



家庭から宇宙まで、エコチェンジ eCO Changes

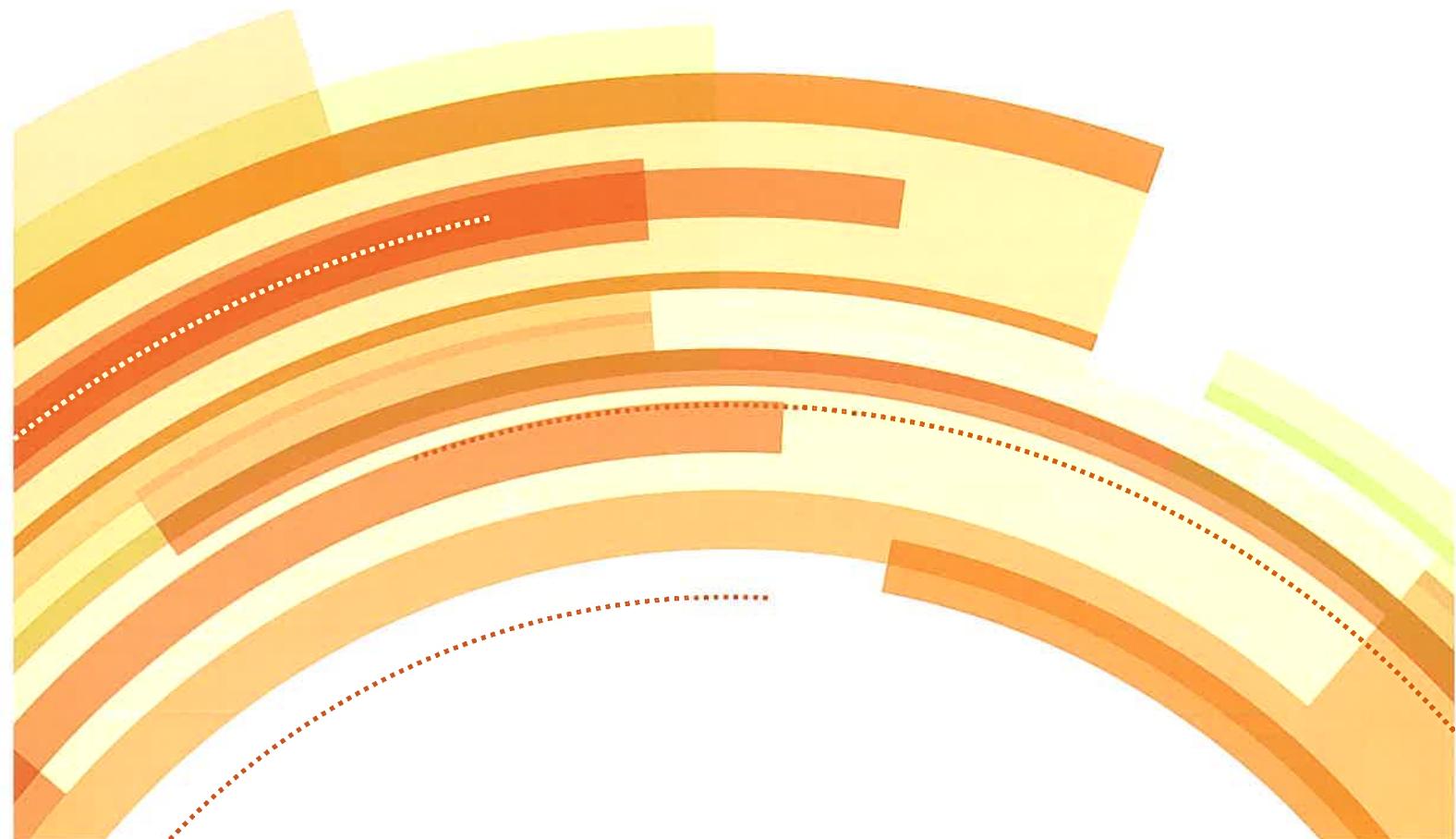
三菱電機技報

7

2014

Vol.88 No.7

「ユーザーエクスペリエンスデザインとその基盤技術」



目 次

特集「ユーザーエクスペリエンスデザインとその基盤技術」	
デザインという言葉の歴史的変化とUXデザイン 1	
山中俊治	
ユーザー エクスペリエンスをデザインする 2	
杉浦博明・小山健一・藤本仁志	
成田国際空港デジタルサイネージ 7	
鶴直樹・米沢みどり	
コンセプトカー“EMIRAI2”的HMI 11	
相川真実・八木澤喬樹・今中英夫	
三菱HEMSのUXデザイン 15	
飯澤大介・北村憲史郎	
音声対話技術 19	
藤井洋一・小路悠介・石井純	
テキスト音声合成技術 23	
大塚貴弘・川島啓吾・古田訓・山浦正	
音のデザイン 27	
木村勝・堀田厚	
映像合成技術 31	
三木洋平・古木一朗・三浦紳	
最新映像符号化技術とスーパーハイビジョンエンコーダ 35	
坂手寛治・井戸貴之・本山信明	
デジタルプリント品質向上への取組み 39	
長瀬百代・古木一朗・三浦紳	
映像同定技術“Video Signature”とその応用技術 43	
工藤大樹・鶴崎真理子・阿倍博信	
タッチパネル感度向上技術 47	
佐々木雄一・根岸博康	

特許と新案

「音声対話装置」「映像音声記録装置」 51	
「合成表示装置」 52	

User Experience Design and Its Fundamental Technologies
From a Historical Viewpoint of "Design" and "User Experience Design"
Shunji Yamanaka

Designing for User Experience
Hiroaki Sugiura, Kenichi Koyama, Hitoshi Fujimoto
Digital Signage for Narita International Airport
Naoki Tsuru, Midori Yonezawa

HMI for Concept-car "EMIRAI2"
Masami Aikawa, Takaki Yagisawa, Hideo Imanaka
UX Design for Mitsubishi HEMS
Daisuke Iizawa, Kenshiro Kitamura

Spoken Dialog Technology
Youichi Fujii, Yusuke Koji, Jun Ishii
Text-to-Speech Technology
Takahiro Otsuka, Keigo Kawashima, Satoru Furuta, Tadashi Yamaura

Designs of Sound Quality
Masaru Kimura, Atsushi Hotta

Video Synthesis Techniques
Yohei Miki, Ichiro Furuki, Shin Miura
New Video Coding Technology and Super High Vision Encoder
Hiroharu Sakate, Takayuki Itsui, Nobuaki Motoyama
Towards Improved Digital Printing Quality
Momoyo Nagase, Ichiro Furuki, Shin Miura
Outline and Application of Video Identification Technology "Video Signature"
Daiki Kudo, Mariko Tsurusaki, Hironobu Abe
High Sensitivity Detection Method for Projected Capacitive Touch Panels
Yuichi Sasaki, Hiroyasu Negishi

表紙：ユーザーエクスペリエンスデザインとその基盤技術

デザインの重要性が指摘されている。三菱電機では、これまで、ユーザビリティ視点、誰でもが使えるユニバーサルデザイン視点で、人間中心のデザインを行ってきた。ユーザーエクスペリエンスデザインはそれを更に発展させた製品・サービスと人との相互作用の全ての側面に着目したデザインである。

成田国際空港デジタルサイネージ(1)では単なる広告提供にとどまらない空間演出を、三菱HEMS(Home Energy Management System)(2)では、機器を初めて手にしたときからのユーザーの体験を、コンセプトカーEMIRAI2(3)では、未来の車社会のコンセプトに基づく安全な操作をデザインしている。

三菱電機では、高度な基盤技術に立脚し、製品・サービスの価値を、ひいては豊かな社会をデザインしていく。



巻/頭/言

デザインという言葉の歴史的变化とUXデザイン

From a Historical Viewpoint of "Design" and
"User Experience Design"

山中俊治
Shunji Yamanaka



誰もが“デザインが大切だ”と口にする時代である。メーカーの経営者はそろって、これからの商品価値はデザインにこそあるといい、雑誌もWebもデザインについて多くの文字数を使って語る。数年前から科研費の新たな分科細目として“デザイン学”が設定されるようになった。明らかに多用され重用されるようになった“デザイン”という言葉だが、その意味は文脈によってかなり異なる。

一般的に“デザイン”は美術の一分野のように語られる。もともと英語圏のDesignという言葉の原義は広く“計画に形を与えること”であり、工学的な設計の意味も含まれるが、20世紀の初頭にインダストリアル・デザイナーという職業が成立したころには、狭義に“意匠”的意味でも用いられるようになった。我が国でも、デザインというカタカナ表記が戦後に導入されて以来、色や形、又は図案、柄(ガラ)等の意味で一般化しており、今日でも“デザイン家電”などと表される時のデザインは、ほぼ“おしゃれな(外観の)”と同義である。

歴史的に見ても、20世紀の産業デザインは、大量生産される商品に常に新しい形を与え続け、消費者の欲望を喚起する装置として機能してきた。アドリアン・フォーティが指摘するように、新しく発明された技術を既存の日常的な家具に見せかけたり、未来を感じさせるようなスタイルで煽(あお)ったり、又は技術に進展はなくとも、見かけだけを変える機能をデザインが果たしてきた(フォーティはこうしたデザインの働きを軽視すべきではなく、それこそが資本主義社会を支えてきた原動力だと言い切る)。その意味では少なくとも、前の世紀のデザインは科学技術開発の下流にあったと言える。

しかし今日、前の世紀の後半に露(あら)わになった様々な要因がデザインを産業構造の上流に押し上げた。ソフトウェア技術とネットの隆盛による製品開発とサービスの一体化。先進国における消費者の要求の高感度化と細分化によって感覚的な商品価値が主役になったこと。地球環境問題が露わになり、成長の限界の中でつくる意味そのものを問われるようになったこと。情報社会の中で個別の商品性能よりもブランドの価値が優先されるようになったこと等々。こうした状況を踏まえてデザインは、何をどのように

に使われるかをあらかじめ計画するプロセスとして、技術スペックや製造方法に先立つようになってきた。同時にデザインは、応用美術としての意匠設計ばかりではなく、人と人工物の関わり全般を計画する技術に変貌しつつある。

特に、90年代に露わになったユーザビリティの問題は、デザインを産業構造の中核に据える原動力になった。“機能満載だがユーザーはそのほとんどを使っていない”商品が蔓延(まんえん)するようになったというD.A.ノーマンらの問題意識は多くの技術者とデザイナーが共有するようになり、その解決手法としてヤコブ・ニールセンらがユーザビリティ・エンジニアリングを発展させる。“ユーザビリティ”的概念を導入したことの最も大きな功績は、スタイルリングと人間工学が明瞭な目的意識の下に合流し、“デザイン”を工学的に理解しやすい概念として拡大させたことにある。この方向で更に拡大されたデザインを象徴する言葉の一つが、今回の特集であるユーザーエクスペリエンスデザイン(UXデザイン)である。

やや逆説的だが、ニールセンらが提唱したユーザビリティ・エンジニアリングが優れて実用的であったゆえんは、ユーザーの行動を科学的に厳密には解明しようしないところにあった。ユーザーテストを繰り返し、観察と改良を繰り返すことで網羅性や再現性に捕らわれることなく実用的な設計解を得るその目覚ましい成果は、科学的な厳密さの袋小路を避ける民俗学的マクロ視点によって成立している。今後の発展方向としてはやはり、反応装置としての“人”モデルをミクロにも解明することになるだろう。時には生物学的な侵襲を伴う高精細なセンサによる多次元パラメータの同時解析手法や、ログ的なビッグデータ解析等、テストに立ち会う開発者達の“観察”に頼っていた人と人工物の関わりに、実験科学の手法が導入されつつある。一方で音声インターラクションの技術や高精細なAR(Augmented Reality)環境は、人を取り巻く環境自体をモデル化することを可能にしつつある。“エクスペリエンス”は、極めて幅の広い境界が曖昧な言葉だが、それでもその詳細な記述と解析の方向性は見えてきたように思う。科学と美術を統合し、人を取り巻く人工環境全体を扱う“デザイン”は始まつばかりである。

巻頭論文

ユーザーエクスペリエンスを
デザインする

杉浦博明*



小山健一**



藤本仁志***

Designing for User Experience

Hiroaki Sugiura, Kenichi Koyama, Hitoshi Fujimoto

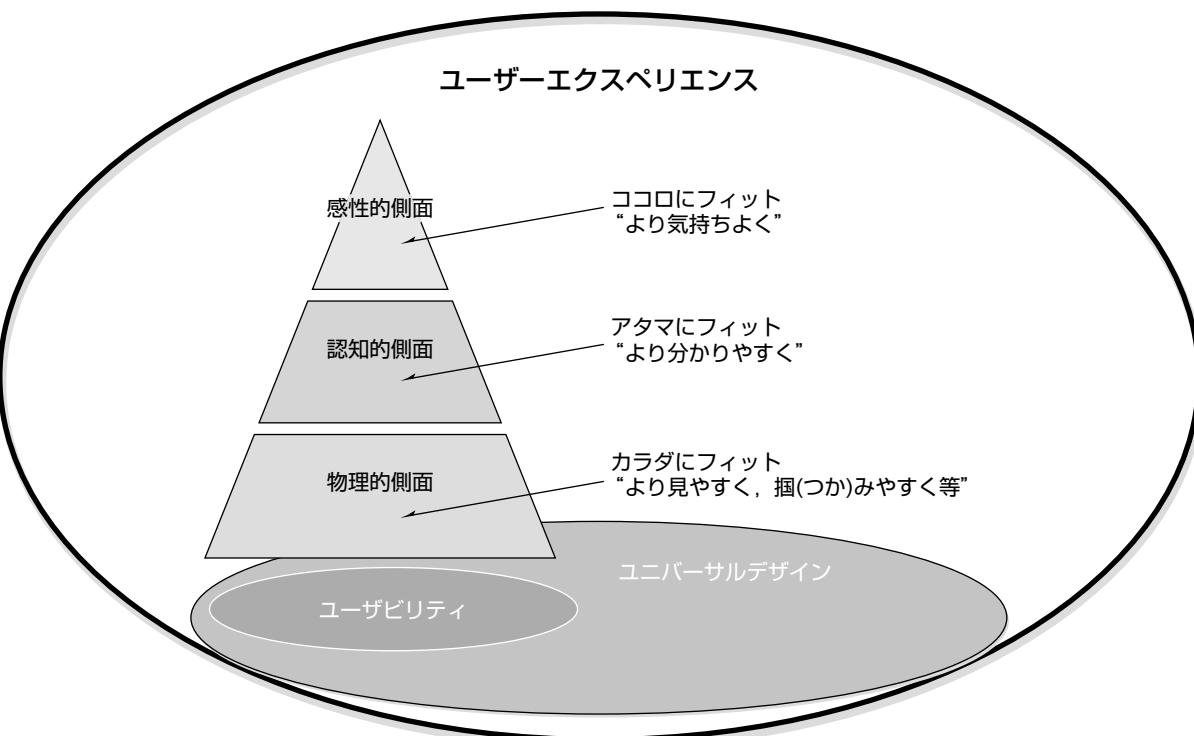
要 旨

三菱電機では、インターフェースデザインを物理的側面“より操作しやすく”，認知的側面“より分かりやすく”，感性的側面“より気持ちよく”的3階層に分け、より上位層の実現に向けてデザインの取組みを進めている。さらに、ユーザビリティの点でより多くの人を対象とするためのユニバーサルデザイン(UD)の普及に努めてきた。

1990年代から提唱され始めたユーザーエクスペリエンス(UX)という言葉は、多くの研究者によって様々な解釈と定義が試みられてきた。当初は、Webデザインなど、主にIT分野でユーザビリティの向上と同義的な使われ方で広まっていた。しかし、IT、通信技術の進歩に伴い、製品の価値が“モノ”すなわち機能から、“コト”すなわち製

品を通じて得られる体験に移るようになり、2010年代を迎えて、UXは人と製品・サービスなどの相互作用の全ての側面として、その概念、評価軸が明確化されるに至っている。デザインの視点からは、実際に利用した体験だけではなく、利用する前の“期待”から、利用後の“振り返り”までの時間的な変化までを含んだ総合的な概念として捉えることで、製品・サービスへの要求を整理し、体験価値を高めるデザイン手法も検討されている。

評価手法など今後の課題も多いものの、当社はUX視点でデザイン活動を捉え直し、さらには、デザイン部門と要素技術開発部門とが密に連携することで、製品価値を高め、豊かな社会の実現に貢献する。



インターフェースデザインの進化

インターフェースのユーザビリティは物理的な使いやすさから、認知的な分かりやすさに進化してきた。そして、ユーザビリティをより多くの人に提供する活動がユニバーサルデザインである。さらに感性的な要素を考慮したインターフェースを目指すが、ユーザーエクスペリエンスは操作の瞬間だけでなく時間的空間的に対象を広げた概念である。

1. まえがき

当社では“デザインの行き先は、人”を理念として掲げている。“楽しい”“快適”“気持ちはいい”など、感性的側面での品質向上を目指してきた。使いやすさや分かりやすさで、より多くの人を対象とするためのユニバーサルデザインに特に配慮した活動を90年代から継続的に行っており、企画や設計の各段階で広く普及している⁽¹⁾⁽²⁾⁽³⁾。

特に、各種の機器・システムの高度化多機能化によって、ユーザビリティ視点での“デザインの重要性”“製品価値としてのデザイン”が強く問われるようになり、ユーザーエクスペリエンス(UX)という言葉が頻繁に使われるようになった。

本稿では、UXの定義と概念、それに向けたデザインプロセス、そして当社の取り組みについて述べる。

2. ユーザーエクスペリエンスとは

2.1 モノからコトへの変化

モノの消費の時代は、ユーザーは生活における不足や不満に対するニーズを充足するために商品やサービスを選択して購買した。このとき、機能や価格といった価値は企業によって与えられてきた。しかし、コトの消費の時代では、生活における自己創造を実現するため、商品やサービスは手段として選択され、これらから得られる体験に基づく価値は、ユーザーが参加することで生み出される。

ソニーはウォークマン^(注1)で音楽を屋外に持ち出す体験を、アップルはiPod^(注2)とiTunes^(注2)でシームレスに音楽を聞く体験そのものをデザインした。体験のデザインは旅館やホテルでのサービス“おもてなし”にも共通するものであり、あらゆる製品やサービスでも一流のデザインが求められ、選択される時代になった。

(注1) ウォークマンは、ソニー(株)の登録商標である。

(注2) iPod, iTunesは、Apple Inc. の登録商標である。

2.2 ユーザーエクスペリエンスの定義

ユーザーエクスペリエンス(UX)という言葉は認知科学者であったDonald A. Norman博士の造語と言われている。Norman博士が設立したコンサルティング会社ニールセンノーマングループではUXを“UXは企業、サービス、製品とエンドユーザーとの相互作用の全ての側面を網羅するもの。”と定義⁽⁴⁾している。そして、2つの要件“混乱や面倒なしで的確にユーザーニーズを満たすこと”と、“所有したり使用する喜びを生み出す簡潔さと優雅さ”をあげている。

また、クオリティの高いUXを実現するためには、“多角的な専門分野のサービス”的シームレスな結合が必要とも述べている。それらの分野とは、エンジニアリング、マーケティング、グラフィックデザイン、インダストリアルデザイン、インターフェースデザインである。ここから次の

2つの留意項目が導出される。

- (1) ユーザインターフェース(UI)設計は重要な部分であるがUXとは区別する必要がある。
- (2) UIの品質属性であるユーザビリティともUXは区別しなければならない。

2.3 UXハニカム

2000年代に入ると、効果が計りやすいこともあり、主にWebデザイン設計でUXが取り入れられるようになった。

2004年にPeter MorvilleがUXの評価軸として、次の7要素をUXハニカム⁽⁵⁾として示した(図1)。

- (1) Useful：ユーザーにとって役に立つ、便利であること
- (2) Usable：使いやすいこと、分かりやすいこと
- (3) Desirable：魅力があること、好ましいこと(感情的な要素)
- (4) Findable：ユーザーが必要な物を見つけやすいこと
- (5) Accessible：誰でもがアクセスしやすいこと(建物でエレベーターやスロープを設置するように、誰でもが分かるよう、可読性の考慮、音声ナビゲーション等の配慮をする必要)
- (6) Credible：信頼できること、安心感があること
- (7) Valuable：ユーザーに価値があると思ってもらえること

(1)から(6)の6要素が、UXの目標として中心にある(7) Valuableを取り囲むかたちになっている。

図1の下部の3項目“Findable”“Accessible”“Credible”はWebなどのオンライン環境に関わりが深い要素であり、それぞれが“ファインダビリティ”“アクセシビリティ”“セキュリティ”に対応付けられる。従来の“ユーザビリティ”“ユニバーサルデザイン”は“Useful”“Usable”“Findable”“Accessible”的品質向上を目指すものである。これらに、より感性的で計りにくい評価尺度ではあるが“Desirable”と、物理側面のセキュリティだけではなく製



図1. UXのハニカム⁽⁵⁾

品に対する信頼という“Credible”的感性的側面が加わって全体を構成している。

2.4 UX白書

2010年にUX関連の研究者30名以上が集まり、基本コンセプトについて議論し、整理したものをUX白書⁽⁶⁾として公開している。ここで、UXは、実際に利用した体験だけではなく、利用する前の“期待”から、利用後の“振り返り”までを、その人の中の評価の変化を経て、間接的なUXが拡張されることが指摘されている。図2はUXを期間で分け、それぞれのUXがもたらすフィードバックを図示したものである。

利用前の予期的UXは、ユーザーはこれから得られるであろう体験を想像し期待する基になるUXであり、この期待が高いほど次の段階のUXにつながる。利用中の一時的UXは、使い始めの分かりやすさや、一つ一つのインタラクションに感じるUXである。利用後のエピソード的UXは、楽しかった、不安なくできた等の特定の利用に関する評価を顧みたり、他人に話したりする段階のUXである。利用時間全体にわたる累積的UXは全体として様々な利用を振り返るときのUXで、ブランド力の醸成など次の予期的UXに大きな影響を与える。

この分類は、UX定義にある“ユーザーと企業、サービス、製品との相互作用の全ての側面を網羅”に基づき、UXをユーザー側から時間の流れに沿って分類したものである。ユーザーが利用している以外の期間での体験を含めた全体がUXであり、どの期間に着目するかによって、考慮すべき要件が異なることが明らかになる。

UXをデザインするに当たり、各期間で考慮すべき要件例を列挙する。予期的UXには“見ただけでほしいと思う形”が必要である。一時的UXに対しては、“迷わず直感的に使える”“思い通りに使える”“やりたいことができる”“気持ちよく使える”を実現する仕組みの提供が要求される。エピソード的UXを向上させるためには、“もっと便利な使い方ができる”“自分なりの使い方ができる”“状況や能力

