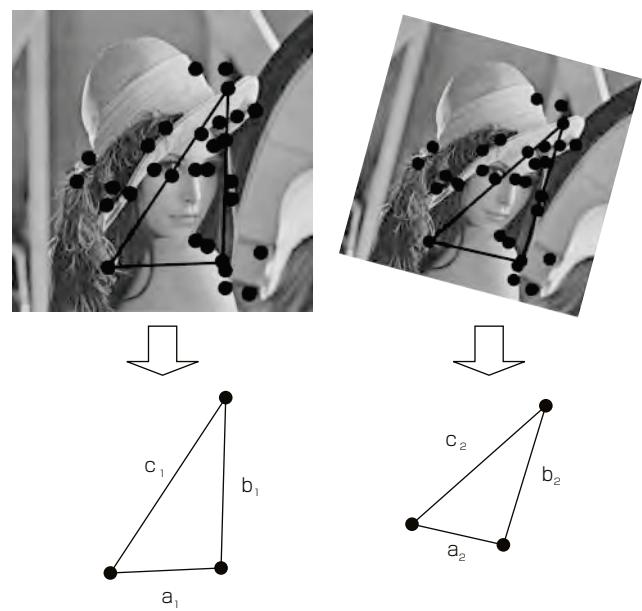


図4. 3特徴点ペアによる幾何学的制限

- ①特徴点ペアを3つ選択する

特微量距離が閾値を下回るペアを3つ選択する。

- ②3特徴点ペアによる幾何学的制限(図4)

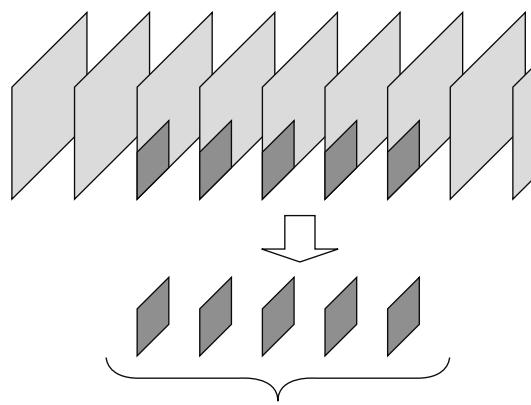
3つの特徴点ペアから三角形を作成する。それぞれの画像中の2つの三角形の辺の大きさ(a₁, b₁, c₁)と(a₂, b₂, c₂)の比を比較し、類似の場合のみ次の③に進む。

- ③4特徴点ペアによる幾何学的制限

特微量距離の近い特徴点ペアをもう一つ選ぶ。4ペアから4つの三角形が得られる。対応する各三角形の面積比を基に最終的なマッチング判定を行う。

このLocal Signatureの採用によって、画像のクロッピングに対する頑健性が強化されている。クロッピング後、画像の幅及び高さが70%以上保持されている場合、94%のマッチング精度を実現している。

ハミング距離を利用するGlobal Signatureに比べてマッチング速度は劣るもの、10万枚/秒と十分に高速である。



Picture-in-Picture の映像のみから抽出

図 5. Picture-in-Picture からの抽出

4. Video Signature

Image Signature同様、各種改変処理が行われた映像に対しても、同一性を保持する特徴量である。図5に示すように、映像中に別の映像を埋め込んだPicture-in-Pictureの子画面部分の検索に対応するため、映像中の特定領域及び特定時間からの抽出も可能である。

Video Signatureは次の3つの要素で構成する。

(1) Frame Signature(画像一枚の特徴量)

Frame SignatureはImage Signatureと同様に画像一枚の特徴量である。映像中の全フレームに付与されることから、Image Signatureに比べて小さな演算量で抽出できる。まず、図6に示すように画像を様々な矩形(くけい)領域に分割する。次に、領域ごとの輝度平均値や、領域間の輝度差分値から380次元のベクトルを得る。取得したベクトルは可変長符号化され、最終的に608ビットの特徴量となる。

(2) Word(Frame Signatureのサブセット)

WordはFrame Signatureのサブセットであり、380次元のベクトルから25個の要素を計40ビットで表現する。

(3) Bag-Of-Words(時間方向の特徴を代表する要素)