に応じた方法がある”“他の製品との組合せに対応できる”“困ったとき、エラー時に対応できる”といった仕組みで、ユーザーの変化に追従することが重要である。累積的UXには“飽きずに使える”“使い込むと味わいが出る”“壊れたときに部品の交換などがスムーズにできる”“これはいいものだと思える”といった意味のある価値が形成されるような仕組みが必要である。

一方、UX白書にも参画した黒須は、図3に示すような時間軸におけるUXの5段階を提案している⁽⁷⁾。右側に企業における対応する活動を併記した。

先に述べた予期的UXをユーザー行動として店舗に出かける前後(電子商取引ならEC(Electronic Commerce)サイトの訪問前後)を2段階に分けている。それぞれに、製品やサービスに触る前にある欲求や期待感、手に入れたときの喜びといった主観を対応付けている。

2.5 標準化状況

2.5.1 ISO 9241-210

人間中心設計(HCD)プロセスを規定したISO 13407の改訂版ISO 9241-210では、2010年にHCDの目的にトータルなUXの実現を組み込んだ。ここではUXを“製品やシステム、サービスの利用、及び／又は予想された使い方によつてもたらされる人々の知覚と反応”と定義している。

- (1) UXとは、使用前、使用中、使用後に生じるユーザーの感情、信念、嗜好(しこう)、知覚、生理学的・心理学的な反応、態度、達成感等の全てを含む。
- (2) UXは、ブランドイメージ、見た目、機能、システムのパフォーマンス、インタラクティブシステムのインタラクティブな振る舞いと支援機能、事前の経験から生じたユーザーの内的及び身体的状態、態度、スキルとパーソナリティ、利用状況の結果である。
- (3) ユーザーの個人的目標という観点から考えたときには、通常はUXに付随する知覚的・感情的な側面をユーザビリティは含む。ユーザビリティの評価基準を用いて、UXの諸側面を評価することができる。

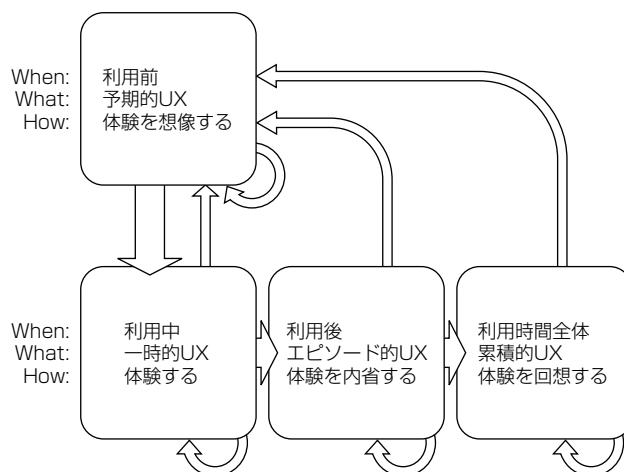


図2. 期間で分類したUX⁽⁶⁾

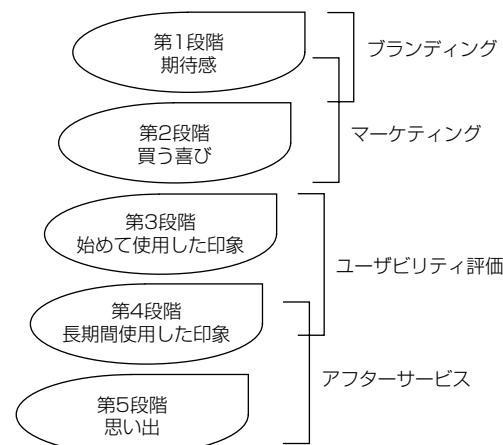


図3. 時間軸におけるUXの5段階⁽⁷⁾

2.5.2 ISO/IEC 25010 : 2011

ISO/IEC 25010 : 2011(JIS X 25010 : 2013)では利用品質という概念を次のように定義している。システムの利用時の品質は、製品(システム又はソフトウェア製品)が利害関係者に及ぼす影響を特性づける。それは、ソフトウェア、ハードウェア及び運用操作環境の品質、並びに利用者、作業及び社会環境の特性によって決定される。これら全ての要因は、システムの利用時の品質の一因となる。

そして、有効性、効率性、満足性、リスク回避性及び利用状況網羅性の5つの特性でその特長を規定し、満足性は更に詳しく次のように分類されている。

- (1) 実用性(usefulness)：利用の結果及び利用の影響を含め、利用者が把握した目標の達成状況によって得られる利用者の満足の度合い。
- (2) 信用性(trust)：利用者又は他の利害関係者が持つ、製品又はシステムが意図したとおりに動作するという確信の度合い。
- (3) 快感性(pleasure)：個人的なニーズを満たすことから利用者が感じる喜びの度合い。個人的なニーズには、新しい知識及びスキル(技術)を獲得するというニーズ、個人のアイデンティティを伝えるというニーズ及び心地よい記憶を引き起こすニーズを含むことができる。
- (4) 快適性(comfort)：利用者が(システム又はソフトウェアを利用するときの)快適さに満足する度合い。

3. ユーザーエクスペリエンスに向けたデザイン

3.1 何をデザインするか

UXは“楽しい”“爽快”“安心”“いろいろ”などユーザーが抱く感情から発して、記憶に残るものである。感情は一瞬であるが、UXとして残るものは繰り返しも含めて、時間的な概念が大切な要素と考えられている。対象と状況によってUXは全く異なる。そもそもUXそのものはユーザーの個人的なものであり直接的に操作できないことが前提となる。

したがって、企業側ではユーザーにどんなUXを体験してもらうことを期待するのかを計画し、その仕組みをデザインする必要がある。それには、従来の製品やサービスのデザインにとどまらず、ブランドイメージ構築から継続的なサービス提供、アフターサービス等を含む広い範囲のUX向上に向けた取組みが求められる。

もちろん、製品・サービスの開発でもUXを意識したデ

ザインを行う。徹底的にユーザーの体験を中心に置き、ユーザーの“期待”“意欲”“理解”といった感性面にも目を向けて、時間軸を意識した捉え方も加味する。

3.2 デザインプロセス

図4は製品開発におけるUXのデザインフレームワークを示している。開発の各段階でUXに対する開発フェーズが対応し、最下段が具体的な手法例である。開発では、次に示す3つのポリシーを守った上で、それぞれのフェーズで各手法を適宜用いる。

(1) UXコンセプトの明確化

個々の製品・サービスのコンセプトより上位に位置付ける、目指すべき体験価値のコンセプトを明確化する。ターゲットユーザーの生活や現場での利用状況を理解し、目指すべき体験価値(提供価値)に基づくコンセプトを定義する。従来のプロダクト中心のコンセプトに比べて、ロバストなコンセプトになりやすい。

(2) UX発想からのブレイクダウン

ユーザー中心の視点でシナリオに分解し、製品や事業提案(デザイン)に落とし込む。機能から発想するのではなく、ユーザーの体験から機能や振る舞いにブレイクダウン(具現化)していく。

①アクティングアクト

UXを可視化する手段として、操作をペーパープロトタイプで表現し、シナリオに沿って演じる手法。

②シナリオライティング

サービス体験を具体的なシナリオにして、コンセプトを徐々に詳細化して記述し製品企画を立案する方法。階層ごとにビジョンを具体化していく。

(3) 当たり前品質の確保

ユーザビリティ品質などの“当たり前品質”は確保した上で“魅力的品質”を目指す。様々なユーザー特性を常に把握しながら、望ましい体験価値を引き出すデザインを創出する。

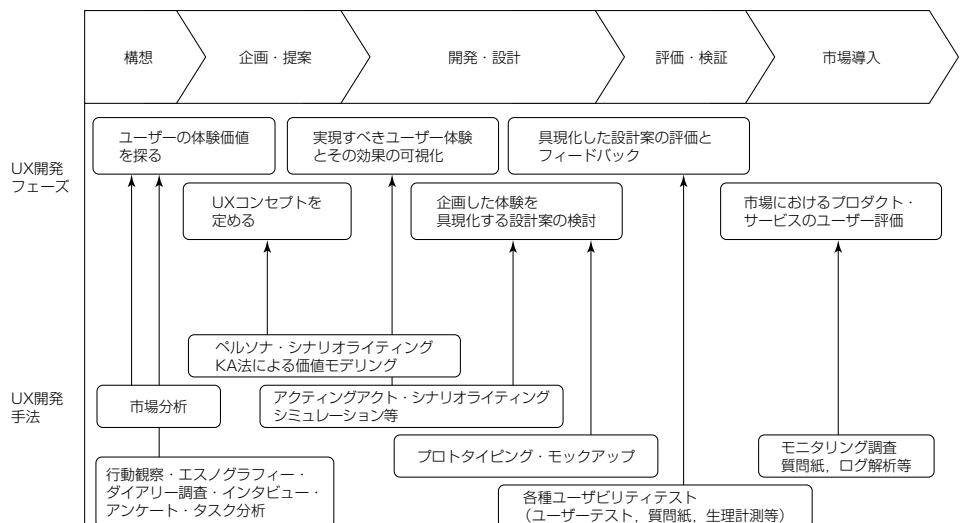


図4. UXのデザインフレームワーク

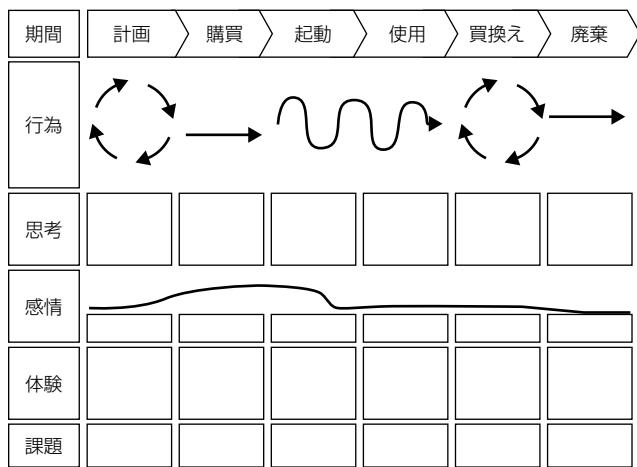


図5. エクスペリエンスジャーニーマップ

よりユーザーの体験、感情を体系的に捉える手法として、図5に例を示すエクスペリエンスジャーニーマップ(カスタマージャーニーマップとも呼ばれる)によって整理することも効果的である。横軸に時間(期間)をとり、ユーザーの行為、考えていること(思考)、感じていること(感情)、課題、UX(体験)を図やグラフ、文字で記載していく。具体的なユーザー像を特定して検討するため、ユーザーの種類が複数想定される場合はその数だけ必要となる。

3.3 今後の課題

UXの概念は明確化されるに至ったが、その評価手法についてはまだ十分な検討がなされていない状況である。ユーザビリティに近い範囲では、まだ製品側からの使いやすさ品質にとどまっている。ユーザー個々人の側に立脚した、使用品質を評価する手法の確立は大きな課題である。これまでユーザビリティ及びUDの向上で用いてきた評価とフィードバックの反復がUXの評価でもより重要となってくると考える。ラピッドプロトタイピングと定量評価の技術を更に高度化することで、UX開発手法を確立していく。

また、UXを生み出すためには製品やサービスを創ることだけでなく、広告のあり方を見直すことも重要で効果が高い。マーケティングでも以前からイベントにおける体験の概念などが検討されており、ユーザーの主観の重要性は認識されていた。体験価値と手段を同時に伝えることで、視聴者はテレビ広告を見ただけで使い方を学び、製品を手にしたときにいかにも直感的に使っているような気分にさせることもできる。これは、予期的UXのマネジメントに当たる。広告に限らずユーザーと企業の接点“コンタクトポイント”でユーザーにどのような体験も総合的に提供できるUXのサプライチェーンを構築する必要がある。

4. む す び

当社では、これまでにもユーザビリティ向上に向けてのデザインに注力してきたが、これらの活動はUXの一部であり、感性面につながるUXをいかに生み出せるかを考え、

デザイン活動に取り組んでいく。

この特集号では、そのデザイン事例として“空港サイネージ”“将来コンセプトカー”“HEMSコントローラ”的3論文で述べるとともに、UXデザインに基づく人工物、人工環境を実現するための基盤技術を述べている。物・情報を理解し、それに働きかけるための技術分野である音声、音響、映像処理、入力デバイスは、従来の単純な蓄積・再生・利用技術から、UXデザインを実現する高度なメディア処理へと着実に進歩している。人と人工物のインタラクション、UXデザイン要求に着目した技術を、“音声対話”“タッチパネル”を例に述べ、内容に着目した構造化技術として“映像同定”技術について述べる。さらに、認知的視点でのメディア合成技術として“音声合成”“映像合成”を、感性に響く高品質化技術として、“音のデザイン”“プリンター”“最新映像符号化技術”を、それぞれの論文で述べる。

製品やシステムのデザインを考える上では、経済は感情で動くという行動経済学にあてはめても、感情への作用を考慮することが今後ますます重要になると考える。卷頭言では“デザイン”的視点・要求が明確に整理されている。そこで指摘されているとおり“科学と美術を統合し、人を取り巻く人工環境全体を扱うデザイン”は今始まったばかりである。

当社は、UXの観点で今までのデザイン活動を捉え直し、ブランド・製品価値を感じてもらい、感情に訴える製品を生み出せるよう、デザイン部門と技術開発部門とが一体となって、人と人を取り巻く人間環境、さらには、製品・サービスを介して結び付く人ととの社会全体をデザインし、豊かな社会の実現を目指す。

参 考 文 献

- (1) 澤田久美子, ほか: ユニバーサルデザイン開発・評価システム, 三菱電機技報, 83, No.12, 738~741 (2009)
- (2) 澤田久美子, ほか: 人にやさしい家電製品を目指したUDの取組み, 三菱電機技報, 86, No.10, 544~547 (2012)
- (3) 藤本仁志, ほか: 家電製品におけるUDの進化, 三菱電機技報, 87, No.9, 488~492 (2013)
- (4) Nielsen, J., et al.: The Definition of User Experience <http://www.nngroup.com/articles/definition-user-experience/>
- (5) Morville, P.: User Experience Design (2004) <http://semanticstudios.com/publications/semantics/000029.php>
- (6) Roto, V., et al.: USER EXPERIENCE WHITE PAPER (2011) <http://www.allaboutux.org/files/UX-WhitePaper.pdf>
- (7) 黒須正明: ユーザーエクスペリエンスのチカラ <http://web-tan.forum.impressrd.jp/e/2012/08/29/13463>

成田国際空港デジタルサイネージ

鶴 直樹*
米沢みどり**

Digital Signage for Narita International Airport

Naoki Tsuru, Midori Yonezawa

要 旨

2012年6月から本格稼働を始めた成田国際空港デジタルサイネージシステムのデザインコンセプト、及びその検討プロセスについて述べる。三菱電機では、2007年からエンジニア、営業、デザイナーというメンバーで構成する次世代空港プロジェクトを発足させ、るべき空港の姿を検討してきた。国内外の空港を調査し、また現地調査や空港関係者へのインタビューによって、空港を取り巻く環境分析を行った上で、るべき空港の姿を描き、その実現のためのステップの明確化を図った。

成田国際空港は出国する日本人、世界各国から来訪する外国人、老若男女様々な人が利用する広大な公共空間である。デザインに当たり、この日本を代表する表玄関が、誰

にでも分かりやすく、安心、便利で楽しい空港するために、デジタルサイネージが担う役割を議論した。

その上で“NARITA FRONTIER VISION”というコンセプトで“世界に先駆ける圧倒的な大画面映像～世界のトップレベル、アジアのリーディング空港を目指して～”と日本の“おもてなしの心”的演出を提案し、これらのコンセプトに基づき詳細デザイン開発を進めた。

これらのプロセスを通じてデザインしたシステムは、単なる情報提供システム、広告提供にとどまらず、空間演出を含めたデジタルサイネージシステムとして高く評価されている。



成田国際空港デジタルサイネージシステム

成田国際空港の第1ターミナル、第2ターミナルの“出発フロア”“到着フロア”及び“レストラン・店舗フロア”に設置された国内最大規模のデジタルサイネージである(端末台数100台、画面枚数334面)。

1. まえがき

三菱電機では、デジタルサイネージシステムを数々の場所に納入してきた。そのデザインでは、様々な実験を通じノウハウの蓄積を行い、ユーザー エクスペリエンス 値値を重視した提案・デザインを行っている。2012年6月から本格稼働を始めた成田国際空港デジタルサイネージシステムは、第1ターミナル、第2ターミナルの“出発フロア”“到着フロア”に設置された大小様々なサイズの広報型サイネージと“レストラン・店舗フロア”に設置されたタッチパネルディスプレイを用いた検索型サイネージで構成された国内最大規模(端末台数100台、画面枚数334面)のデジタルサイネージである。

デザインに当たっては、機能、性能だけに着目するのではなく、旅客を満足させる、感動させる体験を提供することにも着目して、コンセプト構築、端末の配置、プロダクトデザイン、インターフェースデザイン等トータルに検討した。これによって単なる情報提供者の視点だけによる情報提供ではない、旅客の求めている情報が提供できるデジタルサイネージシステムを実現した。

本稿では、成田国際空港デジタルサイネージを事例として、ソリューションデザインの検討プロセス(図1)について述べる。

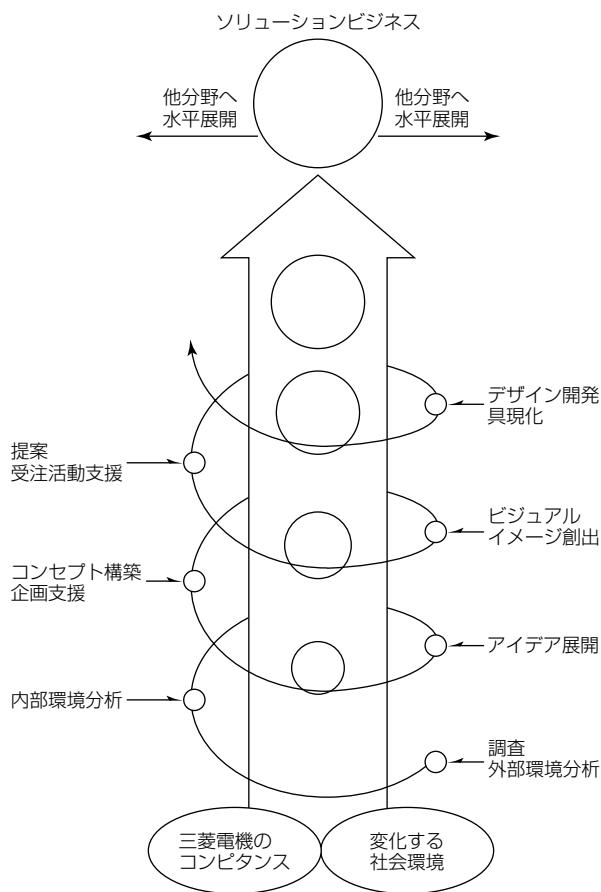


図1. ソリューションデザインの検討プロセス

2. ソリューションデザインの検討プロセス

2.1 調査と外部・内部環境の分析

すべてのユーザーを満足させる、感動させる体験を提供するためには、まず現状の把握が必要である。デジタルサイネージにおけるユーザーには、情報を受け取る旅客と情報を提供する事業者がある。それぞれがどのように考え方行動しているか、何を欲しているかを常に考えながら、現地視察を行うとともに、直接ヒアリングし、将来のサービスを提供する姿を検討した。現地調査として設置エリアの広さなど建物としての特徴や旅客動線等を中心に調査した。また、外部環境として国内外の空港視察の他、旅客満足度調査資料、空港経営の利益状況等から現在の成田国際空港のポジションを確認した。これに内部環境としての成田国際空港の経営ビジョンを重ね、コンセプト構築のための材料とした。

2.2 アイデア展開

(1) 必要情報のリストアップと分類

旅客に必要最低限な情報から、知っておくと便利な情報、事業者が旅客に提供したい情報等をリストアップし、旅客に知らせるタイミングごとに整理した。

(2) ロケーションに対応した案内の検討

空間における様々な場所、それぞれの機能、人の動き、時間軸等、ロケーションに合わせた案内方法を検討した。

(3) デバイスの検討

ロケーションで検討した案内方法に最適なデバイスを検討・選択した。

2.3 コンセプト構築

分析結果、アイデアからキーワードを抽出し、コンセプトを作成した。成田国際空港デジタルサイネージでは、“NARITA FRONTIER VISION～世界のトップレベル、アジアのリーディング空港を目指して～”というコンセプトを創出し、“世界に先駆ける圧倒的な大画面映像”と日本的な“おもてなしの心”的演出を提案した。このコンセプトによる成田国際空港のポジションの変化を示し、コンセプトの有用性をアピールした(図2)。

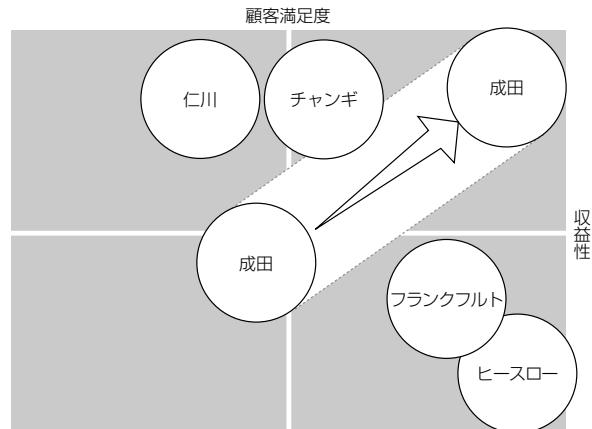


図2. 成田国際空港のポジショニング

“NARITA FRONTIER VISION”は“シームレス・ストレスフリー”“プラスバリュー”“フロンティア”的3つのキーワードで構成されている(図3)。

2.3.1 シームレス・ストレスフリー

旅客が空港内の何處(どこ)にいても、必要な情報を必要なタイミングで入手できるよう考慮した。図4のようにビジネスパーソン、観光客等旅客ごとに異なるニーズに対応した最適な情報提供のイメージを図案化することで、コンセプトを分かりやすく表現した。また旅客動線上では、単体ディスプレイを連続して配置することで移動中の旅客の視界に入りやすくし、滞留スペースでは1度に多くの人に情報を伝達できるよう大画面マルチディスプレイを配置した。人の動き、空間の特徴、目的を考慮しながら設置位置やディスプレイサイズとして適切なデバイスを選択することで、シームレスな情報提供を実現している。このほか、災害時には緊急情報を各端末に一斉同報し、空港全体に瞬時に



図3. コンセプト

に伝達できる仕組みを取り入れている。

コンセプトに基づき、プロダクトデザイン、インターフェースデザインを創出した。デザインでは、ユニバーサルデザインの考えに加え、ユーザーの目的、動き、位置を考慮し、いつでも誰でも使いやすいよう配慮した。各フロアのすべての出入口と通路上に設置される単体ディスプレイの筐体(きょうたい)は現在位置を表示するサイン部分とディスプレイ部で構成しており、設置される空間と旅客の動線、移動状況に応じて2種類の高さの筐体を用意した(図5)。自