Bag-Of-WordsはWordの代表的要素を時間方向にまとめた特徴量である。図7に示すように、90フレームを一つの抽出単位とし、さらに45フレームずつ互い違いに抽出を行う。映像の一部をカットする編集に対してもマッチングが可能である。

Video Signatureでは、はじめにBag-Of-Wordsによって時間方向で一致する位置を見つけ、その後、Frame Signatureによって画像一枚ごとに照合を行うことでマッチングを実施する。

5. む す び

画像が含まれるマルチメディアコンテンツに対する検索技術への要求は高いが、技術的には困難な課題があった。しかし、大量のコンテンツに対して“いつでも”“どこでも”

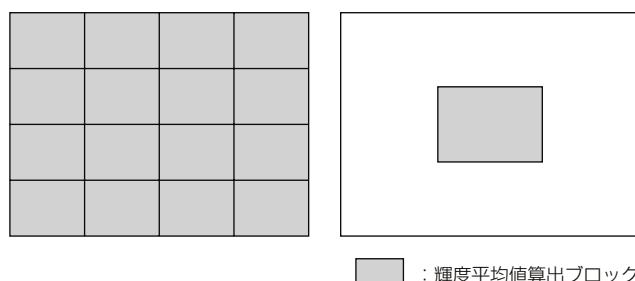


図 6. 領域分割例

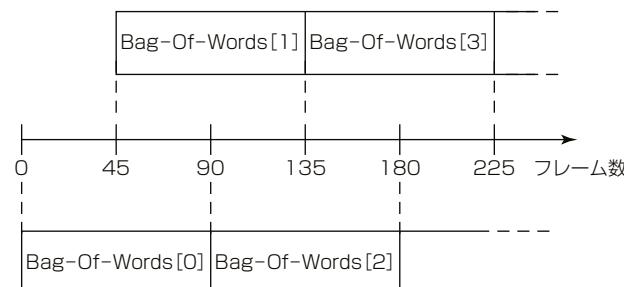


図 7. Bag-Of-Words の抽出

アクセス可能な環境が整い、様々なコンテンツ提供サービスが急速に拡大していく中、画像映像に対する検索・管理の重要性はますます高まっている。

その中で、画像同定方法として当社が国際標準化に大きく寄与したVisual Signatureは、画像の加工編集に対して頑健な特徴を少ない情報量で表現することができ、高精度・高速マッチングを実現している。

また、現在MPEG-7ではCDVS(Compact Descriptors for Visual Search)と呼ばれるモバイル端末向けの画像検索特徴量の標準化が始まっている。

この技術によって、違法コンテンツの検出に留まらず、画像情報からインターネット上の同一画像を検出することで、画像検索型のショッピング、情報検索、さらには、スマートフォンで撮影した看板やロゴなどに関連する情報を提供する情報サービス等、様々なアプリケーションへの適用が可能である。

参 考 文 献

- (1) 西川博文, ほか: MPEG-7技術 – Image Signature : 画像同定技術 –, 三菱電機技報, **82**, No. 12, 763~766 (2008)
- (2) ISO/IEC 15938-3 : 2002/Amd 3 : 2009
- (3) ISO/IEC 15938-3 : 2002/Amd 4 : 2010
- (4) Kadyrov, A., et al.: The Trace Transform and Its Applications, IEEE Trans. Pattern Analysis and Machine Intelligence, **23**, No. 8, 811~828 (2001)
- (5) ISO/IEC TR 15938-8 : 2002/Amd 5 : 2010