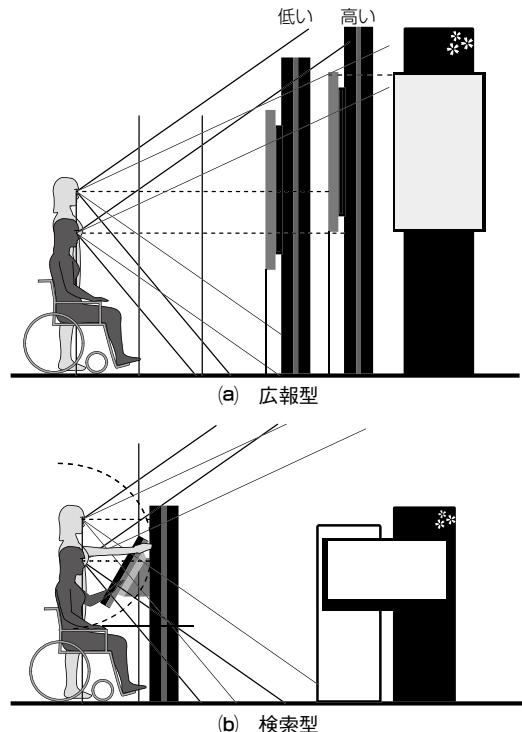


図5. 筐体デザイン

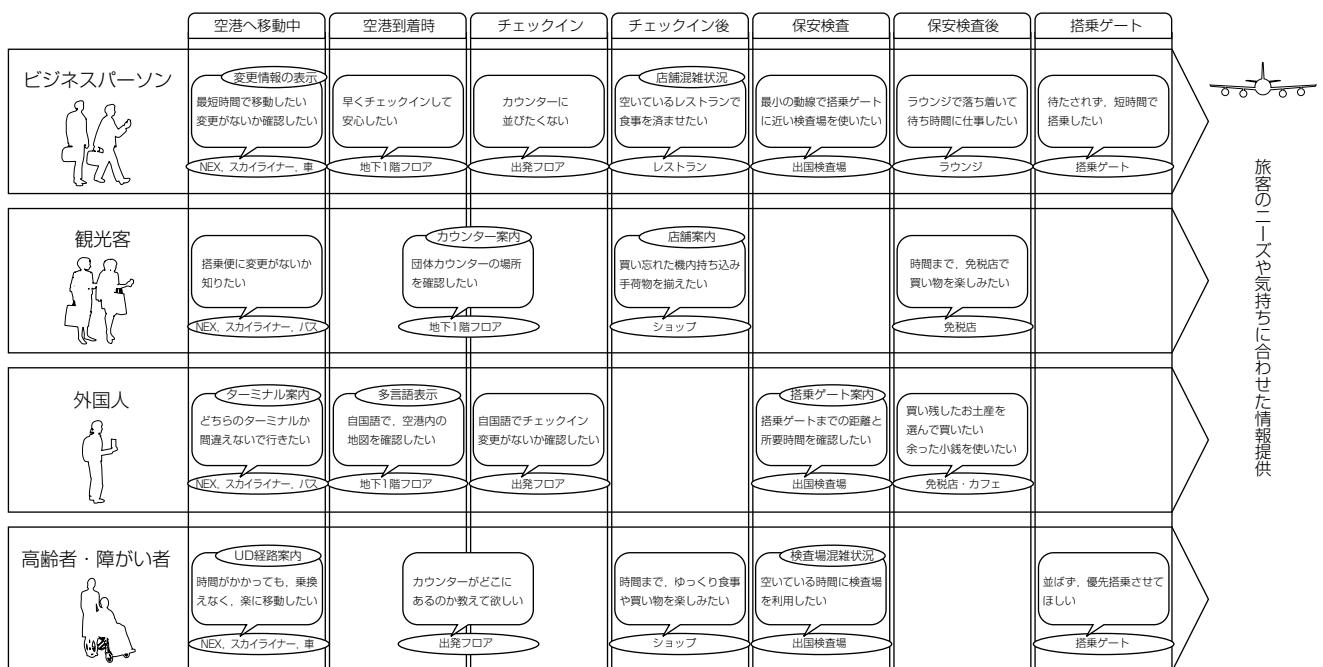


図4. 旅客のニーズに合わせた情報提供

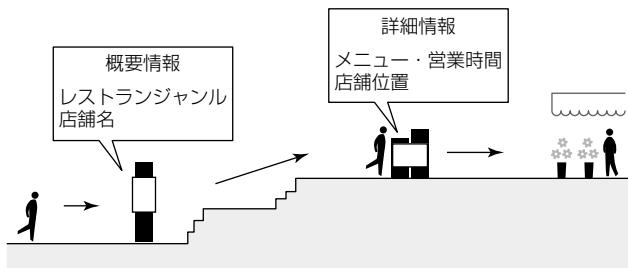


図6. 表示コンテンツの考え方

動ドアのある出入口や鉄道改札のように一瞬足が止まるような狭い空間には、筐体近くで見ることを想定し、背の低いタイプを設置した。開けた通路など見通しの良い動線には背の高いタイプを動線に対して正面に設置している(図5(a))。レストラン・店舗エリアに設置した検索型サイネージは、車いまでの利用を考え、足下に空間を作り、近寄りやすい形状とした(図5(b))。

また、表示コンテンツは図6にあるように手前で概要情報、奥で詳細情報にすることで、旅客がある程度目的を絞り込んでから検索できるようにしている。これによって検索端末に立ち寄る時間を短縮し、より多くの旅客にストレスなく利用してもらえるコンテンツ、筐体配置としている。検索方法についてはスムーズに行えるよう表示内容と機能を限定した。慣れない人でも迷わず利用できるように、①トップページ、②カテゴリー選択ページ、③店舗詳細ページの3階層だけとなっている。トップページは、従来の固定案内板としての機能を持ち、何も操作せずとも各店舗の配置が分かるようになっている。また、日英中韓の4か国語に対応しており、言語ボタンを押すと、すべてのコンテンツ言語が切り替わる。

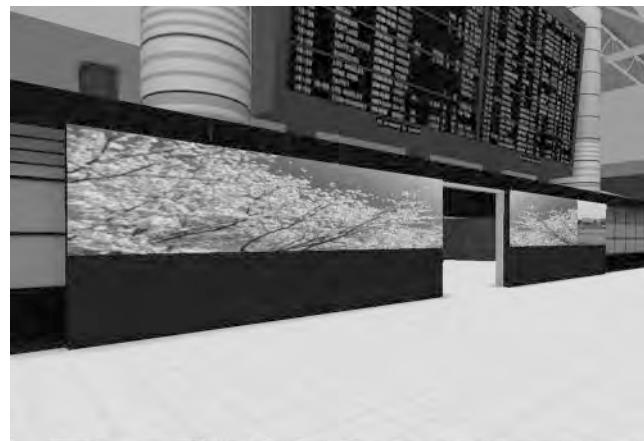
2.3.2 プラスバリュー

日本的な“おもてなしの心”的演出で、日本を訪れる人、世界に旅立つ人にメッセージを伝達する。空港全体で日本を表現し、文化、デザイン、技術が融合した空港として日本の観光産業、航空産業の発展に貢献することを目指した。

また、成田国際空港のブランドネーム“WORLD SKY GATE_NARITA”から、出発口、到着口の大画面マルチディスプレイを日本と世界を結ぶ巨大な映像の門にみたて、出発口を“送る門”，到着口を“迎える門”，レストラン・店舗エリアを“門前市”として表現した。周囲に点在する単体ディスプレイは、灯籠をイメージした形状とし、漆のようなツヤ塗装、アイコンに桜のグラフィックを用いて、日本の玄関にふさわしいデザインとした。

2.3.3 フロンティア

常に時代の先端であり続けるために、デジタルサイネージをステップ1とし、空港機能の拡張、地上交通や他空港との連携、旅客サービスのパーソナル化等、今後予想される次世代空港のあるべき姿を提示した。



(a) CGイメージ



(b) 実物写真

図7. CGイメージと実物比較

2.4 ビジュアルイメージ創出

これまでの提案は、イメージスケッチを紙の提案書としてまとめて提出していたが、成田国際空港への提案では、設置エリアが広く、コンセプトにあるシームレスな情報提供や筐体の位置関係、大画面の効果、コンテンツの表示イメージ等が伝えにくいため、空港内をCG(Computer Graphics)で再現し自由に歩き回れるプレゼンテーションツールを作成し、紙の提案書とともに提示した(図7)。これによってプレゼンテーションで高い評価を得るとともに、採用後、関係者の意識の共有化にも効果を發揮した。

3. むすび

デザインに当たり、旅客を満足させる、感動させる体験を提供するためにデジタルサイネージが担う役割は何かを考えてきた。その実現のため、ユーザーエクスペリエンスの視点から導き出される、サイネージに必要な要素、旅客の要求を抽出し、これを整理して具体化することが極めて重要である。

今後も従来取り組んできたノウハウを蓄積し、より広い分野で活用し展開することで、価値の高いデザインを提供する。

相川真実*
八木澤喬樹*
今中英夫*

コンセプトカー“EMIRAI2”のHMI

HMI for Concept-car "EMIRAI2"

Masami Aikawa, Takaki Yagisawa, Hideo Imanaka

要 旨

三菱電機は2020年の自動車社会を想定し、来るべき車社会に対応した自動車の未来像として“環境”“安全”“快適”を高度に実現するコンセプトカー“EMIRAI2(イーミライツー)”を開発した。EMIRAI2は、パワートレイン系コンセプトカー“xEV(エックスイーブイ)”と運転支援系コンセプトカー“xDAS(エックスダス)”の2台コンセプトカーとして2013年の東京モーターショーに出展した。本稿では運転支援系コンセプトカー“xDAS”に搭載された各アイテムを通じて、当社が考える未来の車環境におけるHMI(Human Machine Interface)について述べる。

“xDAS”には、車外の状況を監視する画像センサやドライバがどこを見ているか判別する視線センサ等の検知デバイス

イス、フロントガラスを利用して車外の風景に重畳する大型のHUD(ヘッドアップディスプレイ)やドライバの状況に応じて形状や表示内容を変化させる三連ディスプレイ等を用いた表示デバイス、状況に応じた入力方法を使い分けられるスマートハンドルといった入力デバイス等、様々なデバイスを搭載している。

コンセプトカーEMIRAI2“xDAS”では、“インテリアが人に調和”“情報が人に調和”“操作が人に調和”的3つのテーマに基づき、これらのデバイスを複合的に活用する様々な運転支援システムを提案している。“xDAS”は、より安全でより快適な、人と調和する自動車用HMIの未来像を示すものである。



運転支援系コンセプトカー“xDAS(エックスダス)”

xDASのHMIは、“車が人に調和”することをテーマとしてデザインし、種々のデバイスを活用した運転支援システムを搭載し、将来の自動車像を提案している。

1. まえがき

当社は、2020年の自動車社会で、“環境”“安全”“快適”を高度に実現する自動車の未来像であるコンセプトカー“EMIRAI2(イーミライツ)”を開発し、2013年の東京モーターショーに出展した。EMIRAI2は、2台のコンセプトカーからなり、その1台である運転支援系コンセプトカー“xDAS(エックスダス)”は“クルマが人に調和する”をコンセプトに、インテリア・情報・操作を人に調和させたEV(Electric Vehicle: 電気自動車)を想定している。xDASのコックピットには、当社の映像表示制御技術、画像・センサ情報処理技術等の様々な将来技術を搭載し、それらを有効に活用するHMIによって、ドライバの安全・快適な運転を支援する。

ここでは個々のHMIを“インテリアが人に調和”“情報が人に調和”“操作が人に調和”的3つのテーマごとに述べる。

2. インテリアが人に調和

ドライバが必要とする情報を、視線移動を減らして分かりやすく表示することによって、便利で快適な運転を支援する“インテリジェントインフォメーションパネル(IIP)”を搭載している(図1)。この章ではIIPにおける“状況適応ディスプレイ”と“個人適応HMI”という二つの特徴について述べる。

2.1 状況適応ディスプレイ

状況適応ディスプレイは、複数の液晶やHUD(ヘッドアップディスプレイ)等のマルチ画面の表示領域を、走行速度の上昇に伴うドライバの視野狭窄(きょうさく)に応じて可変制御することによって、ドライバに提供する情報量を最適化するディスプレイ装置である。

画面部は横配置の三連ディスプレイとなっており、中央に15インチ相当のメインディスプレイを、左右に9インチ相当の左右ディスプレイを装備している。この三連ディスプレイは速度に応じてダッシュボード上に露出する部分が可変する機構を備えており、次のように走行状況に応じて形状を



図1. インテリジェントインフォメーションパネル(IIP)

変化させ、それに応じた情報提供を行う。

(1) 停車時の形状と情報提供

三連ディスプレイすべてが最大表示領域まで露出する。この場合3つのディスプレイは1つの表示領域として制御されており、ドライバは各種アプリケーションのアイコンが浮遊する三次元空間を模した領域を、後述するスマートハンドルのタッチデバイスを用いてフリーカーソル方式で操作することができる。

(2) 市街地走行時の形状と情報提供

三連ディスプレイが(1)から高さ半分の前方視界を確保できる位置まで下降する。左右ディスプレイは車両両側後方の映像を表示し、サイドミラーの役割を担う。メインディスプレイで速度などの車両情報を表示するとともに、HUD部では次交差点案内などの簡易的なナビゲーション表示を行う。

(3) 高速道路走行時の形状と情報提供

左右ディスプレイは(2)と同じ状態で、メインディスプレイは必要最小限の情報が表示できる位置まで下降する。主要な走行関連情報はHUD部で表示する。これによってドライバは最大限の前方視界を確保できる。

このように速度に応じて表示器を最適配置／最適利用することでドライバの負荷を低減し安全運転に寄与する。

2.2 個人適応HMI

ドライバに応じて、IIPの表示内容を変化させるHMIである(図2)。

ドライバが所有するスマートフォンに車両の設定や好みの着せ替えパネルデザイン等、個人ごとの設定情報をあらかじめ入力しておくと、搭乗時にスマートフォンと車両がNFC(Near Field Communication)を用いて同期し、IIPに表示されるメータデザインが好みを反映したデザインにな



図2. 個人適応HMIのメータデザイン

るなど、ドライバに適応したインテリア空間を提供する。

3. 情報が人に調和

車を取り囲む膨大な情報の中から、ドライバにとって危険性の高い情報や、ドライバが欲している情報を抽出して、ドライバの視線に応じてIIPのHUD表示に提示する。

3.1 統合型イメージセンシング

車両に近づく危険を車外センサで検知してHUDに表示すると同時に、車内に設置した視線センサで“ドライバの気づき”を判定して、危険性が高い情報を優先的にドライバに知らせるシステムである(図3)。

車両外側に設置された前方監視カメラによって、危険因子となる可能性のある車両や歩行者の飛び出し等を識別するとともに、ドライバ前方のダッシュボード上に設置された視線検出カメラによってドライバの視線を検出し、その危険因子を視認したかどうかを判断する。視認していないと判断した場合にはHUD上にその危険因子と重畠表示される“よそ見警告”を表示することでドライバに警報を与える。このように車両周辺監視とドライバ監視の2つを組み合わせることによって、ドライバの認知測定を利用した予防安全を可能にしている。

3.2 あれ何？サーチ

“あれ何？サーチ”とは視線センサ情報と高精度自車位置情報から、車外に見える施設について知りたい情報をすぐ検索できるHMIである。



図3. 統合型イメージセンシング



図4.“あれ何？サーチ”

走行中、視線センサを用いてドライバの視線を常時検知することで、ドライバが見ている施設を特定し、HUD上にその名称を表示することができる。ドライバがその施設についての詳細情報を知りたい場合には、視線を施設に向いている時にハンドル上にある“あれ何？”ボタンを押下すると、その施設の詳細情報を合成音声で通知する(図4)。このような方法で、あたかも助手席に“あれは何だろう？”と話しかけるように、車両前方を見たまま簡単、安全に欲しい情報を得ることができる。

4. 操作が人に調和

ドライバの行動を先読みした操作の提案と、タッチパネル操作・手書き文字入力によって、簡単・安全・直観的な操作を実現するナチュラルユーザーインターフェースを取り入れた、レコマンドメニューとスマートハンドルを搭載している。

4.1 レコマンドメニュー

従来の一般的な車載情報機器の操作方法は、ドライバが希望する項目をメニュー画面で辿(たど)りながら機能を絞り込んでいく操作系をとっている。この場合、機器の多機能化が進むにつれ操作するメニュー階層は深くなり、操作時間や操作するために機器を注視する時間も増加する。このような操作系では、走行中にドライバが操作することは非常に危険を伴う。そのため、走行中に操作制限を設け、一定の操作時間以上のメニュー操作や、機器注視時間が長くなるメニュー表示を制限するなど、車載情報機器では走行中にその機能を十分に活用できないケースも見受けられる。

そこでこのレコマンドメニューでは、ドライバの希望する項目を、その時の車載機器(ナビゲーション、オーディオ、エアコン等)の設定状況や、車外環境の状況(車速や現在位置情報、時刻、天候、道路状況等)から推測・先読みして提示する。ドライバの操作は希望する項目を選択するだけの単純な操作に限定することで、操作時間、機器注視時間を劇的に短縮することを目指している。しかしシステム側でドライバの希望する項目を100%正しく推測することは不可能であるため、表示項目から漏れた先読み候補を、再表示できるアプリケーションボタンを設けることでこれを補うよう設計した。

レコマンドメニューは、①IIP上に表示された3択メニュー、②ハンドルを握った時に親指が自然に来る位置に配置された3択ボタン、③レコマンドメニューで操作できる4つのアプリケーション(ナビゲーション、オーディオ、エアコン、電話)に対応した、ハンドルパッド上の4アプリケーションボタンの3つの要素で構成されている。

通常操作では、レコマンドメニューはその時の状況から先読みした優先度の高い上位3つの候補を3択メニューに



図5. 3択メニュー表示



図6. 3択ボタンと4アプリケーションボタン部

表示する(図5)。ドライバはこの3択メニューの中から希望する項目に対応した3択ボタンを押下するだけ、すなわち1回の操作で実行できる。

先読みされた3択メニューにドライバが希望する項目が無かった場合には、4つのアプリケーションボタンの中からユーザーが希望する機能に対応したボタンを押下することで、最初に先読みされた3択メニューよりも優先度の低かった項目から、選択されたアプリケーションに対応する項目が表示される(図6)。

ドライバが希望するアプリケーションを入力することで、推測された次候補を表示する場合に比べ、希望項目が次の操作で選択できる確率は高くなる。レコマンドメニューは、このような操作系とすることで、“操作回数2回以内、総操作時間15秒以内に操作完了の実現”を目指すものである。

走行中のドライバはハンドル上の3つのボタンに対応するメニューを選択するだけという単純な操作で目的を達成できるため、レコマンドメニューは操作回数・時間の観点だけではなく、操作の理解のための認知的負荷も低減可能であり、安全な運転に寄与する。

4.2 スマートハンドル

ハンドルに曲面タッチパネルを搭載したハンドル型操作デバイスである。カーソル操作や手書き文字入力等の直観的な操作性を実現するとともに、任意の曲面形状を実現するタッチパネル技術によってデザイン性の高さを両立させる(図7)。

スマートハンドルのタッチパネルは、停車時にIIPの表示と連動して使用する。ホーム画面上アプリケーションアイコンのカーソル操作や各種アプリケーションでの手書き文字入力による検索機能に対応している。また、マルチ



図7. スマートハンドル



図8. 文字入力時のスマートハンドル

タッチに対応しており、ピンチイン・アウトによる画面拡大・縮小操作や、スワイプ、タップ、ダブルタップによる選択・決定操作や文字入力操作等、スマートフォンと同等のタッチパネル操作を可能にしている(図8)。

また、操作する画面と実際の操作部が離れている遠隔操作タッチパネルとなるため、ユーザーが実際に触れているポイントを発光させタッチした軌跡を表示することで、より分かりやすい操作感を実現している。

5. むすび

新しいHMI提案を盛り込んだ運転支援系コンセプトカーEMIRAI2“xDAS(エックスダス)”は、パワートレイン系コンセプトカーEMIRAI2“xEV(エックスイーブイ)”と合わせて2013年東京モーターショーで展示し、一般来場者並びに自動車メーカー関係者に実際に体験いただき、高い評価を得るとともに種々の意見も得られた。今後はこれらの意見を考慮し、来るべき未来の車社会に向けて、より安全でより快適な自動車用HMIの開発を推進する。

三菱HEMSのUXデザイン

飯澤大介*
北村憲史郎*

UX Design for Mitsubishi HEMS

Daisuke Iizawa, Kenshiro Kitamura

要 旨

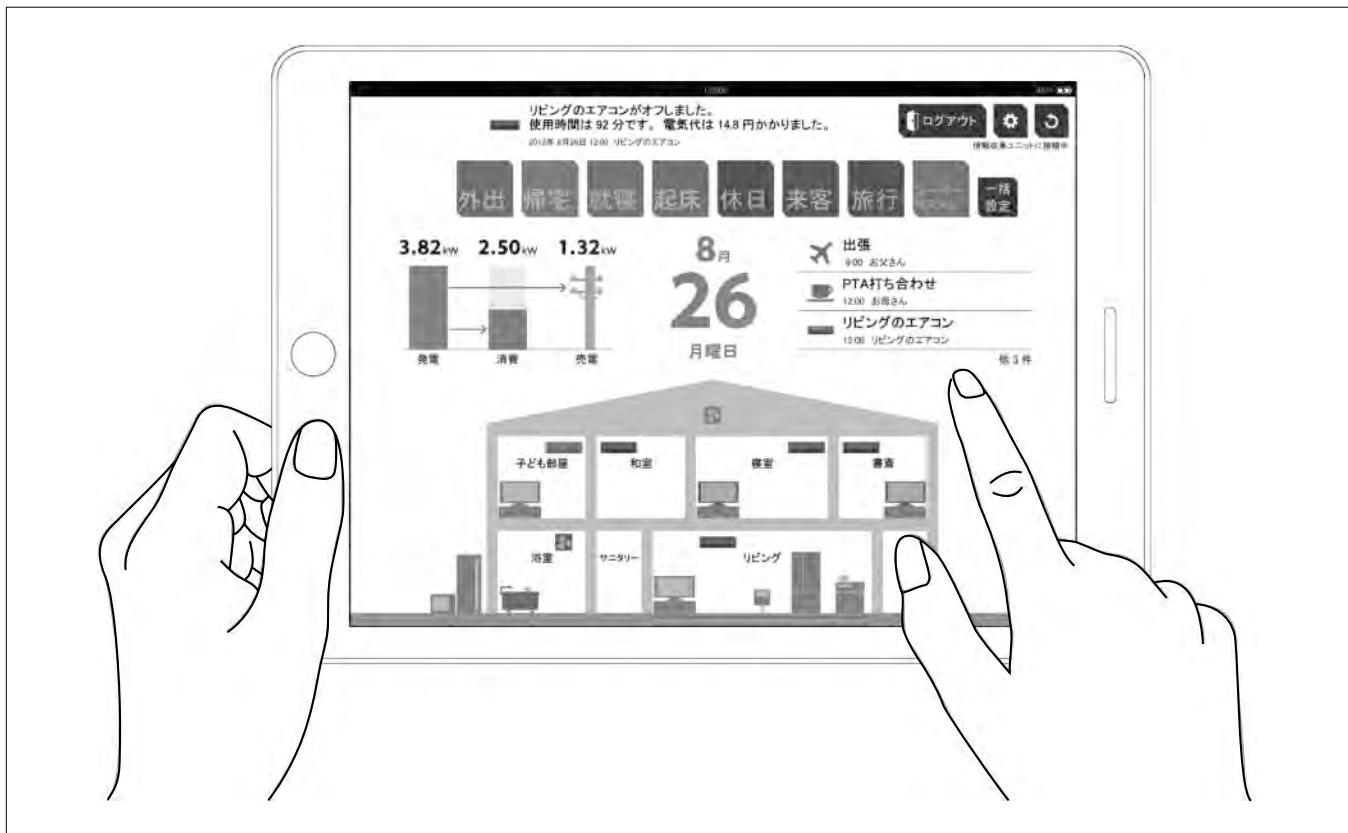
2013年8月に発表した三菱HEMS(Home Energy Management System)に対応するタブレット端末アプリケーションは、使用電力の管理や機器制御といったエネルギー・マネジメントだけでなく、家族のコミュニケーションなどのライフマネジメントというコンセプトを実現している。