モービルマッピング技術

前原秀明*
石渡要介**

Mobile Mapping Technology

Hideaki Maehara, Yosuke Ishiwatari

要旨

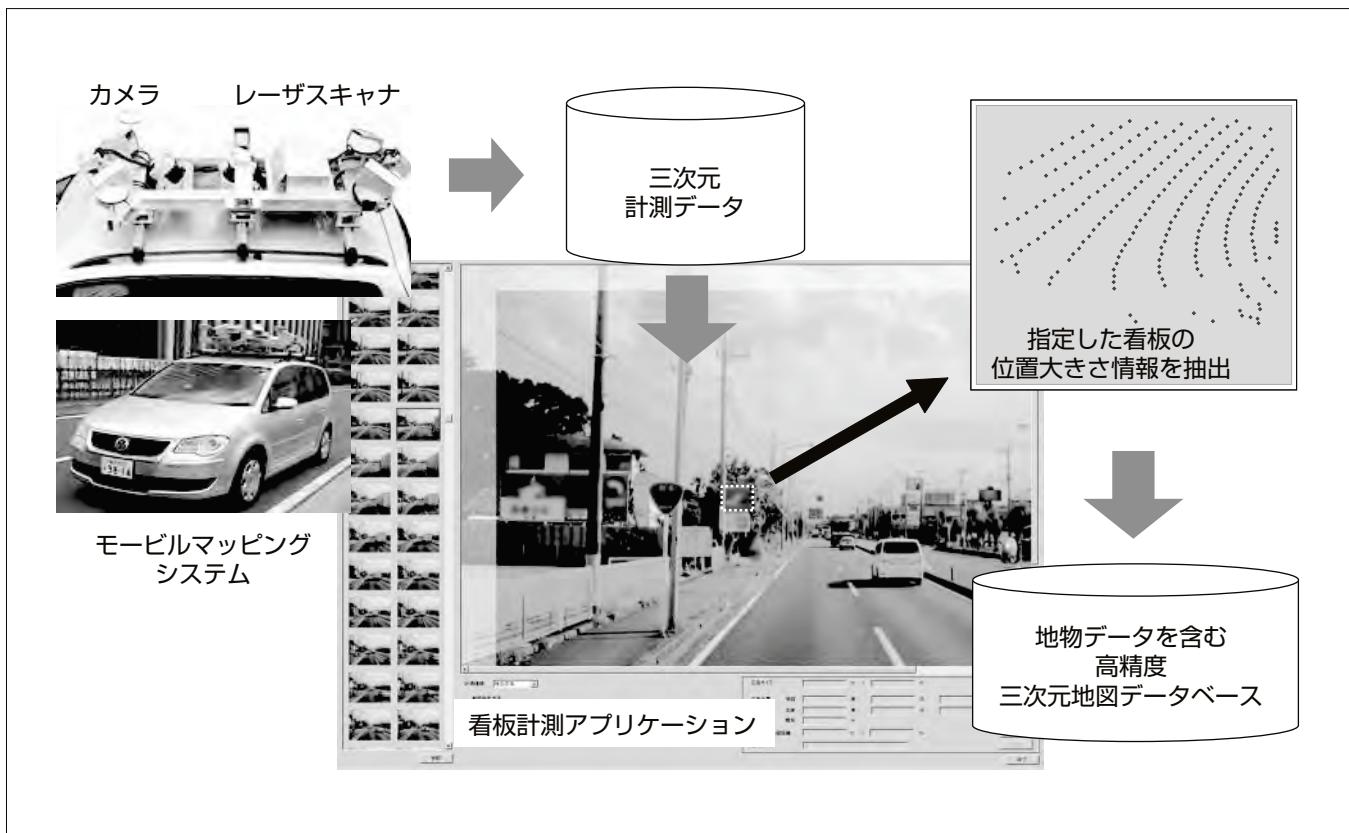
あらゆる生活空間でIT(Information Technology)を使って活動を支援するユビキタスサービスの普及が期待されている。スマートフォンなどのナビゲーションでは、現実の映像に情報を付加提示する拡張現実が使われるなどヒューマンインターフェースの進展も目覚ましい。このようなサービスでより良いインターフェースを実現するには、あらゆる生活空間のデータを管理するため、空間内の物体の構造データなどを含めた精密な三次元地図データベース(以下“三次元地図”という。)整備が必要である。

そこで、三菱電機は、車載カメラと、レーザセンサを用い、道路周辺の情報を効率的に収集する“モービルマッピング技術”を開発している。今回、物体の自動抽出を目的

とし、道路周辺の看板の位置・大きさを測定する実験を行った。この実験では、看板位置・大きさ計測の際、看板の点群の前方にある電柱や電線といった障害物の点群を排除し目的の看板の点群を抽出することが重要となる。

この課題に対し、当社は、映像から看板を含む領域を指定し、その範囲の点群に対してグルーピングを行った後、画像上の分布範囲を用いて看板から得られた点群を抽出する方法を開発した。この方法を用いた看板の位置・大きさを計測するアプリケーションの試作、評価を行った。

真値との比較の結果、誤差5%以下で計測できることを確認した。今後は、計測精度の向上を図り、ユビキタスサービスの基盤となる高精度三次元地図の効率作成を目指す。



モービルマッピング技術を用いた看板計測アプリケーション

カメラ画像と、センサから得られる点群を利用して、看板の位置、大きさの測量が行えるアプリケーションを試作した。簡単に高精度の三次元地図が作成でき、拡張現実を用いるサービスなどに適用できる。

1. まえがき

あらゆる生活空間でITを使って活動を支援するユビキタスサービスの普及が期待されている。例えば、スマートフォンを使ったナビゲーションサービスなどが、その代表例であるが、そこではカメラによる現在地周辺の実写画像にナビゲーション情報が重畳表示される拡張現実技術が使われるなど、新たなヒューマンインターフェースが種々試みられている。

拡張現実を用いるサービスで、実写画像などに情報を重畳するためには、空間内物体の構造を把握し、管理する精密な三次元地図を整備することが必要である。しかし、三次元の都市地図の整備は屋外、屋内ともに手間がかかり、まだ限定された地域で、建物外観などのデータのみが実現されているにすぎない。

そこで、精密な三次元地図実現のために、車載センサを使って道路周辺の情報を効率的に収集する技術である“モービルマッピング技術”を開発している。モービルマッピング技術は、慣性計測装置(Inertial Measurement Unit: IMU)やカメラ・レーザレンジファインダ等を車両に搭載し様々な計測を行うものであり⁽¹⁾⁽²⁾⁽³⁾⁽⁴⁾⁽⁵⁾⁽⁶⁾、主に道路面調査や道路周辺環境の計測等に適用されている。

当社はこのモービルマッピング技術を利用し、道路周辺の物体、すなわち地物の抽出を目的として、データとして重要な看板の位置、大きさを測定する方式を開発し、評価実験を行った。この看板計測では“点群の中から看板の点群だけを抽出すること”が課題となる。特に、看板前の障害物の除去をいかに行うか、どの点群が看板でどの点群が障害物であるかを正確に判定することが重要である。

この課題に対し、ユーザーの指定情報を利用した半自動点群抽出手法を提案する。この手法は、看板を含む画像からユーザーに看板の範囲を指定させ、その範囲情報と、画像上における点群の分布状況から看板の点群を抽出するものである。この手法を実装したアプリケーションを試作して看板の幅・高さ算出を行い、真値と比較した。

2. 点群と看板計測アプリケーション

2.1 点群と地物

点群は、車両に搭載したレーザレンジファインダによって得られる測量結果(車両からの相対距離)を三次元座標に変換した、空間絶対座標上の点の集合である。それぞれの点は道路周辺環境上の1点を示し、その集合が道路・建物等を示す。そのため、適切に点群を抜き出すことで、道路周辺の特定の地物を構成することが可能となる。図1は道路周辺点群全体から行き先表示板の点群を抜き出した例である。

2.2 屋外広告物調査と看板計測アプリケーション

図1のように地物に対応する点群を抜き出し、点群の位置情報を利用することで地物の存在する場所・大きさ等を求めることができる。このことを利用し、“屋外広告物調査”実証実験を行った。この実証実験システム(図2)の一部として、カメラ画像と点群を利用して屋外広告物(看板)の位置・大きさ等の測量を簡易に行える看板計測アプリケーションを試作した。現状は調査ごとに、調査員が現地に赴き、看板の位置・大きさを個別に計測している状況であるが、データ収集車両が走行し、収集したデータを蓄積しておけば、必要な看板の計測がいつでもできるようになる。

なお、このアプリケーションでは、看板の形状・大きさが多岐にわたることを考慮し、画像から看板を自動認識させる全自动計測は想定しておらず、ユーザーに画像上の看板の範囲を指定させ、その看板の位置・大きさを計測する半自動計測を採用している。