スマートハウス、エネルギー・マネジメントシステムの導入は拡大しているが、一般ユーザーにとってHEMSコントローラはいまだに馴染(なじ)みが薄く、機器制御などの機能に関する知識も十分ではない。このような状況でライフマネジメントを実現するアプリケーションを設計するためには、ユーザーが実際に製品を手にする前から利用後までの長い期間におけるユーザーと製品とのインタラクション

を想定し、ユーザー経験をデザイン(UXデザイン)することが不可欠である。

そこで、アプリケーションとユーザーの接点を“導入期”“利用初期”“利用中期”“利用後期”的4段階に分類し、それぞれの接点におけるデザイン上の配慮すべき点、要求を整理し、これに基づいてデザインを行った。

開発したタブレット端末アプリケーションは、HEMSに関する知識があまり多くないユーザーが、エネルギー・マネジメントの見える化を確認するといった簡単な利用から、家族とコミュニケーションしながら機器の節電運転をカレンダーに登録する比較的高度な操作までを、段階的に学んでゆくUXデザインを実現している。



三菱HEMSのタブレット端末アプリケーションのトップ画面

タブレット端末アプリケーションのトップ画面のデザインは、画面を構成する情報の優先度を慎重に検討した上で、色や線、影等の視覚情報の量を極力抑えたシンプルなデザインを採用している。また、家の立面図をモチーフとした間取りコントローラを中心レイアウトすることで、ユーザーに馴染みのある印象を与えることを意図している。これらのグラフィックデザインとしての配慮は中長期的なUXデザインの一例である。

1. まえがき

社会のエネルギーに対する意識が高まる中、三菱電機は2012年に社会全体の課題を解決しながらひとりひとりの“暮らしのクオリティ”を高めていくという“スマートクオリティ”コンセプトを発表し、2013年8月にはスマートハウス関連事業への本格参入と同時に“三菱HEMS”を発表した。

HEMSとはホームエネルギー・マネジメントシステムの略であり、家庭で消費又は発電されるエネルギーを管理して効率よく運用するためのシステムのこと、近年のスマートハウスの導入拡大に合わせて市場への導入が進んでいる。

本稿では三菱HEMSのコントローラとなるタブレット端末アプリケーション(以下“アプリ”という。)におけるユーザー経験のデザイン(UXデザイン)の取組みについて述べる。

2. 三菱HEMS

三菱HEMSは7種類のHEMS対応製品に加え、太陽光発電システム、蓄電池、ガス流量計、水道流量計を接続することが可能で、快適性を維持しながらトータルエネルギー・マネジメントを実現するシステムである。

7種類のHEMS対応製品はルームエアコン、エコキュート、IH(Induction Heating)クッキングヒーター、換気システム、ヒートポンプ式温水暖房システム、冷蔵庫、液晶テレビであり、それらの機器情報は、無線LAN機能を持つ専用のアダプタを介し、情報収集ユニットに集約される。また、エネルギー計測ユニットが分電盤から家全体の使用電力量を計測し、さらに水道・ガスのパルス式流量計からそれぞれの使用量を計測して、情報収集ユニットに送信する。

これら情報をもとに、タブレット端末にインストールしたアプリで、エネルギー使用量の見える化グラフ、節電目標に応じた自動機器制御、各家庭の間取りを再現した画面上での機器操作、家族の生活に合わせた節電を実現するファミリーカレンダー等の機能を利用することができる。

3. UXデザイン

三菱HEMSのタブレット端末アプリにおけるインターフェイションデザインのコンセプトは、ユーザーにとって負担と感じる可能性のある節電行為を、安心・安全の確認から、積極的な行動へ、そして最終的には家庭内コミュニケーションといったライフマネジメントにまでつなげるというものである。このコンセプトを実現するために開発初期段階からユーザー・エクスペリエンス・デザイン(以下“UXデザイン”という。)に注力したインターフェイションの設計を行った。

一般ユーザーにとって、HEMSの機器構成や制御内容は非常に複雑なものであるが、目標電気代に応じた自動節電機能のように、複雑な制御を自動化することで簡単な操作性を実現している。一方で、接続された製品のコントロ

ルなどユーザーが手動で操作することで利便性や快適性を更に向上させることも可能である。しかしこれを実現するためには、アプリを通じたHEMSの活用方法をユーザーに理解してもらう必要があった。開発当初はHEMSというシステムがまだ世の中で一般的でなかったため、HEMSを利用する前の導入期から、利用初期、利用中期、利用後期までの期間⁽¹⁾に、ユーザーがシステムに関する知識や節電行動の方法を段階的に獲得できるように配慮した設計を行った。

3.1 導入期

三菱HEMSの購入検討の段階で、カタログやホームページ、広告等の媒体を通じて、ユーザーがアプリ画面などの情報に触れる際のUXデザインについてまとめる。

3.1.1 シンプルなグラフィックデザイン

タブレット端末アプリのグラフィックデザインとしての課題は、HEMS対応製品7機種最大16台の接続機器という多量の情報を分かりやすく伝えることであった。画面内の表示情報や複雑な階層構造はHEMSの導入を検討しているユーザーに、難しい機器であるという印象を与えるため、誰にでも親しみやすい家の立面図をモチーフとした間取りコントローラを中心とした画面をレイアウトした。さらに、全ての画面のグラフやボタンなどの構成要素でも、塗りや線、影等の装飾的な要素を極力少なくしたシンプルなデザインを採用して視覚情報の削減に努めた。

間取りコントローラの画面の一例を図1に示す。親しみやすさと見やすさを両立させ、導入期のユーザーに簡単なアプリであるという印象を持ってもらうことを意図したデザインとしている。

3.1.2 ライフマネジメント

タブレット端末アプリでは、一般的なエネルギー・マネジメントと機器操作以外に、ユーザーの生活を支援するライフマネジメントという新たなコンセプトを追加している。

その一例は家族間のコミュニケーションと節電行動を組み合わせた、業界初^(注1)のファミリーカレンダー機能である。できるだけ多くの家族が、家庭の節電や快適な暮らしの実現に参加できるように搭載した機能で、生活を豊かにするライフマネジメントの先駆けとなる機能である。これによって、タブレット端末アプリが、利用者を限定しない幅



図1. 機器構成をシンプルに表現した間取りコントローラの画面

広いユーザーのためのアプリであることを明確にしている。

(注1) 2013年8月26日現在、当社調べ

3.1.3 汎用タブレット端末の利用

タブレット端末アプリはアンドロイド^(注2)OSを搭載したタブレット端末にアプリをインストールし使用する方式を採用している。専用のコントローラハードウェアを必要とせず、ユーザーが既に持っているタブレットでHEMSを操作することができるので導入の敷居を下げるとともに、日常的な利用を促進する効果を期待している。同時に、複数人の家族がそれぞれのタブレット端末で操作を行うことが可能な点も、ライフマネジメントにとって重要である。

(注2) アンドロイドは、Google Inc. の登録商標である。

3.2 利用初期

三菱HEMS導入後、ユーザーがアプリを操作し始める際のUXデザインについてまとめる。

3.2.1 ガイダンス

ユーザーが三菱HEMSを導入して初めてアプリを起動した際に、使い方を文章とアニメーションで述べるガイダンス画面(図2)が表示される。このガイダンスの目的は2つある。1つは、三菱HEMSが行う様々な制御の説明である。操作直後にその場で起こる変化だけでなく、一定時間経過後に制御される内容などについてアニメーションを活用することで、分かりやすく説明している。もう1つは操作方法のシミュレーション効果を狙ったものである。ガイダンスのアニメーションで、GUI(Graphical User Interface)の操作手順を示すことによって、機能利用の障壁を下げるとともに積極活用の意欲を向上させる効果を狙っている⁽²⁾。

このガイダンスは初回起動後も、トップページ上のリンクからいつでも再生できるようにしている。

3.2.2 家族登録

アプリ利用開始時に必要な一連の設定手続きの中に、ファミリーカレンダーに表示する家族情報の入力がある。ファミリーカレンダーは先に述べたとおりライフマネジメント機能であるが、ユーザーにとっては他のアプリなどで利用したことのない新しい機能である。そのためUXデザインの視点としては、初回登録時に少しでも機能利用のイメージを持ってもらい、継続して利用してもらうようにす



図2. ガイダンス画面

ることが重要である。そこで、何もない状態から家族を登録させるのではなく、家族登録欄の初期状態の表示人数を3名とし、家族構成をイメージさせる顔アイコンを用いて、家族での利用を促すデザインとした(図3)。家族の追加や削除が容易に行えるだけではなく、顔アイコンはシルエットを用いたものを複数準備し、タブレット端末や任意の場所からユーザーが所持している顔写真なども活用できるようにして、機能に対する親しみと積極的に利用したいという意識の向上を意図した設計としている。

3.2.3 目標値設定

アプリ利用開始時の設定項目には電気代の目標値設定も含まれている。ここで設定された目標金額を実現するためには、システムが日々の電力使用状況から使用量を予測して機器の自動制御を行う。これは自動制御のための設定という位置付けであるが、UXデザインの視点としては、ユーザーにとってのゴールセッティングの役割も持っております。電気代という明確で分かりやすい目標値を設定することで、節電行動へのモチベーションが高まることを期待している。

3.3 利用中期

利用開始の段階を経て、日常的な利用を繰り返す際のUXデザインについてまとめる。

3.3.1 段階的なUXの設計⁽³⁾

タブレット端末アプリに搭載された機能は、ユーザーのシステムやアプリに対する知識やスキルと、節電に対する意欲、それぞれの向上の度合いに合わせて、次第に使いこなせるようなUXに配慮した構成となっている。エネルギーの見える化グラフを“見るだけ”から始まり、カレンダーでの節電“行動”につなげ、また、デフォルト設定の活用による簡単操作から、詳細なカスタマイズを含めて操作ができるよう、次第に理解と意欲が高まるUXデザインを行っている(図4)。

3.3.2 お知らせによるきっかけ

タブレット端末アプリには、お知らせ機能が搭載されており、目標の達成度などをシステム側から伝達することで、ユーザーが節電行動を起こすきっかけを与える役割と、お知らせをタイムライン上に表示することで、システムの動作を見える化し、ユーザーがそれを学習する機会を提供する役割を担っている。このように、この機能には3.3.1項で述べた



図3. 家族登録初期画面

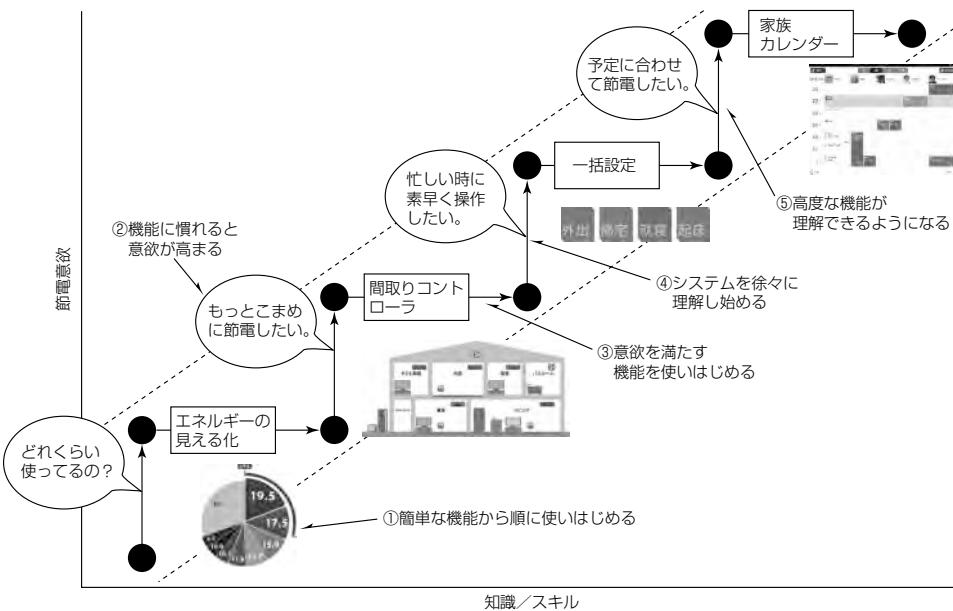


図4. 段階的なUXの設計



図5. ファミリーカレンダーによる節電行動の見える化

段階的なUXの設計を推し進める役割としての側面もある。

3.3.3 家族間コミュニケーション

ファミリーカレンダー機能は、家族間で予定を共有するとともに、行動予定に合わせて行ったHEMS対応機器の設定など、節電行動を見える化、把握・共有ができる(図5)。家族でのコミュニケーションの場で、お互いの節電意欲に影響を与え合う効果も期待できる。

3.4 利用後期

日常的な利用がある一定期間経過した後に起こる変化を想定したUXデザインについてまとめる。

3.4.1 間取りの編集

間取りコントローラ機能は、ユーザーの住む住宅の間取りを立面図上で再現することで、機器構成の分かりやすさを実現しているが、それと同時にそのユーザーのためだけにカスタマイズされた表示は、ユーザーが愛着を持ってHEMSを利用してもらうための重要な要素となっている。長期的な利用後に発生する、新たなHEMS接続機器の追加、部屋の模様替えや機器の移動に対応できるように、シンプルで簡単なUI(User Interface)で間取りや機器の編集操作を実現している。

3.4.2 ファミリーカレンダーによる思い出の蓄積

ファミリーカレンダーへのイベント登録は、節電に関連するものだけでなく、一般的なスケジュールも登録できるように設計されているため、利用を続けた結果、アプリ内に家族の思い出となる情報が蓄積されることになる。これはタブレット端末アプリが単なるシステムのコントローラの役割だけでなく、家族全体のライフマネジメントの役割を担うという位置付けとなることを意図している。

4. むすび

HEMSアプリ向けのUXデザインについて、デザインプロセスとコンセプトについて述べた。複雑なシステムとユーザーとの関係を中長期的な視点から捉え、利用経験の中で、理解が深まり、高度な機能が自然に、かつ積極的に利用できるようになることを目指している。今後は、設計意図についての評価を行い、評価結果に基づく再設計のサイクルについても検討を行う予定である。今後も、UXデザインによるエネルギー管理とライフマネジメントを両立させる製品開発を通じ、快適で、安全・安心な暮らしの基盤づくりに貢献していく所存である。

参考文献

- (1) Newbery, P., et al.: Experience Design: A Framework for Integrating Brand, Experience, and Value, John Wiley & Sons Inc. (2013)
- (2) Fogg, B.J. : A Behavior Model for Persuasive Design, Proc. of Persuasive2009, 40 (2009)
- (3) Csikszentmihalyi, M. : Flow : The Psychology of Optimal Experience, Harper Perennial Modern Classics (1990)

音声対話技術

Spoken Dialog Technology

Youichi Fujii, Yusuke Koji, Jun Ishii

藤井洋一*
小路悠介**
石井 純*

要 旨

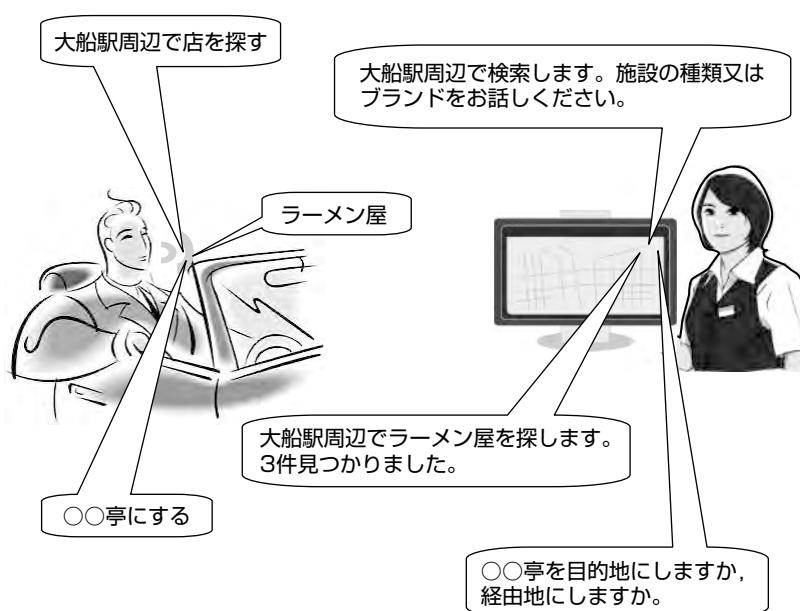
情報システム、カーナビゲーション・AV(Audio Visual)機器等の情報家電、さらにはルームエアコンなど電化製品も、多機能化が進んでいる。これらのシステム・機器のインターフェースを従来のGUI(Graphical User Interface)で実現しようとすると、メニュー構造が複雑化し、ユーザーは実行したい機能に辿(たど)り着くまでに“どう操作していいかが分からない”という問題が生じる。

この課題を解決するための技術として、音声認識への期待は高い。しかし従来型の音声コマンドでは、多機能化によってコマンド数が増え、GUI同様に“何を言っていいかが分からない”という問題を引き起こす。これを解決するには、自由な発話から意図を推定してシステム・機器を動

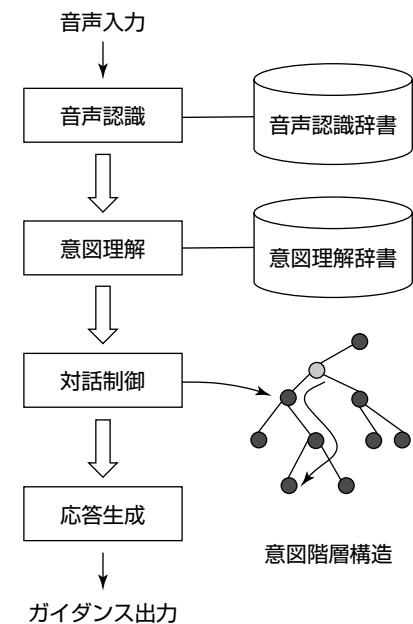
作させることが必要である。さらに、ユーザーの発話で動作に必要な情報が不足していたり、意図が曖昧であったりしたときにこれを解消するための対話技術が必要となる。

そこで、我々は、意図階層構造を用いる音声対話方式を開発した。意図階層構造は、操作対象機器に関する機能定義から生成され、上位の抽象的な意図から下位の実行可能な意図までを階層的に表現する。発話からこの構造中の意図を推定し、推定に曖昧性があれば確認し、下位の実行可能な意図に到達するように、対話が制御される。

この方式によって、複雑で多機能なシステム・機器に効率的で、自然なインターフェースを実現することができる。



音声対話例



音声対話ブロック図

音声対話の対話例とブロック図

音声対話システムでは、ユーザーが自由に発話した内容から意図(ユーザーが実行したい機能)を推定するために意図理解辞書を使う。対話制御部では、推定した意図から実行が可能な下位の意図表現に進むように対話を制御する。末端の実行可能な意図へ到達した時点で対応付けられた機能を実行することでユーザーの目的を達成する。

1. まえがき

各種の情報システムや、カーナビゲーション(以下“カーナビ”という。)・AV機器等の情報家電、さらにはルームエアコンなど電化製品にいたるまで、システム・機器の多機能化が進んでいる。これらのインターフェースを、GUIやスイッチ類で実現すると、メニューの階層が増えて複雑化するため、ユーザーが実行したい機能に辿り着くためにどう操作すればいいかが分からずという問題が生じる。さらには、どのような機能が存在するのかさえ分からずの場合もある。ユーザーが機能と操作を理解するには取扱説明書を読み、何ができるかを把握し、その手順を覚える必要がある。しかし、多機能化に伴い膨大となった説明書を読むユーザーは稀(まれ)であり、多様で高度な機能の恩恵をユーザーが享受できない状態になる。

この問題を解決し、ユーザーが所望の機能を簡単に実行するための手段として音声認識への期待が高い。音声認識の場合、操作の直接的な表現を用いることができ、多くの語彙を扱えることから有効な手段になり得る。しかし、従来型の操作に対する音声コマンドを設計者視点で決めてしまう方式では、機能が増えれば、コマンド数が増え“何を言つていいかが分からない”というGUIと同じ問題を引き起こすことになる。

そこで、誰でもが音声によって簡単に操作ができるよう、ユーザーの自由な発話を理解する技術による音声対話インターフェースが研究・提案されている⁽¹⁾。音声対話には、ユーザーがシステム・機器を操作して目的を達成するための課題遂行型音声対話と、対話自体を楽しんだり、対話中に情報を得ることを期待したりする非課題遂行型音声対話がある。今回の対象は、課題遂行型音声対話である。

課題遂行型音声対話では、まずユーザーの発話から意図(ユーザーが実行したい機能)の推定を行う。その上で、ユーザーが機能についての理解が不十分であったり、発話自体が曖昧であったりした場合には、その確認を行い、目的の実行に不足する情報があれば、その入力を求め、機能の実行に導くことが対話の目的となる。これを実現するために、すべての対話状況を網羅的に記述することは現実的ではない。