2.3 点群の取捨選択

点群から看板位置・大きさの計測を行うためには、まず、看板を示す点群を選び出す必要がある。

カメラ撮影時の位置・向き等の情報によって、点群はカメラ画像上にマッピングすることができる。これによって、画像上の範囲内に収まる点群を選ぶことができ、また、目

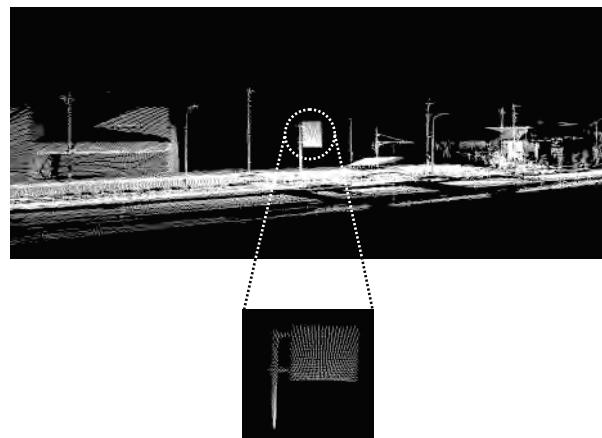


図1. 点群から行き先表示板を抜き出した例

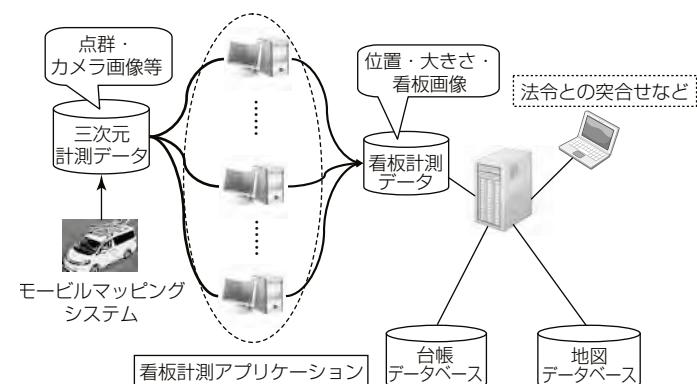


図2. 実証実験システムの構成

的の看板の点群はこの範囲内に収まる。ただし、画像から範囲を指定するだけでは対象物までの距離が判別できないため、看板の点群だけを選び出すことはできない。これは範囲内の点群すべてが目的の看板の点群とは限らず、他の地物が画像範囲内に含まれる可能性があるためである。例えば、目的の看板の点群の前方に電柱や電線といった障害物があれば、それらの点群が含まれ得る。図3は図中の取得対象範囲(看板の一部分)に収まる点群を抽出し、点群内の各点の位置(緯度経度方向)をプロットしたものであるが、看板のほかに手前の電柱・電線の点群も含まれることが分かる。看板測定のためには障害物の点群を取り除く必要がある。

3. ユーザー指定情報を利用した点群抽出手法

このアプリケーションでは、画像上の看板範囲は、ユーザーによって指定されるため、“看板以外の地物で看板より手前にあるものは、指定範囲の一部だけを覆う”ということになる。例えば、ユーザーは、目視によって地図上の範囲指定を行うため、選択された領域には、看板前方の電線や電柱等の障害物があっても看板の範囲を比較的正確に指定することができる。

目的の看板の点群は、指定範囲全体に分布するが、看板以外の地物は、指定範団の一部の領域に分布するという特徴の相違となる。図3の例における看板と電柱等の点群の分布の違いを図4に示す。

一方、看板の後ろの他の看板や建物の壁等の地物の点群は、指定範囲全体に分布し得る。図5に事例を示す。

これらの特徴から、目的の看板の点群を選び出すためには、“各点群の部分集合を画像上にプロットした上で画像上の分布を抽出し、‘範囲全体を覆う分布となる部分集合のうち、画像撮影位置から最も近いものを選択する’”ことで、目的の看板の点群を抽出することができる。点群抽

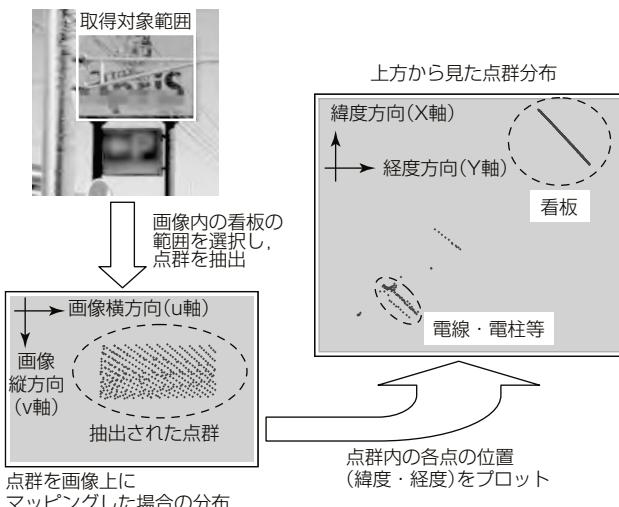


図3. 選択点群と点群の示す地物

出手順としては“入力された点群を、点群間距離でグループ化”し、各グループに対して“点群を画像上にマッピングし指定範囲にどれだけ収まるかを計算”，その占有度合いが閾値(しきいち)を超える、最も画像撮影位置に近いグループを看板の点群として出力する⁽⁷⁾。

4. 試作アプリケーションと評価

4.1 試作アプリケーション

試作した看板計測アプリケーションの画面を図6に示す。このアプリケーションでは、カメラ画像上の看板の範囲をユーザーが矩形(くけい)で指定することで、その看板の点群を抽出して大きさ・位置を計測する。

なお、このアプリケーションでは業務効率化のための機能として、一度計測した看板を明示する機能も実装している。これは、計測した看板の三次元位置を記録しておき、画像撮影時の車両位置、カメラ向きの情報から、画像上の位置を求めるため、どのような映像にも表示することができる(図7)。

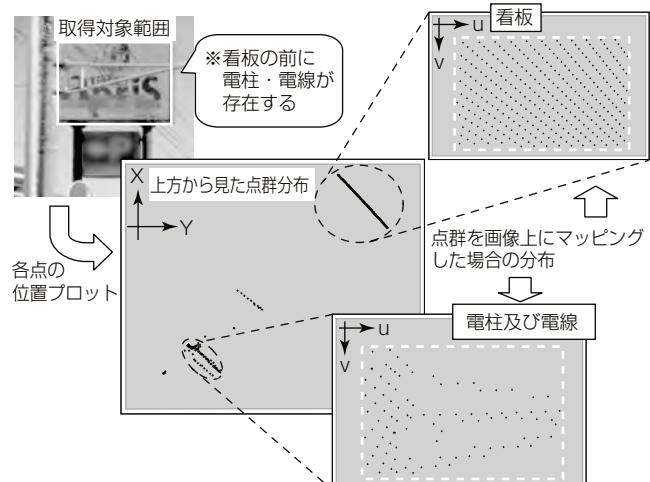


図4. 看板とそれ以外の地物の点群の分布例(1)

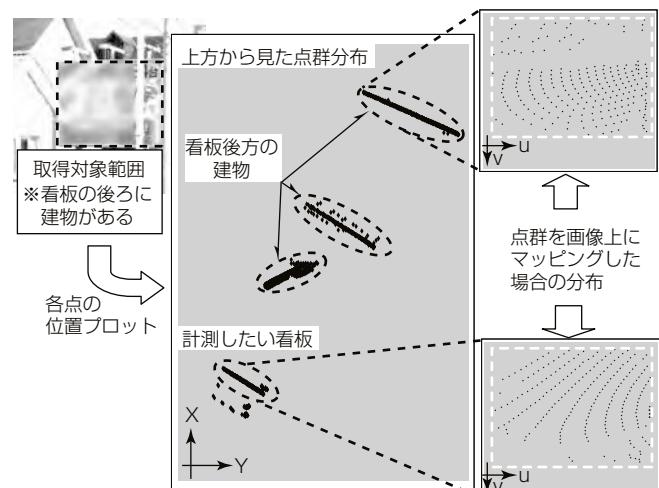


図5. 看板とそれ以外の地物の点群の分布例(2)



図6. 看板計測アプリケーションの画面

4.2 評価実験

千葉県内の国道を走行、データを収集し、看板の計測を行った。大きさと位置情報を抽出し、看板の管理台帳に記載されている値を真値として比較した。比較例を表1に示す。

どの測量も誤差5%以内で測定できていることが分かる。図3～図5で示すように実際に看板前方には数多くの障害物が存在していたが、看板の点群の抽出は、いずれの場合にも正しく行われた。