そこで、我々は、対象システムの機能の表現でもあり、ユーザーの対話行動の表現モデルでもある意図階層構造に基づく音声対話システムを開発した⁽²⁾。

2. 音声対話システム

2.1 音声対話システムの特長

開発した音声対話システムの特長は、次の点にある。

(1) 対象システムの機能定義を利用してることで、意図の表現が容易にできる。

(2) 意図を階層化し、対話行動を表現する意図階層構造を構築することで、対話手順を記述することなしに汎用的な対話制御が可能である。

(1)の意図は、設計仕様書や取扱説明書から作成する。意図は単純な表現形式であり設計者は簡単に記述ができる。意図の間の関係性も考慮する必要があるが、これらの書類の構成や、GUI上の関係性等を参照することで、容易に記述することが可能である。ここで定義された意図は、3.1節で示す方法で階層構造化される。(2)の対話制御は、この意図階層構造を使って対話を進める。階層化は上位の抽象的な表現から下位の実行可能な意図表現を関連付けているため、対話の目的である、曖昧性の解消は、発話から意図が一意に推定できない場合の確認として、また、情報の取得は、上位意図から下位意図への遷移として実現できる。意図階層構造が対象機能並びに対話のモデルとして動作するため、対話手順を記述することなく音声対話システムが構築できる。

2.2 音声対話システムの構成

音声対話システムは、大きく分けて音声認識、意図理解、対話制御及び応答生成の4つの部分から構成される(図1)。音声認識は、ユーザーの自由な発話を受け付けることができるディクテーション型の認識エンジンを用いる。次に意図理解は、音声認識結果のテキストから、ユーザーの意図を推定・理解する。意図理解は、単語と意図の関係の統計的性質を記述した意図理解辞書によって実現する⁽³⁾。対話制御は、2.1節で述べたように意図階層構造を下位へ辿ることで、ユーザーの意図を具体化する。応答生成は、システムが次に期待する発話をユーザーに促すための応答を生成してガイダンスを出力する。この4つの処理を繰り返すことで対話を進める。最終的に、末端の意図に到達したら、その意図に割りつけられた機能を実行して対話を終了する。

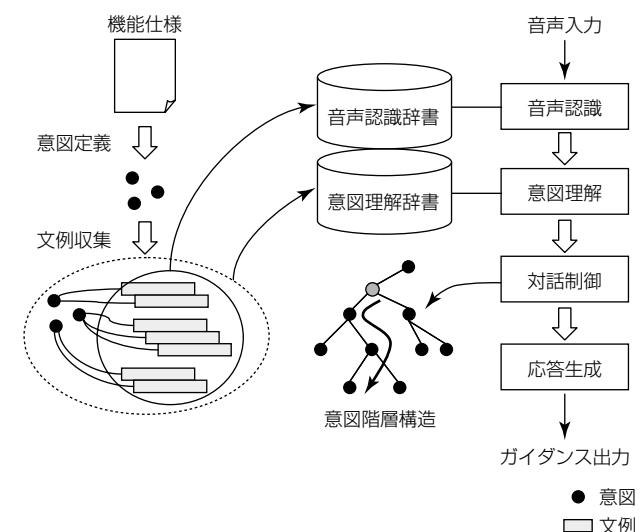


図1. 音声対話システムの構成

3. 意図の階層性を利用した対話制御方式

3.1 意図の定義と意図階層化方式

課題遂行型音声対話では、発話の目的はシステム・機器の機能を実行することであるため、これを発話の“意図”として概念化した。カーナビを例として述べる。例えば“現在地周辺のコンビニを検索したい”という意図は図2のように記述する。意図は意図の分類、又は機能を表す主部と、主部の実行に必要な情報を表すスロット部からなる。図2の例では主部が“周辺検索”で、基点というスロットの値が“現在地”，ジャンルというスロットの値が“コンビニ”である。意図は対象とするカーナビの設計仕様書、取扱説明書等の文書から抽出できる。

意図の集合が決まると、3.3節で述べる対話制御で利用するための意図の階層構造を生成する。階層構造の例を図3に示す。この意図階層構造は上位にいくほど抽象的な表現となる意図の上位・下位関係を示す構造である。意図はこの階層構造のノードであるため、構造上は意図ノードと呼ぶ。中間に位置する意図ノードはスロット情報の一部が空(スロット名=NULLと表現)のノードとなる。中間の意図ノードは最下層の意図ノードから自動で生成できる。

3.2 意図理解方式

ユーザーの自由な発話から、先に述べた意図を推定する意図理解には統計的手法を用いる。

この手法ではあらかじめ収集した発話文例から抽出した単語と、発話の意図を対応付けた学習データを用意する。例えば、“スカイツリーを目的地に設定”という文例から“スカイツリー”“目的地”及び“設定”という単語を抽出し、“目的地設定(簡単のためスロット部は省略)”という意図に

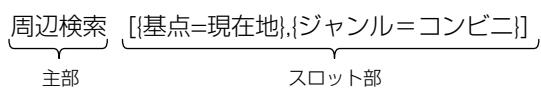


図2. 意図の例

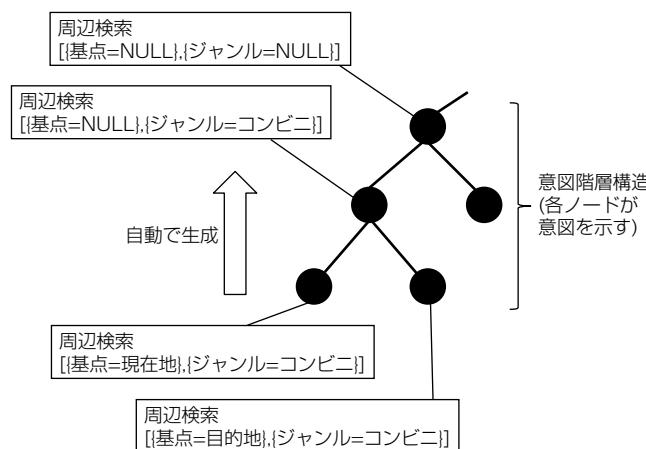


図3. 意図階層構造の例

対応付けておく。このデータから各単語が出現したときの対応付けられた意図の確率を学習し意図理解辞書とする(図4)。

実際に意図理解するときは、発話中に出現した各単語に対する各意図の確率を意図理解辞書から取得し、確率の高い意図を意図理解結果として出力する。例えば“目的地を削除”という発話の場合、意図理解辞書を参照し“目的地”に対応して“目的地設定”や“目的地削除”という意図の確率が高く、“削除”に対応して“目的地削除”や“経由地削除”という意図の確率が高い値が得られる。意図ごとの確率を掛け合わせることで“目的地削除”的確率が最も高いという結果が得られる(図5)。

この手法の特長は、学習データに含まれない単語の組合せの発話であっても、その単語が学習データにあれば意図を推定することができる点である。そのため、決められた音声コマンドではなく自由な発話から意図を推定することが可能となる。

3.3 対話制御方式

対話制御では、3.2節で述べた意図理解方式を用いてユーザーの最初の発話から、意図階層構造上の最も近い意図ノードを推定する。そこから階層を下りるための対話を繰り返し、最終的に末端の意図ノードへ遷移することでユー

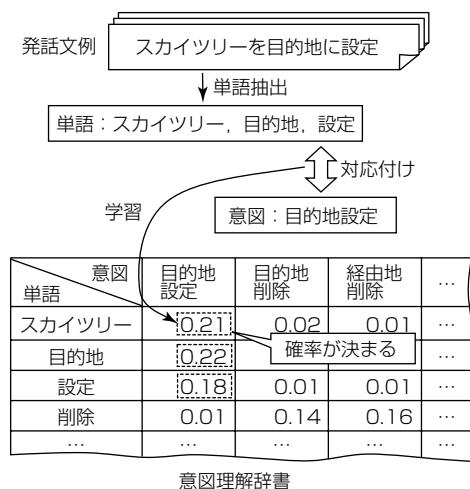


図4. 意図理解辞書の作成例

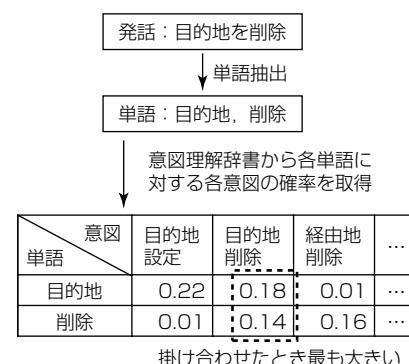


図5. 意図理解の例

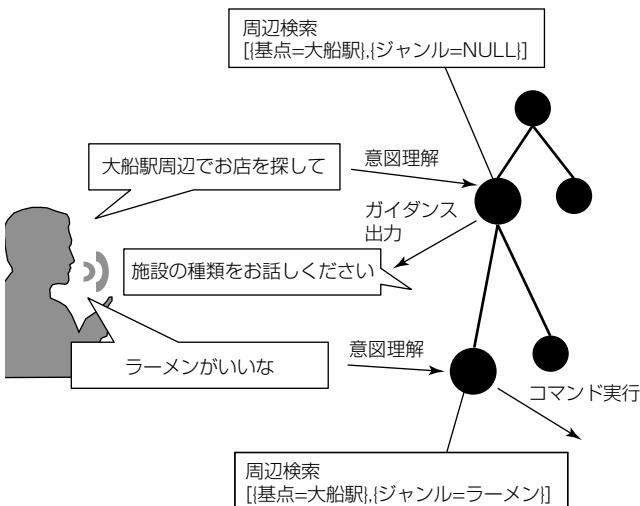


図 6. 対話制御処理の例

ユーザーの意図を具体化し、機能を実行する。

図 6 に例を示す。例えば、“大船駅周辺でお店を探して”とユーザーが発話した場合、対応する意図ノードが“周辺検索[基点=大船駅], {ジャンル=NULL}]”であると推定する。この意図ノードはジャンル情報が不明な中間ノードで、下位の意図ノードはジャンル情報が特定された意図ノードである。そのため、階層を降りるためにジャンルの情報が必要である。そこで、システムは“施設の種類をお話しください”とガイダンスを出す。ユーザーが“ラーメンがいいな”と発話すれば、ジャンルが“ラーメン”であることが特定でき、末端のノードへ遷移し、大船駅周辺のラーメン屋を検索する機能を実行する。

この例は、ユーザーの発話意図が唯一に推定できる場合であるが、実際には発話の意図を唯一に推定できない曖昧な発話の場合もある。そのような場合には曖昧性を解消する処理を行う。具体的には、推定した複数の意図ノードから一つを選択させるための対話をを行う。例えば、ユーザーが“スカイツリー”とだけ発話した場合、スカイツリーを目的地に設定したいのか、スカイツリーを登録地に追加したいのか曖昧である。このような場合、システムは“スカイツリーを目的地に設定しますか、登録地に追加しますか？”とガイダンスを出す。ユーザーが“目的地に設定”と発話すると前の発話である“スカイツリー”と合わせて意図理解を行い、スカイツリーを目的地に設定する機能を実行する。

4. 今後の展開

この音声対話方式は、多機能化する機器のインターフェースとして有効であるが、幾つか課題も明らかになっている。意図の定義を機能仕様から定義して階層化するため設計しやすくなっているが、意図理解辞書の学習のために各意図に対してある程度の規模で文例を収集する必要がある。文例データを集めるには、アンケートなどの方法をとるが、

意図と文例との対応付けは人手で行う必要があるため、開発にかかる作業量が大きくなる。文例収集及び意図対応付けが、簡単にできる方法の確立が必要である。

また、意図の定義は機器の機能から作成されるので、ユーザーが機能にない発話を行った場合には必ずしもうまく対話が行えない。例えば、“経由地の削除”と“経由地の追加”は機能として存在するが、“経由地の変更”は2つの機能の複合であるため、1機能として用意されていない場合がある。このような場合にユーザーの“経由地の○○を□□に変更して”という発話を、“経由地の削除”と“経由地の追加”が複合した意図であると理解できる仕組みが必要である。

さらに、この音声対話方式は、意図階層性を使って下位の階層に遷移することでユーザーの意図を具体化していくので、一度間違った意図に遷移すると1つ前に戻って別の言い方をするなど、対話が冗長になる場合がある。

今後これらの問題を検討し、より使いやすい音声対話インターフェースの実現を図る。

5. む す び

音声対話技術として、機能と対話行動のモデルである意図階層構造を利用する対話方式を提案した。この方式は、自由な発話からの意図理解に基づき、対話の目的である、機能実行のための情報の取得と、曖昧性の解消を、その手順を記述することなく実現でき、ユーザーには分かりやすく、効率的なインターフェースを提供できる。さらには、ユーザーがシステム・機器の機能について十分な理解がなくとも、発話の曖昧性解消の対話によって、機能を知ることができる。したがって、機能とその操作手順を事前に学習しなくとも、利用経験を通じて、機能を知り、かつ、より効率的な発声すなわち意図の表現方法も学べることがユーザー視点での大きな特長になる。

さらに、よりユーザーの感覚に近い対話を実現するため、音声だけではなく、ジェスチャー、システムの状況、ユーザー操作履歴等の複数の入力情報・文脈情報・知識を活用できる対話方式の研究・開発を行い、自然でユーザー満足度の高いインターフェースを実現する予定である。

参 考 文 献

- (1) 河原達也：話し言葉による音声対話システム、情報処理, 45, No.10, 1027~1031 (2004)
- (2) 藤井洋一, ほか：取扱説明書の構造を利用した音声対話の評価、電子情報通信学会総合大会講演論文集D-14-3 (2014)
- (3) 藤井洋一, ほか：自由発話による家電機器操作の多言語適用に関する検討評価、電子情報通信学会総合大会講演論文集D-14-6 (2012)

テキスト音声合成技術

Text-to-Speech Technology

Takahiro Otsuka, Keigo Kawashima, Satoru Furuta, Tadashi Yamaura

大塚貴弘* 山浦 正**
川島啓吾**
古田 訓**

要 旨

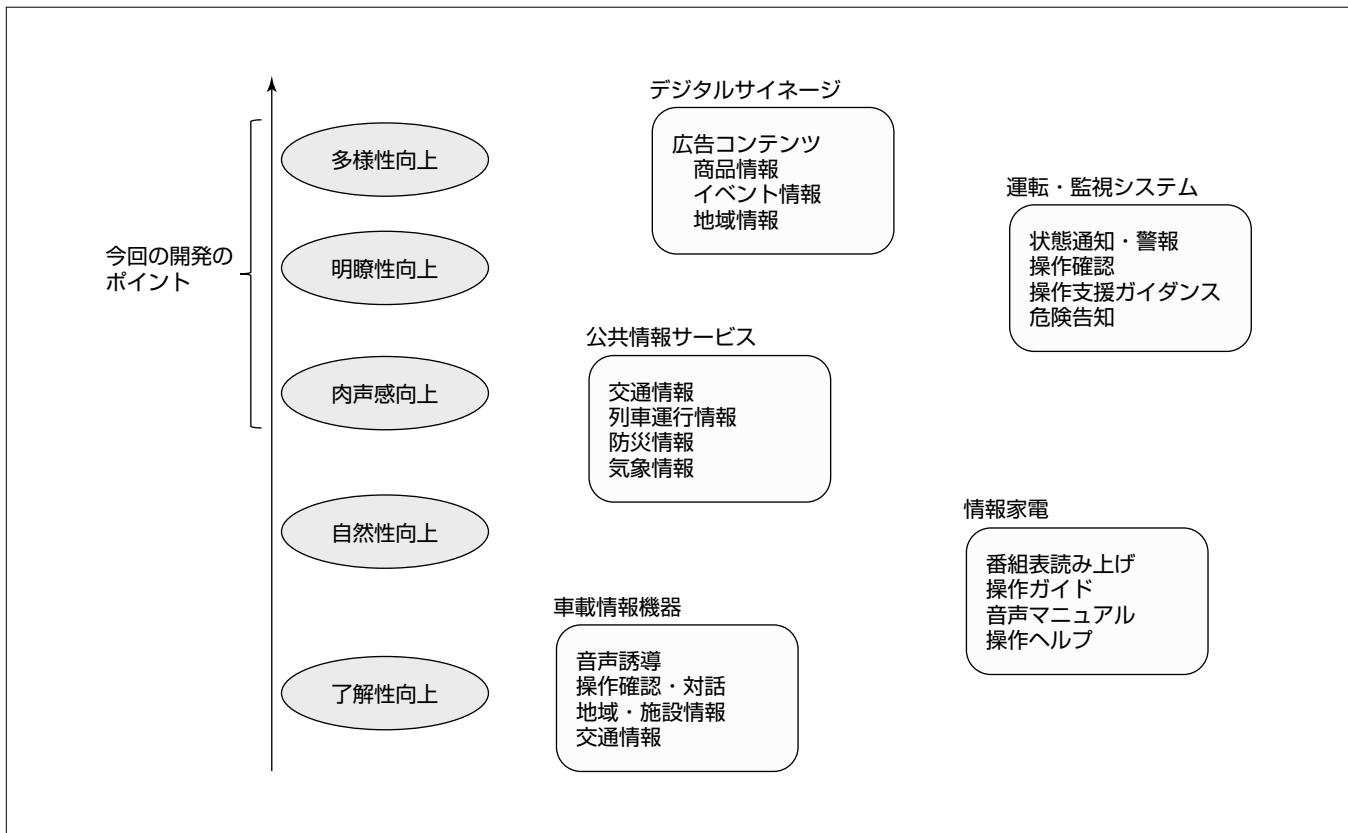
テキスト音声合成は、入力された任意のテキストから音声を自動生成する技術である。音声による情報提供は、伝達したい内容をタイムリーに分かりやすく伝えられるばかりではなく、注意を喚起できるため、作業中の人们にも、さらには同時に多くの人们にも、確実に情報を伝えることができる。三菱電機では、テキスト音声合成技術を幅広い分野で活用するため、合成音声の高音質化を進めるとともに、様々な用途での要求に応えるため、合成音声の機能性(明瞭化、多様化)の向上に取り組んでいる。

開発したテキスト音声合成技術は、人们が発声した大規模な音声データベース(音声コーパス)を用いたイントネーション・リズム制御と、音片選択とその接続による波形生

成によって合成音声を得る方式で、次の特長を備える。

- (1) 公共交通機関における放送音声で求められる人間らしい肉声感の高い音声(高音質)
- (2) 子音特徴を強調することで高齢者でも聞きやすい音声(明瞭性)
- (3) 注意喚起の際に聞き逃しを防止する緊張感のある音声など、用途に応じて特徴を制御できる音声(多様性)

このテキスト音声合成技術は、カーナビゲーションシステム、AV(Audio Visual)家電、公共放送を始めとし、様々な分野・製品に適用が進んでおり、マンマシンインターフェースの不可欠な要素技術となっている。



テキスト音声合成技術の進展と当社製品への適用

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生成する技術であり、音声を収録する必要がないのでコストがかからず、情報をタイムリーに分かりやすく、作業中の人们や多数の人们にも提供できるシステムを実現できる。高機能化する家電の使い方のガイダンスや、運転中の人们への情報提供などに適用されており、高音質化、明瞭化、多様化を進めることで、デジタルサイネージ、運転・監視システム等への適用が期待されている。

1. まえがき

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生成する技術である。音声による情報提供は、伝達したい内容をタイムリーに分かりやすく、確実に、多数の人にも提供できるため、高機能化する機器・システムの操作方法、点検業務のガイダンスや、作業者への支援、公共施設における放送等で要求が高い。これらの領域では人間らしい肉声感の高い高音質な合成音声が求められるとともに、高齢者でも聞き取りやすい合成音声、注意喚起するなど状況に応じた音声の特徴が制御可能な合成音声も求められる。

本稿では、これらの要求に応えるために開発したテキスト音声合成方式での、高音質化(2.2節)、明瞭化(2.3節)、多様化(2.4節)を実現する技術について述べる。

2. テキスト音声合成方式

2.1 概要

図1に、テキスト音声合成方式の構成を示す。テキスト音声合成方式は、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、言語辞書を参照して、テキストの解析(読み・アクセント解析)を行い、テキストに対応した読み(音素)、アクセント位置、品詞等を決定する。これらの音声合成のために必要な情報は中間言語と呼ばれ、アプリケーションによっては、この中間言語を直接専門家が記述・編集し、合成音声を得ることも可能である。

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に対応する音声の基本周波数(声の高低)と、発話速度とそのリズムに相当する音素の継続長のパターンを生成する。それらの情報をもとに、音片選択・波形生成では、人間が発声した音声から切り出された子音や母音に対応する短い波形である音片を、波形辞書から選択し、接続することで、音声波形を生成、出力する。

2.2 ナレーター音声に迫る高音質な合成音声

従来のテキスト音声合成方式⁽¹⁾⁽²⁾では、小規模な波形辞書から選択した音片を接続し、目的の基本周波数、継続長

の音声を得るために、音片に信号処理による変形を行い、基本周波数、継続長を変化させていた。この変形によって、得られる音声波形は品質が劣化するため、高い音質が求められる公共施設における放送音声などへの適用は不可能であった。また、高級車のカーナビゲーションシステムにおける運転者への経路誘導メッセージでも、高級車のイメージを損なわないように、合成音声ではなく、ナレーターによる録音音声が好まれていた。しかし、詳細な情報の提供のためには、多数の音声の録音が必要であり、コストがかかるばかりではなく、最近のクラウドサービスなど、あらかじめ録音ができないようなサービス・コンテンツでは、音声での情報提供ができないという問題があった。

近年、音声コーパス(音声データを多数集めた音声データベース。例えば音声時間10時間以上)を用意し、大量の音片を蓄積しておき、波形の変形を行わないことで、肉声感の高い音声を合成できるテキスト音声合成方式が研究されている⁽³⁾。

このテキスト音声合成方式は、音声コーパスの中から、入力されたテキストに基づき、適切な音片を選択し、信号処理を行わずに接続し、合成音声を得る。この方式では、大量音片からどのようにして、適切な音片を選択するかが重要な課題となる。一般には、次の尺度で音片を選択する。

- (1) 接続する前後の音片との連続性が高い音片(自然つながり)
- (2) 音片の継続長が、接続時に目的の継続時間のパターンに近くなる音片(自然なリズム)
- (3) 音片の基本周波数が、接続時に目的の周波数パターンに近くなる音片(自然なイントネーション)

従来、これらの尺度に対して、コスト関数、すなわち目標値からの差異を評価する関数を経験的に定義し、コスト最小となる音片を選択していた。このコスト関数の重要な部分である目標値の算出では、リズムやイントネーションに影響を与えると経験上考えられる限られた言語要因(入力されたテキストから得られる情報の種類・組合せ)から目標値を推定する学習が行われていたが、十分な要因を考

慮することができず品質劣化の原因となっていた。

そこで、我々は、与えられたテキストの様々な言語要因を条件として、音片系列全体の音響的特徴量の確率分布を音声コーパスから学習する方法を考案した(図2)⁽⁴⁾。この方法は、従来の経験に基づく、限定された要因による局所的な評価尺度の最適化ではなく、種々の言語要因の種類・組合せを同時に考慮し、音声の種々の特徴を確率的に学習し、最も高い確率で観測される音声特徴を再現する音片を

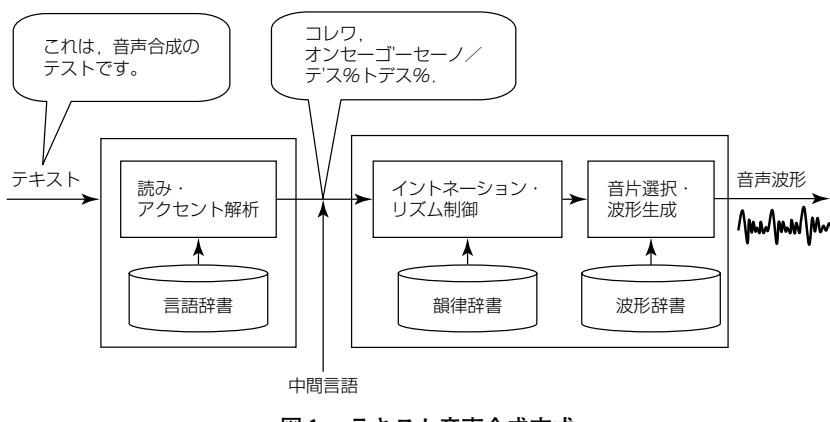


図1. テキスト音声合成方式

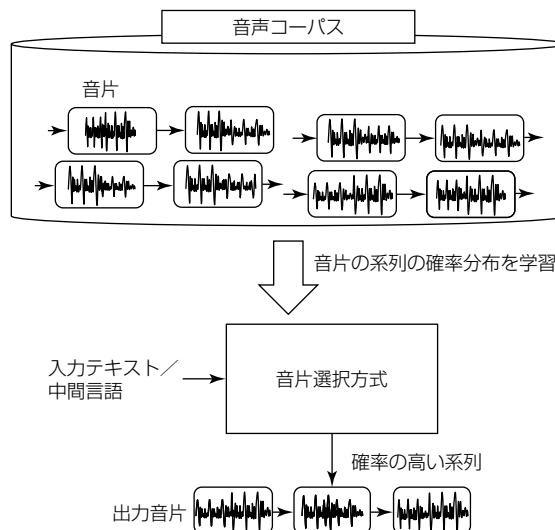


図2. 学習型の音片選択方式

選択する。この方法は、人間の発声で確率高く現れる現象が、品質の高い人間らしい肉声感のある声だという考えに基づく方法と言える。

この音片選択方式では、音片を選択するときに考慮する言語要因は、音片の系列の確率分布を特徴付ける素性関数によって表す。素性関数は、共起条件を満たす場合に反応する関数(共起条件を満たす場合に1, 満たさない場合に0を与える)であり、この共起条件を用いることで、音片の音声特徴と、それらに関連する言語要因とを考慮した確率分布を学習することが容易となる。ここで、音声特徴は、音声分析によって得られる情報で、例えば、韻律(基本周波数・継続長)、スペクトル振幅等である。また、言語要因は、入力テキスト又は中間言語から得られる言語情報で、例えば、音素の種類や、音高(声の高い、低い)、音片の文内・単語内の位置等である。

例えば、接続する音片の基本周波数の差に着目した場合、それに関連する言語要因である先行音片と後続音片の音素の種類、音片の位置の3種を考慮する素性関数 ϕ_i ($i=1, \dots, N$)は次のように設定する。

条件*i*：先行音片と後続音片の基本周波数の差 = I

&先行音片の音素 = P1

&後続音片の音素 = P2

& 音高 = T

ここでI, P1, P2, Tは、接続される可能性のある音片の組合せから得られる値である。すなわちNは、I, P1, P2, Tの組合せの数である。このように定義した素性関数によって、種々の言語要因の種類・組合せを考慮した学習を行う。

開発方式による効果を確認するため、主観評価試験を行った。評定者8名で、提示された14音声に対して、非常に悪い(-3)～非常に良い(3)の7段階で評価した。図3に

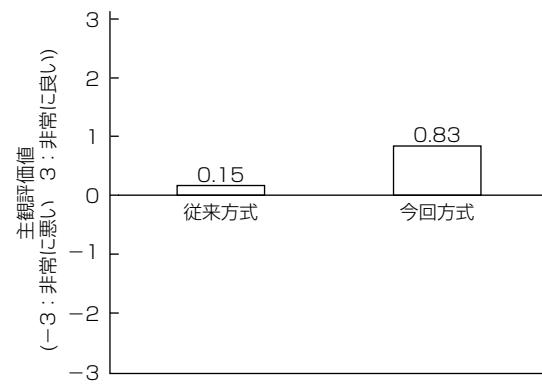


図 3. 主観評価の結果

示すように、今回開発した方式は、従来方式⁽²⁾と比べ、大幅な改善(0.68の向上)が得られ、高い肉声感が得られることが確認できた。

2.3 音声強調による合成音声の聞きやすさの改善

一般に高齢者は加齢による聴覚器官の衰えによって聴力が低下し、特に高域(高い周波数帯域の音声)が聞き取り難(にく)くなる傾向がある。高域の明瞭性を改善する方法として、イコライザによって高域を強調し、明瞭性を高めることが考えられるが、DA(デジタル・アナログ)変換器などの再生機器の制限で音声再生周波数帯域を狭くせざるを得ないシステムでは、既に合成音声を生成する音片から高域の情報が欠落しているため、イコライザによる方法は明瞭性の改善にはつながらない。そこで、音声の再生周波数帯域より高域の信号を、あらかじめ音片辞書の再生周波数帯域内に重畠しておくことで、音声合成のためのメモリ量や処理量に負担をかけず、疑似的な高域特徴を再現し、合成音声の明瞭性を改善する音声強調方式を開発した⁽⁵⁾。

図4に開発方式の処理の概要を示す。

- ①入力音声をスペクトルに変換し、音声再生周波数帯域外の高域スペクトルを切り出す。
 - ②入力音声の特徴を分析し、音声再生周波数帯域内で、高域を重畠するための適切な周波数境界（重畠帯域周波数Fc）を決定する。
 - ③重畠帯域周波数Fcに基づいて、再生周波数帯域内の上限周波数帯域内に収まるように切り出した高域スペクトルを周波数方向に線形圧縮する。
 - ④得られた圧縮スペクトルを、再生周波数帯域内に重畠する。

これは、本来の高域特徴に応じて、帯域内の高域を適応的に強調していることになり、制限された帯域内だけで疑似的な高域感が得られる。この方式によって、予音の明瞭性が向上することを確認した。

2.4 韻律変形による注意喚起可能な合成音声

公共施設における放送音声などは、聞き手が必要な時にだけ耳を傾けるため、災害時の避難誘導案内などの緊急放

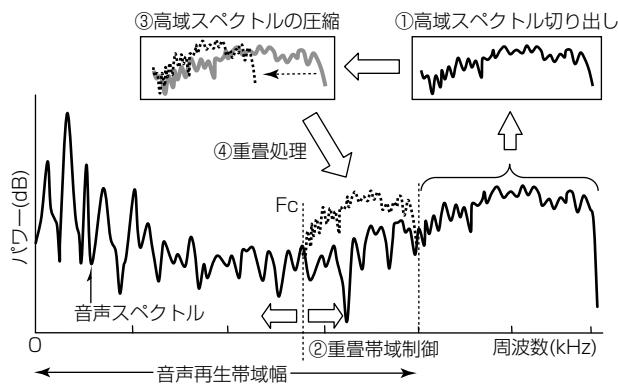


図4. 音声強調方式の処理

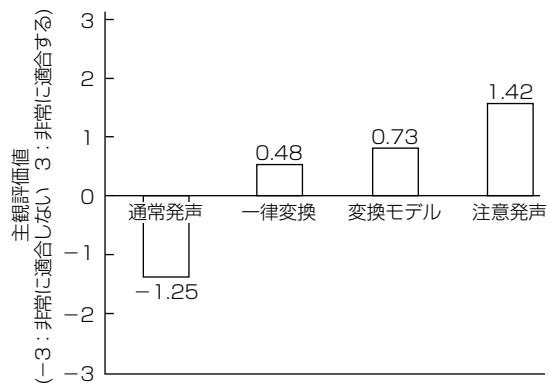


図5. 主観評価の結果

送を通常のアナウンスと同様の音声で放送すると、聞き逃すおそれがある。これを防止するには、発話の特徴を変化させ、緊張感のある注意喚起可能な音声を合成する必要がある。

3名のナレーターの通常のアナウンス(通常発声)と注意喚起を行うアナウンス(注意発声)を収録・分析したところ、注意発声は、通常発声と比べて継続長が短く、基本周波数が高いこと、またその特徴を再現すると、注意喚起の効果が知覚できることを確認した。しかしこの特徴を持つ合成音声を生成するために、注意発声の音声コーパスを作成することは、大量の音声データの録音のコストからも、さらには合成音声が収録した声の特徴再現にしかならないため、多様性の制御の観点からも現実的ではない。

そこで、通常発声の韻律と注意発声の韻律の間の相対的な変化を予測し、通常発声(従来方式)の韻律を変換することで、自然性を保つつ、注意喚起可能な韻律を再現する方式を開発した。特に句頭と句末の部分で変化が大きいことに着目し2段階の変換モデルを構築した。

- (1) 発話全体の継続長・基本周波数の平均値を変換
- (2) 句頭と句末の継続長・基本周波数を局所的に変換

評定者6名で、提示された音声に対して、緊急時の注意喚起の放送音声としてのふさわしいかどうかを“適合度合い”的7段階で判定する主観評価試験を行った。評価音声は、緊急放送文を含む10文を用いた。適合度は、3:非常に適合する、2:適合する、1:やや適合する、0:どちらでもない、-1:やや適合しない、-2:適合しない、-3:非常に適合しない、の7段階である。

次の4方式の音声を比較した。

- (1) 韵律変換を行わない合成音声(通常発声)
- (2) 継続長・基本周波数の平均値だけを変換(一律変換)
- (3) 平均値に加え句頭と句末を変換(変換モデル)
- (4) 韵律の変換モデルの学習に使用したナレーターの発声(注意発声)

図5に示す結果から、通常発声では適合しないという判

断であったのに対し、一律変換だけでも1点以上大きな改善効果があった。さらに変換モデルによって0.2点以上の改善があり、適合性を向上させることができた。開発した変換モデルによって、合成音声によってユーザーに注意を促すアナウンスが可能となった。

3. む す び

当社で開発したテキスト音声合成方式の最近の取り組みとその成果について述べた。開発したテキスト音声合成方式は、様々な製品への展開が可能である。

合成音声による情報伝達は、ここで述べた“注意喚起”にとどまらず、発話スタイルの制御や、強調等のより詳細な制御ができるようになれば、極めて豊かな情報を提供できる。ユーザーも、なんら学習する必要なく、人と人とのコミュニケーションで得た経験から、情報の意図や、種類、重要性を容易に理解できる。様々な分野で、効果的、かつ豊かなユーザーインターフェースが実現できるよう、今後も品質の更なる改良に加え、多様性の向上を図る予定である。

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) 大塚貴弘, ほか: テキスト音声合成技術, 三菱電機技報, **85**, No.11, 641~644 (2011)
- (3) N. Campbell, ほか: CHATR: 自然音声波形接続型任意音声合成システム, 電子情報通信学会技術研究報告SP, **96**, No.39, 45~52 (1996)
- (4) 大塚貴弘, ほか: 条件付き確率場に基づく波形接続型音声合成, 2014春季-日本音響学会講演論文集, 1-R5-10 (2014)
- (5) 古田 調, ほか: 入力信号の帯域外の高域成分を低域重畠する音声強調, 電子情報通信学会ソサイエティ大会講演論文集, A-4-19 (2013)

音のデザイン

木村 勝*
堀田 厚*

Designs of Sound Quality

Masaru Kimura, Atsushi Hotta

要 旨

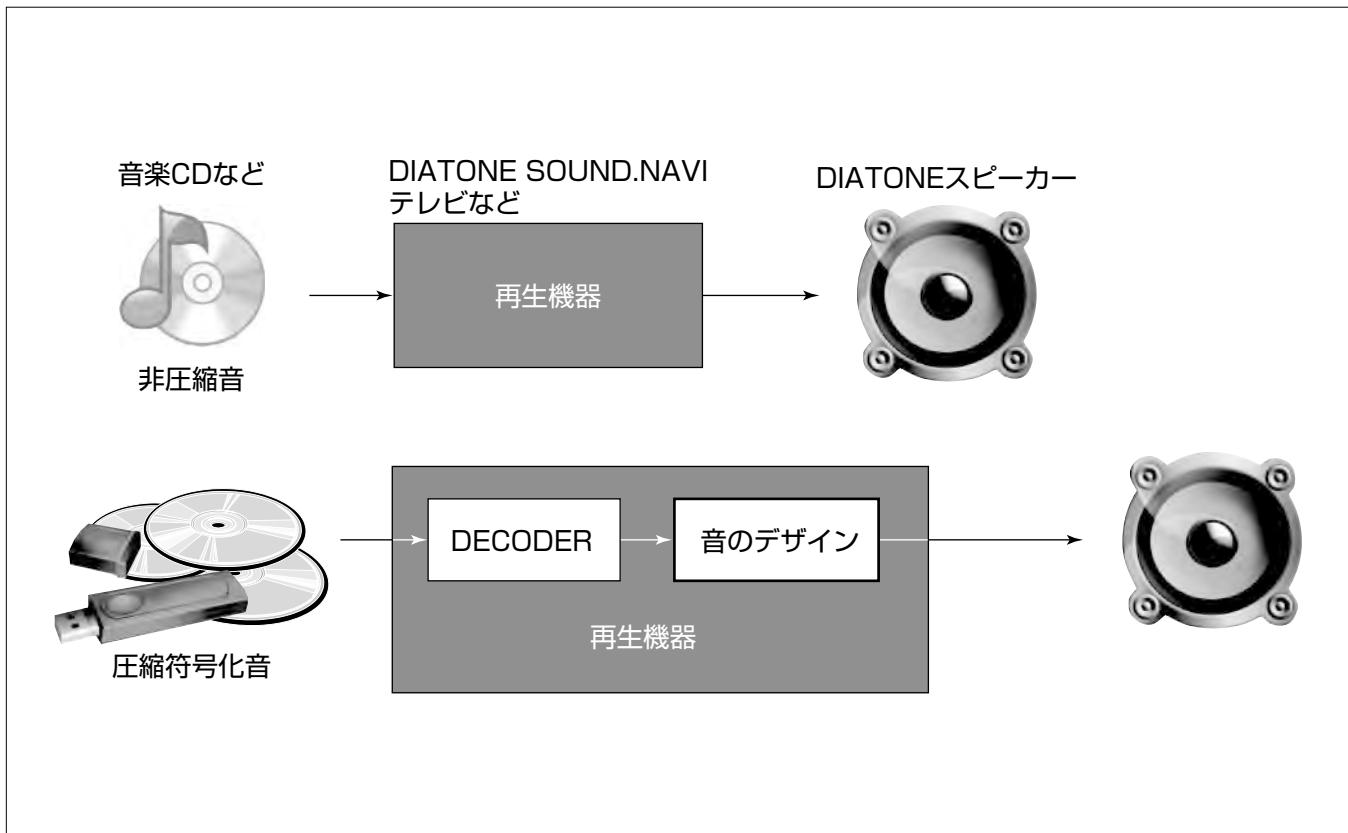
音楽の圧縮符号化技術は、地上デジタル放送や音楽配信サービス、DVD(Digital Versatile Disk)のようなパッケージ等、広範囲に使用されるようになってきた。

三菱電機では、“DIATONE”スピーカーに代表されるように、再生音を忠実に再現する技術について研究開発を続けてきた。しかし、圧縮符号化された音は、CD(Compact Disk)などの非圧縮音と比較すると一般に音質が劣化しているため、高忠実度再生技術だけでは高音質を実現することが困難になってきている。このため、音質・特性を積極的に変更する“音のデザイン”についても研究開発を行っている。本稿では、当社の音のデザインを支えている主要技

術の一つである圧縮符号化音の高音質化技術について技術的な内容と特長について述べる。

圧縮符号化音の高音質化技術は、圧縮符号化の音質劣化傾向に対応した2つの処理から構成されている。1つ目はステレオ回復処理であり、ステレオ左右信号の相関を利用し、豊かで立体感のあるステレオ音像を再現する。2つ目はBWE(Band Width Extension)であり、符号化で失われた高域成分を推定・再現することで、歯切れの良い素直な音を実現する。

主観評価の結果、圧縮符号化音を大幅に上回り、CD原音同等、又はそれ以上の音質が再現できることを確認した。



オーディオ技術のイメージ

当社が取り組んでいるオーディオ関連技術のイメージを示している。音楽CDなどの再現では、忠実な再現が必要であり、DIATONEスピーカーに代表される忠実再生技術に取り組んでいる。一方、音質劣化がある圧縮符号化音に対しては、音質・特性を積極的に加工し、あるべき姿にデザインすることで、CD以上の高音質化、又はユーザーの嗜好(しこう)にあった音の再生の実現を目指している。

1. まえがき

音楽の圧縮符号化技術は、地上デジタル放送や音楽配信サービス、DVDのようなパッケージ等、広範囲に使用されるようになってきた。当社では、再生音を忠実に再現する技術だけでなく、音質が劣化する圧縮符号化音に対しては、再生音の音質・特性を積極的に加工し、あるべき音にデザインする技術について研究を行っている⁽¹⁾。

本稿では、音のデザインを支えている主要な技術の1つとして、圧縮符号化音の高音質化技術について述べる。

2. 圧縮符号化の特性

2.1 音質劣化の傾向

圧縮符号化は、聴覚特性を利用した圧縮技術によって、大幅な情報圧縮を行っても、大きな音質劣化がないことが知られている。しかし、劣化が全くないわけではなく、注意深く聞くと原音とは異なる音質であったり、ステレオ音像の広がりが減少したように聞こえたりする等の音質劣化が知覚されることが多い。圧縮符号化を行ったオーディオ信号には、次の特徴が現れる。

- (1) 高域成分の欠落
- (2) 左右差信号の欠落

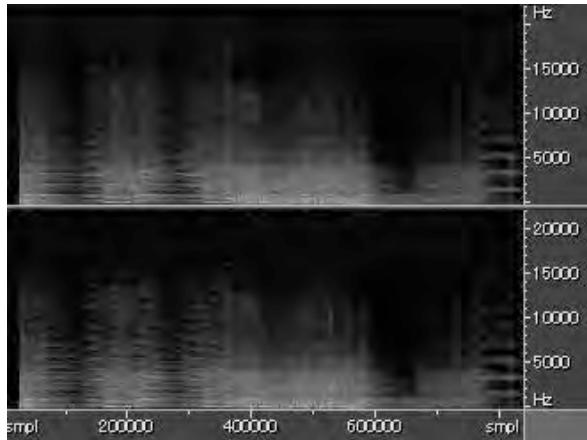


図1. 圧縮符号化のスペクトログラム(MP3)

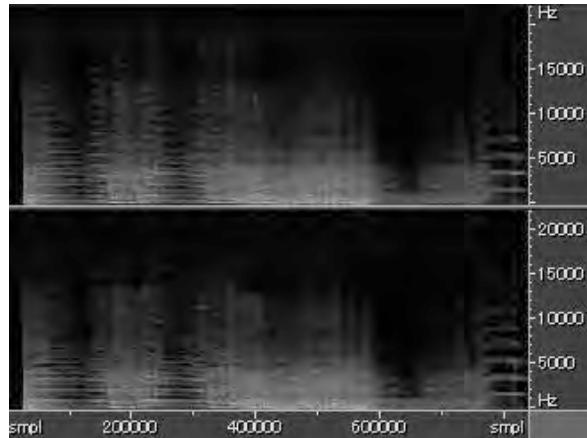


図2. 圧縮符号化のスペクトログラム(AAC)

これらの特徴が音質劣化の要因となり得ることを述べる。

2.2 高域成分の欠落

図1、図2は、圧縮符号化方式MP3(MPEG(Moving Picture Experts Group) AUDIO Layer 3, 本稿ではLAME Ver.3.98.2で、128kbpsでエンコードしたものを使用), AAC(Advanced Audio Codec, 本稿ではiTunes Ver.8.0.2.20で、128kbpsでエンコードしたものを使用)で符号化した信号のスペクトログラム(時間・周波数特性)を示している。図で、上段が圧縮符号化前(原音)の特性、下段が圧縮符号化後の特性である。

図から、MP3、AACとともに圧縮符号化によって、およそ16kHz以上の高域成分が欠落していることが分かる。ビットレートが下がるとこの傾向は更に顕著になり、64kbpsではおよそ12kHz以上の全ての高域成分が失われる場合もある。

高域成分は、音の華やかさや立ち上がりの鋭さ、楽器の生々しさ等に関係するため、高域成分が欠落することは、音質劣化の要因となり得る。

2.3 左右差信号の欠落

再生するオーディオ信号に空間的な情報が多く含まれるほど、ステレオ音像の広がり感や奥行き感が得られる。また、空間的な情報は左右信号の差が大きいほど豊かになる。

図3、図4はMP3、AACで符号化したオーディオ信号

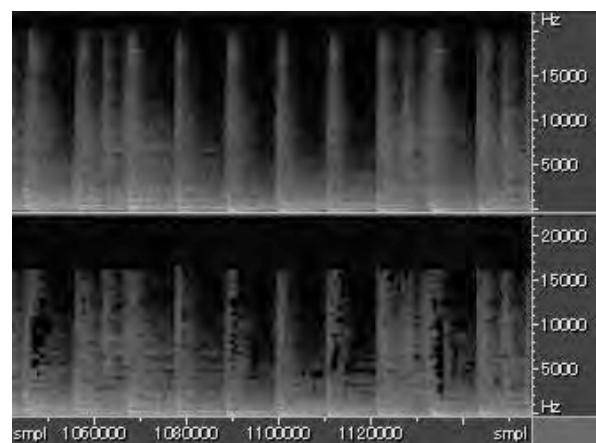


図3. 左右差信号の特性(MP3)

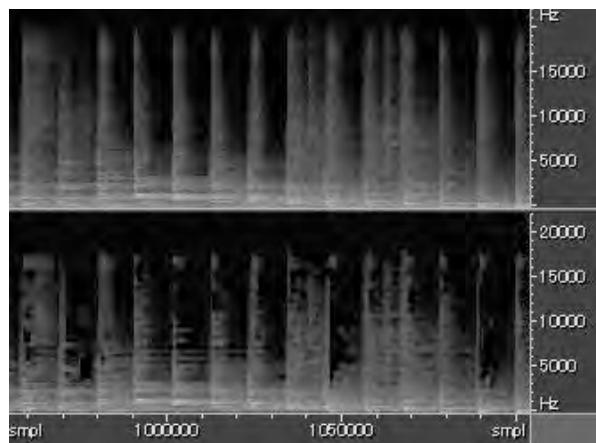


図4. 左右差信号の特性(AAC)