誤差は、ユビキタスサービスの拡張現実インターフェースに適用するには十分の精度であると考えられるが、絶対精度で考えると大きなばらつきが生じている。これは、画像上のユーザーの指定領域の誤差によるものと考えられる。画像上の1ピクセルが、今回の実験では1～2cmになるため、現状のユーザー指定領域内のみを扱う方法では、精度向上は困難であり、範囲指定方法や扱う点群の範囲決定方法等を検討する予定である。

5. むすび

カメラ画像と点群を利用した、看板の位置・大きさ測定のための点群抽出手法について述べた。ユーザーが視認した看板領域を指定することで、領域と点群の関係を用いて、レーザ計測点から看板以外の障害物、背景物のデータを取り除き、対象看板の点群を精度よく抽出できる。開発しているモービルマッピング技術を活用し、高精度三次元地図データベースを整備することで、実世界と情報世界を統合するユビキタス情報サービスとそのヒューマンインターフェースを実現する予定である。

この研究の一部は、総務省委託研究“ユビキタス・プラットフォーム技術の研究開発”(平成20～22年度)の一環として実施されたものである。

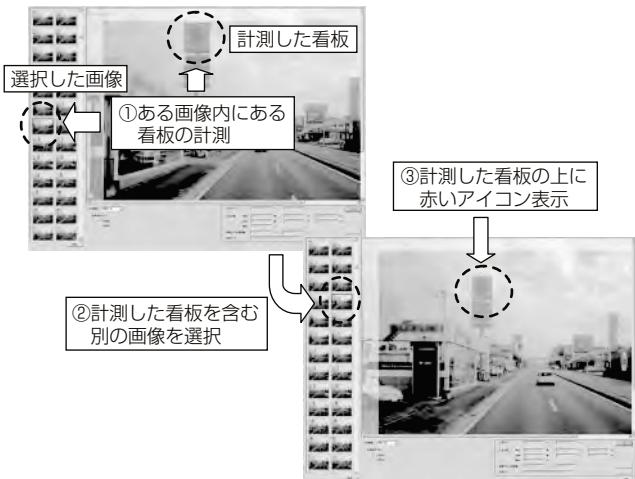


図7. 地物の画像上での配置

表1. 測定結果・誤差

No.	真値(幅、高さ)	測定結果(幅、高さ)	幅誤差率	高さ誤差率
1	3.50m, 3.80m	3.54m, 3.87m	1.10%	1.80%
2	1.80m, 3.60m	1.74m, 3.56m	3.30%	1.10%
3	4.00m, 3.27m	3.91m, 3.28m	2.30%	0.30%
4	5.00m, 4.00m	5.21m, 4.09m	4.20%	2.30%

参考文献

- (1) 三菱電機：三菱モービルマッピングシステム 高精度GPS移動計測装置
<http://www.mitsubishielectric.co.jp/pas/service/mms.pdf>
- (2) 吉田光伸, ほか：モービルマッピングシステム, 三菱電機技報, 81, No. 8, 513～516 (2007)
- (3) 山口博義, ほか：多眼ステレオ3D計測システム, 写真測量とリモートセンシング, 47, No. 5, 18～21 (2008)
- (4) 史中超, ほか：マルチセンサ3次元計測システム, 写真測量とリモートセンシング, 47, No. 5, 22～25 (2008)
- (5) 石川貴一朗, ほか：モービルマッピングシステムと三次元空間情報管理システム(G-VIZ™)を用いた実都市空間モデリング&ウォータースルーシステム, 第13回ロボティクスシンポジア (2008)
- (6) Ishikawa, K., et al.: A study of precise road feature localization using mobile mapping system, 2007 IEEE/ASME International Conference on Advanced Intelligent Mechatronics, 1～6 (2007)
- (7) 石渡要介, ほか：オクルージョンを考慮したレーザー点群抽出に基づく看板計測アプリケーション, 情報処理学会研究報告. ITS, [高度交通システム]2011-ITS-44(1), 1～7 (2011)