の左右差信号のスペクトログラムである。図で、上段が圧縮符号化前(原音)の特性、下段が圧縮符号化後の特性である。

図から、左信号と右信号の差分の中高域のスペクトルが部分的に欠落し、いわゆる“歯抜け”的な状態になっていることが分かる。スペクトル上の欠落は、その帯域の音信号が左右同じ音になっていることを示すものであり、ステレオ音像の広がり感、奥行き感が劣化する要因になり得る。

ビットレートが下がるとこの傾向は更に顕著になり、64kbpsではかなり広範囲の周波数成分が失われることもある。

3. 高音質化技術

3.1 処理構成

図5は圧縮符号化音の高音質化技術の処理構成である。この技術は、圧縮符号化による2つの音質劣化要因に対応し、失われた左右差信号成分を推定補完するステレオ回復器と、失われた高域成分を推定して帯域を拡張するBWEから構成される。

3.2 ステレオ回復器

ステレオ回復器では、左信号と右信号の無相関成分を抽出し、これを左右差信号の推定信号として代用する。これによって、左右差信号を強調し、ステレオ音像の広がり感を回復している。

一般的に、ステレオ音声の左信号と右信号のスペクトル強度がほぼ等しいとき、左右信号の和信号は相関成分、差信号が無相関成分を表すと解釈できる。この方法は、その特徴を利用し、無相関成分を算出し、差信号を再現し強調する方法である。

3.3 BWE

BWEでは、入力信号を、調波構造を持つ周期成分と、立ち上がり成分などの非周期成分とに分割し、周期成分には調波構造を保持するように帯域拡張を行い、非周期成分には立ち上がりを強調する処理を行っている。

一般的に、高域成分が欠落すると、人の声や楽器等の周期性のある音では高調波成分がなくなることで、華やかさ、伸びやかさが欠如し、打楽器などの非周期的な音では、立ち上がりが鈍くなる。これに着目し、周期的な音と非周期的な音で異なる高域の再現を行い、品質改善を図る方法である。

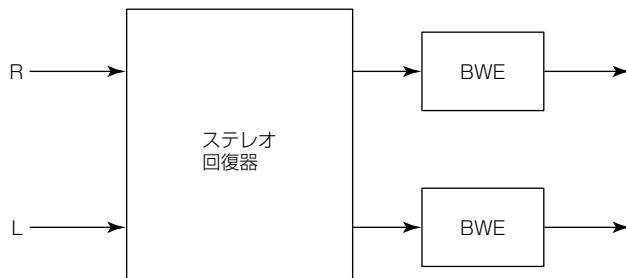


図5. 圧縮符号化音の高音質化処理

4. 評価

高音質化技術の効果を評価するため、音楽信号を用いて周波数特性の確認と、主観評価を行った。

4.1 処理後の特性

図6、図7は、MP3、AACで符号化したオーディオ信号(上段)と、高音質化技術で処理した信号(下段)のスペクトログラムを示している。なお、評価曲として欧州で標準的に音質評価に用いられているSQAM(Sound Quality Assessment Material)のtrack 59(Violin)を用いた。

図から、MP3、AACどちらの方式でも処理前は16kHz以上の高域成分がないが、この処理によって20kHz程度まで高域成分が伸びていることが確認できる。これらの図は一例であるが、SQAMに収録された70曲で同じ傾向となることを確認している。

また、図8、図9はMP3、AACで符号化したオーディオ信号の左右差信号(上段)と、高音質化技術で処理した信号の左右差信号(下段)のスペクトログラムを示している。なお、評価曲としてSQAMのtrack 69(ABBA)を用いた。

図から、この処理によって中高域成分の“歯抜け”が改善

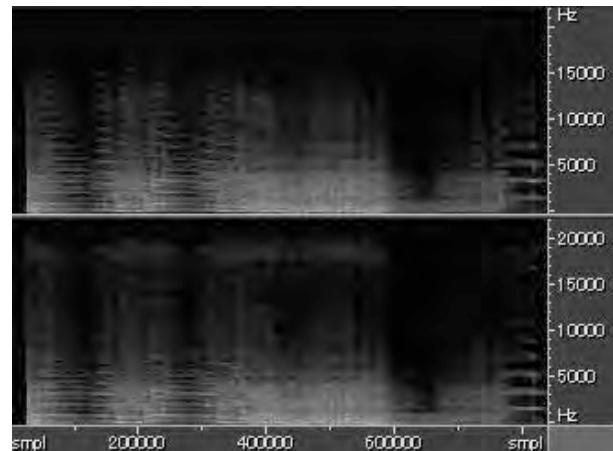


図6. 圧縮符号化の高音質化処理(MP3)

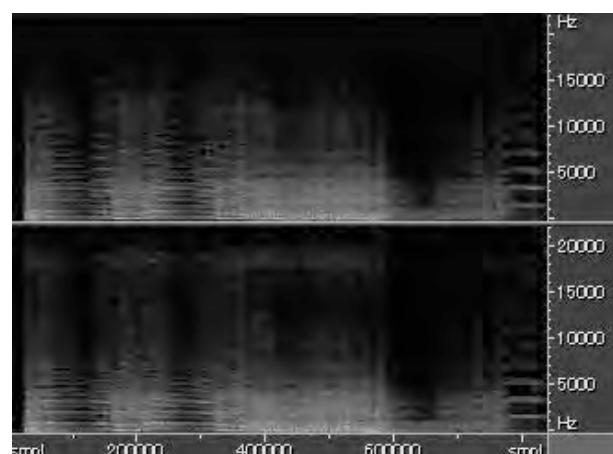


図7. 圧縮符号化の高音質化処理(AAC)

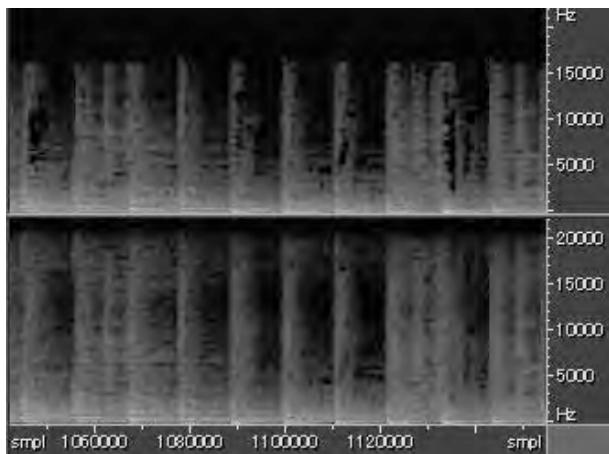


図8. 圧縮符号化の高音質化処理(MP3)

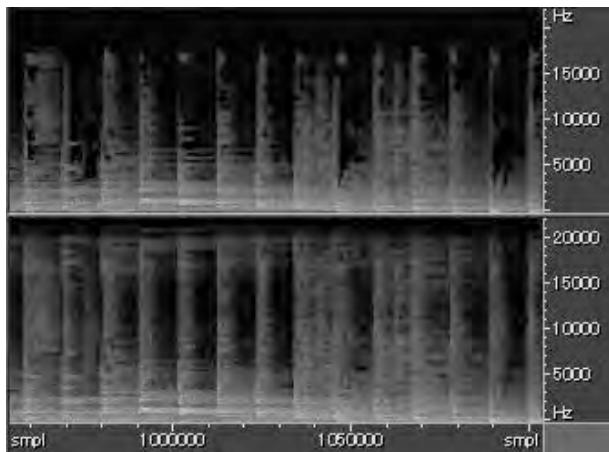


図9. 圧縮符号化の高音質化処理(AAC)

されていることが確認できる。なお、SQAM収録曲70曲で同様の傾向となることを確認している。

4.2 主観評価

オーディオ信号の主観評価方法を規定したITU-R (International Telecommunications Union-Radiocommunication sector)勧告BS.1534⁽²⁾で用いられている“隠れ基準及び隠れアンカー付き多重刺激二重盲検法”を参考に主観評価を行った。すなわち、被験者は、評価音が基準音よりも良ければ50点以上の評点を、悪ければ50点以下の評点を0~100の範囲で採点し、その際、安定した評価ができるよう、明らかに品質の低いサンプルを含めて評価する。このサンプルを隠れアンカーと呼ぶ。評価では、基準音と同程度であれば50点と採点する。隠れ基準はCD音声、隠れアンカーは、6 kHz以上の高域成分を緩やかにカットした曲とした。

また、評価曲として、SQAMの中から(1)track 48 (Vocal), (2)track 59(Violin), (3)track 66(Orchestra), (4)track 70(Eddie Rabbitt)の4曲を選定した。

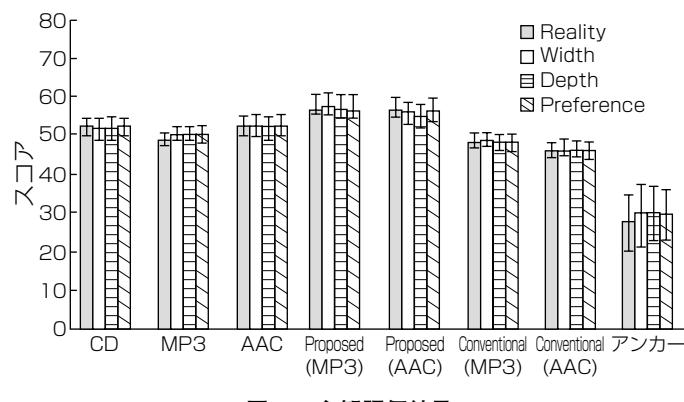


図10. 主観評価結果

評価項目は、“楽器・声のリアルさ(Reality)”“音像の広がり(Width)”“音像の奥行き(Depth)”“好き／嫌い(Preference)”の4項目である。被験者はオーディオエンジニア7名、また比較のため、従来技術による高域再現手法⁽³⁾を含めた8種類の比較を行った。

評価結果を図10に記す。図では、隠れ基準音、アンカーを含めた各評価音のスコアの平均点を示している。また、t検定に基づく95%信頼性区間も示している。

図10から、高音質化技術(図中Proposed)によって元の圧縮符号化音に比べ、8ポイント近くの大きな改善効果があり、従来方法(図中Conventional)からも優位に改善されていることが確認できた。

さらに、提案方式は、全て、CD原音と同等以上のスコアを得ており、被験者が全般的に良い評価をしたと言える。

5. むすび

当社の音のデザイン技術を支える圧縮符号化音の高音質化方式について述べた。この方式は、圧縮符号化によって生じる2つの音質劣化傾向に対応してステレオ回復器とBWEとから構成されており、主観評価によって有効性が確認できた。

今後、圧縮音声に限定せず、より高い品質の提供、ユーザーの嗜好や曲にあわせた音質の制御等、オーディオ全般について音のデザインの研究開発を継続していく。

参考文献

- (1) Kimura, M., et al.: Improvement in Stereophonic Sound Images of Lossy Compression Audio, IEEE 2nd GCCE, 90~91 (2013)
- (2) Recommendation ITU-R BS.1534 : Method for the subjective assessment of intermediate quality levels of coding systems (2003)
- (3) 佐藤 蜜:圧縮で失われた音の復元技術, 電気学会誌, 127, No.7, 411~412 (2007)

映像合成技術

Video Synthesis Techniques

Yohei Miki, Ichiro Furuki, Shin Miura

要 旨

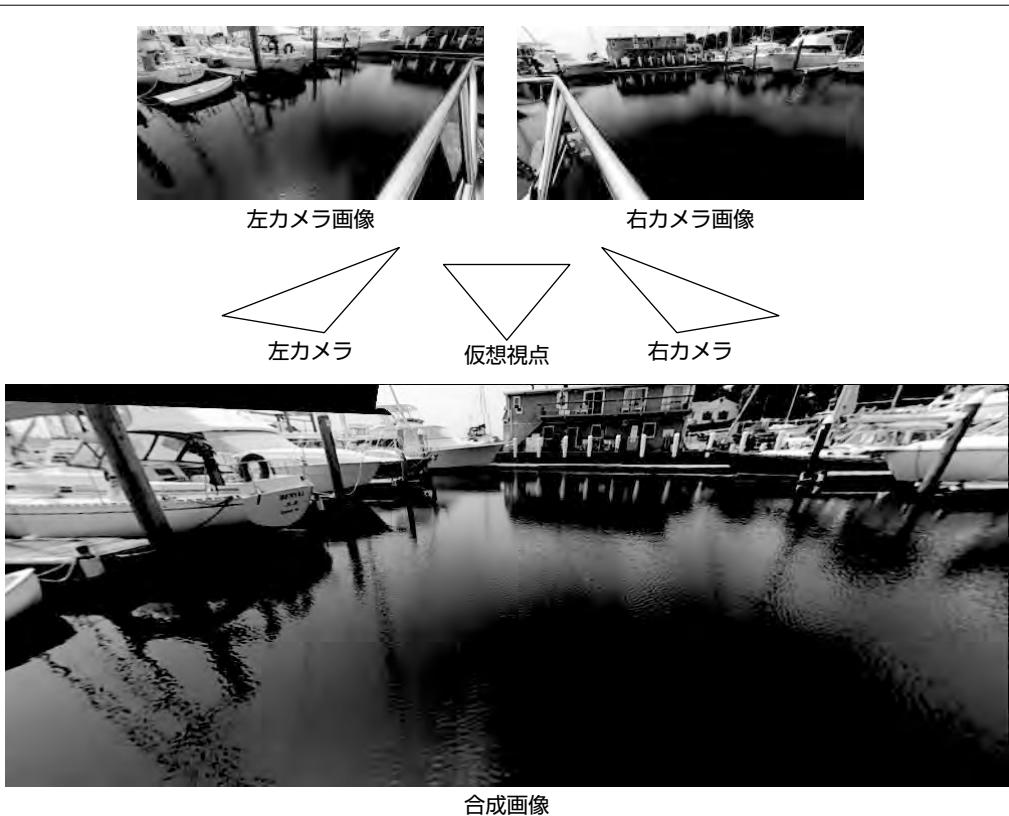
安全・安心に対する意識の高まりがあり、カメラ映像を用いてユーザーが状況を確認するシステムが増えている。カメラ映像の利用によって、ユーザーは直接見ることができない死角領域も確認できるようになる。

しかし、カメラの視野は限られているため、ユーザーが見たい範囲がカメラの視野より広い場合は、複数台のカメラが必要となる。複数台のカメラ映像をそのまま並べて表示すると、ユーザーは各カメラの設置位置とその関係を考慮して見る必要があり、状況の把握は必ずしも容易ではない。この問題を解決するためには、複数視点のカメラ映像から見やすい合成映像を生成する必要がある。複数カメラ映像の合成、仮想視点からの映像合成は、既に実用化も進

んでいるが、カメラの設置間隔が離れ、カメラー被写体間の距離が短い場合を含む任意の被写体を対象とすると、カメラ視差による影響によって、従来技術では、画像の重畳領域でズレが生じるという問題がある。

そこで、画像を合成する際、継ぎ目部分が自然な画像を作成するために、2台のカメラ視野が重畳する領域の被写体までの距離を算出し、算出した距離を利用して、仮想視点からの合成画像を生成する方法を開発した。

合成画像から得られる映像は、高い品質で、船舶などの死角が多く存在する環境でも、広い視野の映像を安定して提供することが可能である。種々の領域でユーザーの安全・安心を映像技術によって支援する。



映像合成技術

2台のカメラの映像を利用した映像合成技術適用例である。2視点のカメラ映像の各フレーム画像について、画像が重畳する領域の距離情報を求め、カメラ視差の影響を取り除いた広い視野角の合成画像を生成する。これを実時間処理することで、被写体によらず、死角のない見やすい合成映像を提供することができる。

1. まえがき

安全・安心に対する意識の高まりによって、カメラ映像を用いてユーザーが状況を確認するシステムの実用化が進んでいる。カメラ映像を利用することで、ユーザーは直接見ることができない広域、複数区域、死角領域を確認できる。さらには、状況を把握しやすい任意の視点からの映像の生成や、映像に情報を重ねて提供するシステム等、今後、映像情報の利用はますます広がると考えられる。

一般に、カメラの視野は限られているため、ユーザーが見たい範囲がカメラの視野より広い場合は、複数台のカメラが必要となる。複数台のカメラ映像をそのまま並べて表示した場合、ユーザーは各カメラの設置位置やその関係を考慮して見る必要があり、ユーザーにとって状況の把握は必ずしも容易ではない。この問題を解決するため、複数視点のカメラ映像から見やすい合成映像を生成する必要がある。

しかし、映像の合成では、各カメラの画像の継ぎ目部分にずれや消失が生じることがある。本稿では、被写体までの距離を算出し、それに基づく画像を生成することでこの問題を解決する方式を提案し、その有効性を示す。

2. 複数カメラの映像合成

2.1 目的

カメラ映像の利用は、多種多様な領域で実用化が進んでいる。代表事例に、自動車における後方映像利用による駐車支援がある。最近では、俯瞰(ふかん)映像の提供など、映像処理技術によって、より状況を理解しやすい映像の提供も実現している。しかし、自動車の俯瞰画像のようにカメラと被写体が固定の位置関係ではなく、任意の被写体を対象とした場合には、自然で連続性の高い映像合成は容易ではない。我々は、カメラ設置位置が離れ、被写体が、視差の影響のある近距離を含めた任意距離にある場合について、品質の高い実用的な合成映像を生成することを目的として映像合成方式を検討している。

例えば小型船舶などでは、もともと死角が多く、さらに波や船舶自体の角度の変化で著しく視界が悪くなることがあり、広い視野の明瞭な前方視界を提供する技術が強く要求されている。

2.2 課題

自動車の駐車支援システムで用いられている映像合成手法は、車両に広角カメラを複数設置し、カメラ映像の毎フレームの画像を、地表面に投影する変換と接続を行い、俯瞰画像を生成する手法が用いられる。地表面とカメラは、固定の位置関係にあるため、それぞれのカメラ画像と、地表面の画素の関係は、事前に設定することが可能である。したがって、合成された画像では、地表面付近については、

連続性のある明瞭な俯瞰画像になるが、他車や通行者などの高さのある被写体が入り込んだ場合には、それらの被写体は歪(ひずみ)が生じ、俯瞰した画像にはならない。

今回対象とする環境は、カメラの設置間隔が離れており、カメラー被写体間の距離が短い環境である。このような状況で、固定の投影面へのマッピング手法を適用すると、地表面以外の物体と同様、仮定した投影面までの距離が異なる場合、画像には歪が生じ、画像の重畳領域・継ぎ目では、被写体が重複して2重に表示される、又は消失する場合がある(図1)。

固定面へのマッピング以外の方法としては、複数画像を、別の被写体全体をカバーできるオーバービューカメラを用いて合成する方法がある⁽¹⁾。この場合、オーバービューカメラ画像を参照画像として、各カメラ画像を対応付けるため、事前にカメラ位置・角度に基づくマッピングを規定しておく必要がなく、厳密なカメラの位置・角度合わせは必要がない。この方法では、1台のオーバービューカメラでは得られない高解像度画像を、複数カメラを用いることで簡単に得ることができる。しかし、この手法は、画角が広

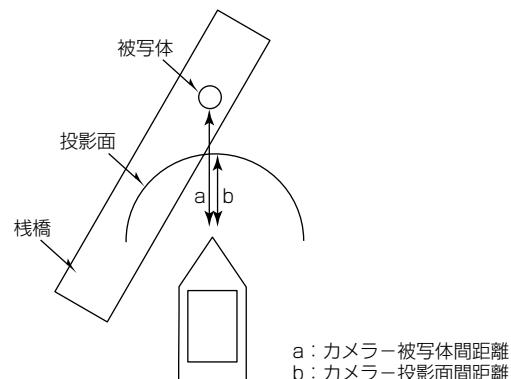


図1. 固定投影面で画像継ぎ目部分に被写体がある場合

いパノラマ画像を得ることが目的であり、被写体は十分に遠いか、各カメラの位置がそれほど離れていないことが画像合成の条件となる。

自動車、船舶などの運転支援のための映像提供を考えると、運転者・操縦者にとっての死角は、自車、自船自体が障害物となることで生じる。1か所に複数カメラを設置して画角を広げただけでは、死角の少ない映像情報を提供することはできない。また、被写体は、カメラから近いところにもあり得るため、異なる位置に設置したカメラの視差による映像のズレを解決することが必須である。

レンズや鏡を用いる光学系によって光学中心を一致させ、視差の影響のない映像を合成する取組みもあるものの、装置のコストがかかること、得られる映像が暗いものになる場合があること、設置条件が限定されること等問題が多い。そこで、これらの問題を、映像の合成時に被写体に対して適切な投影面を求める問題と定義して、視差による映像のズレ・消失を解決する。

3. 映像合成方式

課題を解決するために、2台のカメラ映像の毎フレームの画像について、重畠している領域の被写体の距離を算出し、求めた距離をもとに合成画像を生成する手法を提案する。

次に2台のカメラの場合について、画像合成手法の詳細を述べる。

3.1 距離情報算出

2台のカメラの相対位置・姿勢を用い、カメラの視野が重畠している領域の被写体までの距離を求める。この場合、複数の被写体が写り込んだ複雑な画像から、距離を簡単に求めることはできない。そこで、複数の距離仮説をたて、それぞれに対する合成画像を評価することで、距離を求める方法を提案する。

図2に示すように2台のカメラの間に仮想視点を定義し、仮想視点から見たときに正対する仮想平面を定義する。この仮想平面は仮

想視点から距離が異なる位置に複数面定義する。

はじめに、図3に例を示すように、各仮想平面に左右のカメラの画像を投影し、仮想視点からの画像を作成する。この作成画像について、左右のカメラ画像の同じ画素位置に対する輝度の一致度を評価する。すなわち、左右の画像の重複部分の輝度の正規化相互相関を計算する。最も相関が高い投影面がその被写体までの実際の距離に最も近いものと判断する。仮想視点から見た画像については、複数の被写体が含まれるため、この距離計算を全座標について行い、重複画像の距離情報とする。

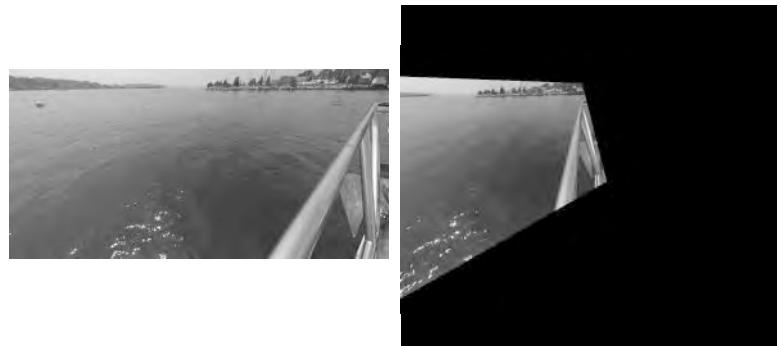
3.2 合成画像生成

距離情報の算出と同様、左右カメラの前方に仮想平面を複数定義する。重畠領域の各画素に対して、左右カメラの画像を算出された距離の仮想平面上にマッピングすることで画素値を算出し、これらを統合することで合成画像を作成する(図4)。

3.3 処理手順

提案する画像合成方式を適用して、合成映像を生成する処理手順を図5に示す。映像合成処理には、毎フレームの画像に対する前処理が必要であり、さらに、距離情報を算出するためには、事前にカメラの位置・姿勢情報を抽出しておく必要がある。

前処理では、レンズ歪の補正と不要領域の除去を行う。レンズ歪の補正是、被写体の距離が既知の場合に投影面に



(a) カメラ画像(左カメラ)
(b) 仮想視点からの画像

図3. 投影面を用いた仮想視点からの画像

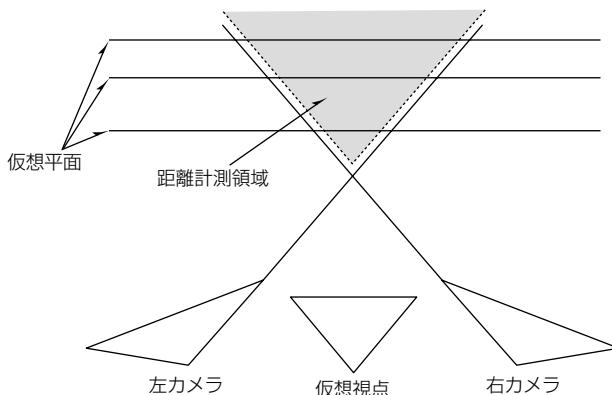


図2. 仮想平面への画像の投影

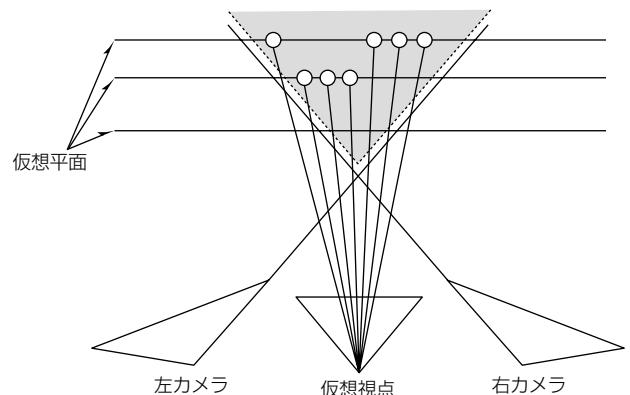


図4. 合成画像生成

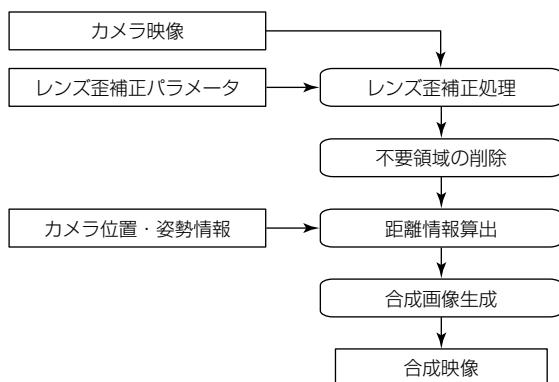


図 5. 映像合成処理



(a) カメラ画像



(b) マスク画像

図 6. 不要領域の削除

投影した 2 カメラの画像の画素が一致するよう、事前にレンズ歪を計測し、この補正を行うものである。

また、カメラ画像には、図 6(a)に示すように、自船の装備品などの、合成時には不要な領域が含まれることがある。このような領域は、カメラ設置時に抽出し、図 6(b)に示すようなマスク画像を用意し、画像合成に使用する領域から削除する。

距離計算並びに画像合成では、2 つのカメラの相対位置と、姿勢すなわちカメラの向きの情報が必要である。この情報抽出には、画像の中の複数の特徴点について、2 つのカメラ画像の投影画像上で誤差を求める、合計誤差が最小になるように、変換行列を繰り返し計算することで自動的に



図 7. 元画像



図 8. 合成結果

取得できる。

この処理を映像の各フレーム画像に対して行い、合成映像を得る。既にシステムを構築し、リアルタイム処理を実現している。

4. 適用例

開発システムを小型船舶に適用し、評価を行った。図 7 に設置した 2 カメラの元画像例を、図 8 に合成結果例を示す。図 7 からは、両画像に明らかな視差があることが分かるが、合成画像では、ずれや消失がなく、船舶自体の構造物が写り込まない画像が生成できている。小型船舶は、視界が狭く、死角が多いだけではなく、波や船舶の角度の変化などで前方視界が得られにくいことが多い。得られた映像を操縦者に評価してもらった結果、安全な操縦に有効性が高いとの評価を得た。

5. むすび

複数のカメラ画像から視差の影響を除去した合成映像を作成する技術を開発した。複数の距離仮説に対して、投影画像の相関を用い画素ごとに距離を算出し、その距離に基づいた投影面の画像を合成する方式によって、死角がない映像の提供が可能である。今後は、評価と改良を行い、様々な領域に適用する予定である。実カメラ映像よりもユーザーの理解、状況把握が容易な映像情報を提供し、安全安心な社会の実現に貢献する。

参考文献

- (1) 原田雅之, ほか: 大画面映像合成技術, 三菱電機技報, 82, No.12, 751~754 (2008)

最新映像符号化技術と スーパーハイビジョンエンコーダ

坂手寛治*
井對貴之*
本山信明*

New Video Coding Technology and Super High Vision Encoder

Hiroharu Sakate, Takayuki Itsui, Nobuaki Motoyama

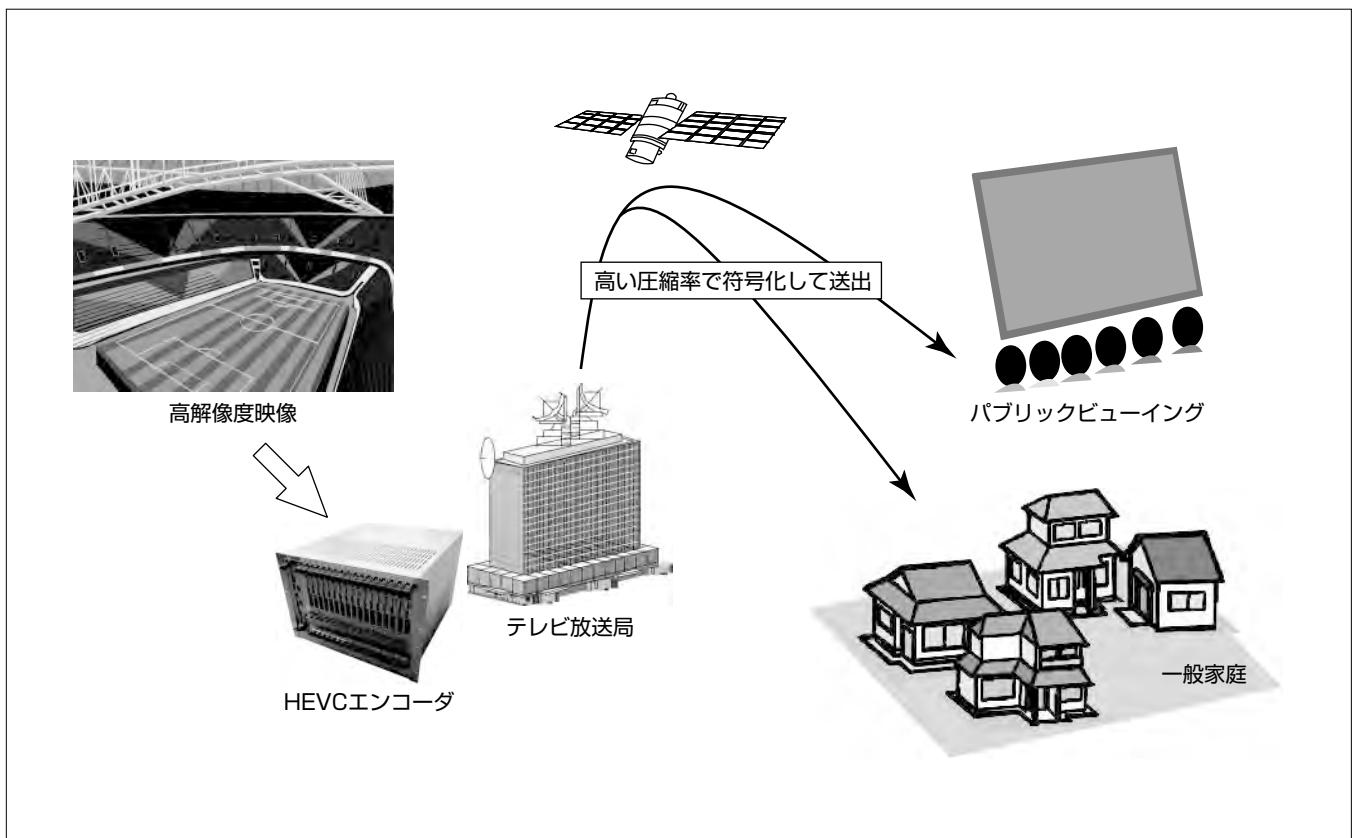
要 旨

映像符号化技術HEVC(High Efficiency Video Coding)が2013年に新たな国際標準として誕生した。HEVCは、従来方式H.264/AVC(Advanced Video Coding)の2倍の圧縮性能を持っており、次世代テレビ放送などの新たなサービスを実現するための基盤技術として期待されている。三菱電機は標準化作業に参画し、技術・運営の両面から貢献してきた。

HEVCの大きな特徴として、4K／8Kなどの高解像度映像への対応が挙げられる。当社は日本放送協会(NHK)と共同で、HEVCを用いたスーパーハイビジョンエンコーダを開発した。この装置は、8K解像度のスーパーハイビジョン映像をリアルタイムに符号化する能力を持つ。17枚の基板による並列処理構成を採用し、それぞれの基板で分割さ

れた画像の符号化処理を行う。このとき、基板間で画像データや符号化パラメータを共有可能なアーキテクチャすることで、画面分割に伴う符号化制約を少なくし、圧縮性能の低下を防いでいる。さらに、符号化アルゴリズムについては、HEVCの構成要素のうち、有用なものだけを選別することで軽量化を図った。これによって、符号化全体の演算量をH.264/AVCの2倍程度に抑えることに成功した。ハードウェア実装面でも、回路規模やメモリ転送帯域をデバイス間で平滑化する工夫や、データ転送負荷を基板ごとに分散させる工夫を行い、装置サイズを小さく抑えている。

この映像符号化技術とその実装技術によって、今後の様々な映像情報サービスの実現に貢献する。



HEVC符号化技術による次世代テレビ放送

従来方式の2倍の圧縮性能を持つHEVCは、既存システムの高性能・高品質化や新たなサービスを実現する基盤技術である。実用化に向けた準備が進められている次世代テレビ放送では、衛星回線など伝送帯域が限られた通信路で情報量が膨大な4K／8K解像度の映像を送出するためには必要不可欠な技術となる。

1. まえがき

映像符号化方式HEVCが2013年に新たな国際標準として誕生した⁽¹⁾。HEVCは、AV(Audio Visual)機器や監視システム等で広く用いられているH.264/AVC方式の2倍の圧縮性能を持っており、既存の映像システムの高性能・高品質化や、新たなサービスを創出する基盤技術として期待されている。当社は、HEVCの標準化作業に参画し、技術・運営の両面から貢献してきた。

HEVCの活用が期待される新しい映像サービスの一例が次世代テレビ放送である。2014年内にHDTV(High Definition TeleVision)の4倍の4K解像度での放送が始まり、さらに2020年には4倍の8K解像度での放送が計画されている。HEVCは、次世代放送サービスの実現に欠かせない主要技術の1つと考えられている。

本稿では、HEVC符号化技術の概要を述べるとともに、NHKと共同で開発したスーパーハイビジョンエンコーダについて述べる。

2. HEVC符号化技術の概要

2.1 HEVC規格の範囲

HEVCは、これまでMPEG(Moving Picture Experts Group)規格を策定してきたISO/IEC(International Electrotechnical Commission)とH.26x規格を策定してきたITU-T(International Telecommunication Union-Telecommunication standardization sector)の2つの機関で標準化されている。規格名は、ISO/IECではMPEG-H(ISO/IEC 23008)、ITU-TではH.265であるが、両者ともHigh Efficiency Video Codingが規格書のタイトルとなっている。

HEVC規格で定義されているのは、H.264/AVCなどの従来の方式と同様、ビデオストリームのデータ構造とその復号方法が主である。符号化のアルゴリズムや符号化/復号装置(コーデック)の実装方法については厳密な記述はなく、多くは開発者の裁量に委ねられている。例えば、低速デバイスで処理させるために低演算量のアルゴリズムを用いたり、リアルタイム性を実現するためにパイプライン処理による装置構成を採用したりするなど、開発者が用途や開発条件に合わせて実現方法を決めれば良い。

表1にHEVCの適用範囲を示す。4K/8K解像度が含ま

表1. HEVCの適用範囲

解像度	QVGA～4K, 8K
色密度	4:2:0, 4:2:2, 4:4:4 ※メインプロファイルは4:2:0だけ
ビット階調	8～14ビット/pixel ※メインプロファイルは8ビット、 メイン10プロファイルは8～10ビット
フレームレート	静止画～60frame/s(以上)

QVGA: Quater Video Graphics Array

れていることが大きな特徴である。H.264/AVCなどの従来の方式は、主としてSDTV(Standard Definition TeleVision)サイズ以下の映像を用いて評価を行いながら規格化されたものであるため、高解像度映像に対して良好な性能が得られる保証はなかった。それに対して、HEVCは4K/8K解像度が方式検討段階から含まれており、高解像度映像でも最適な性能が得られるようになっている。

色密度やビット階調等も将来の高精細映像サービスを見据えた仕様となっている。現在定義されているメインプロファイルやメイン10プロファイルでは、色密度は4:2:0だけ、階調は10ビットまであるが、HEVC規格自体は、高色密度・高ビット階調にも対応できる形式となっている。静止画が対象として含まれているが、HEVCは静止画用符号化方式として高い性能を誇るJPEG(Joint Photographic Experts Group)2000よりも圧縮率が高いとの実験結果が報告されている⁽²⁾。デジタルスチルカメラや医療システム等の用途でも活用が期待される。

2.2 拡張規格と関連技術

2013年のメインプロファイルの規格化完了を受け、HEVCの拡張規格の標準化が進められ、2014年4月に、放送設備などを主用途とする高色密度・高ビット階調対応の拡張プロファイルが成立した。そのほか、スケーラブル符号化や多視点映像向けの符号化方式も検討が進められている。また、パソコンの画面や医療、衛星センサ画像等の非自然画像に対して、画質を主観的ロスレス状態に保ったまま符号化する方式も検討されている。

HEVCを活用した映像システムに関連して、MMT(MPEG Media Transport)と呼ばれる新たな伝送プロトコルの国際標準が2013年に誕生している。ビデオストリームの伝送には、1994年に標準化されたMPEG-2 System規格が長らく使われてきた。MMTは今般の通信ネットワーク形態に合わせて、映像情報をより柔軟に伝送できるように策定されたものである。例えば、ビデオストリームやメタデータを放送網と通信網といった複数の異なる伝送路で配信することが容易となる。MMTは放送・通信連携サービスであるハイブリッドキャストなどで利用される。

2.3 処理の複雑度

HEVCは、圧縮性能が高い反面、従来方式よりも処理の複雑度が増加している。表2は符号化方式の仕様を比較したものである。MPEG-2やH.264/AVCと比べ、予測や直交変換処理のブロックサイズの大きさや予測モード数が増加している。また、SAO(Sample Adaptive Offset)などの新たな機能が追加されている。

一般的には、H.264/AVCと比較し、符号化処理で約2～5倍、復号処理で約1～3倍の演算量が必要と言われている。ただし、アルゴリズムや実装方法の工夫次第で低演算量化が可能である。当社は、2012年度に総務省から受託

表2. 符号化方式の仕様比較

	予測ブロック サイズ	動きベクトル 精度	参照ピクチャ数	イントラ予測	直交変換 ブロックサイズ	エントロピー 符号化	ループフィルタ
MPEG-2	16×16	半画素	2 以下	DC予測	8×8	2D-VLC	—
H.264	4×4～16×16 & H/V shape	1/4画素 (6tap)	5	9 モード 空間予測	4×4, 8×8 整数変換	CABAC CAVLC	デブロッキング フィルタ
HEVC	4×4～64×64 & H/V shape & AMP	1/4画素 (8&7tap)	5	35モード 空間予測	4×4～32×32 整数変換	CABAC	デブロッキング フィルタ, SAO

DC: Direct Current, 2D-VLC: 2 Dimensional-Variable Length Coding, H/V: Horizontal/Vertical,

CABAC: Context-based Adaptive Binary Arithmetic Coding, CAVLC: Context-based Adaptive Variable Length Coding,

AMP: Asymmetric Motion Partitions

した研究開発で、標準化作業段階にあったHEVCエンコーダを実験装置の1つとしていち早く試作した⁽³⁾。演算量が参照ソフトウェア⁽⁴⁾の1/2以下となる軽量な符号化アルゴリズムを開発し、233×220mmサイズの基板1枚でHDTV映像のリアルタイム符号化を実現した。

3. スーパーハイビジョンエンコーダ

2020年の東京オリンピックに合わせて、スーパーハイビジョン放送が計画されている。スーパーハイビジョン放送は、HDTVの16倍の8K映像による次世代の放送サービスであり、視聴者がその場にいると錯覚するほどの超高臨場感を実現する⁽⁵⁾。

情報量が膨大なスーパーハイビジョン映像を、現在の地上デジタル放送と同等の通信方式で伝送するには、映像情報の圧縮率を6倍(圧縮率: 約1/600)近くまで高める必要がある。その実現の基本方式としてHEVCが期待されている。当社はNHKと共同で、スーパーハイビジョン放送の今後の実用化研究の基盤として活用するため、世界初^(注1)となるスーパーハイビジョン向けHEVCリアルタイムエンコーダを開発した。

リアルタイム処理を実現するためには、ハードウェア構成、実装技術に加え、符号化の処理量低減も必須である。以降、開発したエンコーダの構成と、実装技術を述べる。

(注1) 2013年5月9日現在、当社調べ

3.1 エンコーダの構成

表3にスーパーハイビジョンエンコーダの諸元を示す。図1には装置を示す。1枚の制御CPU基板と17枚の符号化基板で構成されている。図2に符号化基板を示す。

図3に示すように、7680×4320画素のスーパーハイビジョン映像を短冊状に17分割し、17枚の符号化基板によって並列に符号化する。各基板が独立に動作するのではなく、分割の境界付近の画像データや符号化パラメータを基板間で相互に転送できるアーキテクチャとなっている。これによって、動き補償予測やデブロッキングフィルタなどの符号化処理を、基板を跨(また)いで実行することが可能となり、分割による圧縮性能の低下を防いでいる。

制御CPU基板は、外部端末からビットレートなどの符号化条件を受信するなどユーザーインターフェースの機能を持つほか、17枚の符号化基板で生成されたビデオストリー

表3. 装置諸元

映像フォーマット	7680×4320画素, 60frame/s
入力インターフェース	3G-SDI(SMPTE425M)を17本
符号化方式	HEVC メイン10プロファイル
ビット階調・色密度	10ビット, 4:2:0
ビデオレート	最大340Mbps
装置サイズ	480(H)×370(V)×545(D)(mm)

3G-SDI: 3GB/s Signal/Data Serial Interface
SMPTE: Society of Motion Picture and Television Engineers



図1. スーパーハイビジョンエンコーダ



図2. 符号化基板

ムを1本に束ねて装置外部へ出力する機能を持つ。また、分割画面間で映像の複雑度に偏りがあった場合に、基板ごとに画質の差が生じないように、各基板のビットレートを動的に補正する制御も行う。

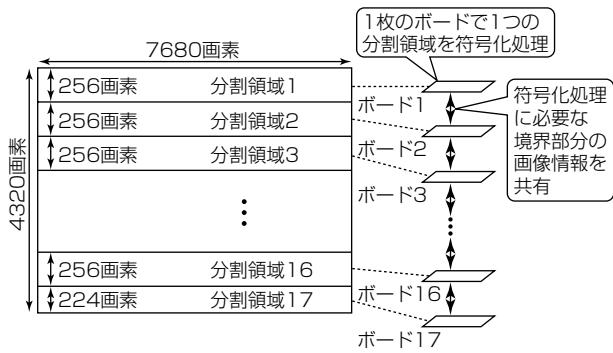


図3. 画面分割構成

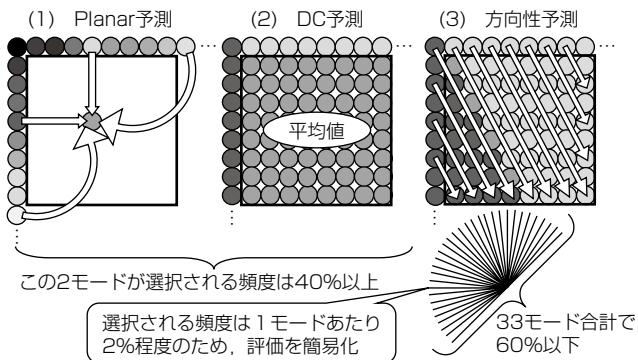


図4. イントラ予測モード選択の軽量化の例

3.2 実装技術

3.2.1 符号化アルゴリズムの軽量化

複雑度の高いHEVC符号化処理をリアルタイムに実現するためには、符号化アルゴリズムの軽量化が必須である。スーパーハイビジョンエンコーダを開発するに当たり、HEVCを構成する個々の要素の有効性を評価し、HEVCの性能を発揮するために真に必要な要素の選別を行った。

例えば、イントラ予測は、処理単位を 4×4 画素、 8×8 画素、 16×16 画素、 32×32 画素、 64×64 画素の5種類の中から選択することができ、また処理ブロックごとに35種類の予測モードが定義されている。スーパーハイビジョン映像で1フレーム当たり8,160個存在する全処理ブロックについて、処理単位と予測モードの最適な組合せを5種類×35モードの中から逐一選択していくと演算量は非常に膨大となる。しかし、実際に選択される予測モードを調べると、図4に例示するように偏りがある。この傾向を利用し、方向性予測として評価するモード数を絞り、処理を簡易化することで演算量を削減している。

イントラ予測以外に、例えばスーパーハイビジョン映像では効果が低い小さいブロックサイズでの予測や変換処理を省略している。こうした削減によって、符号化全体の演算量をH.264/AVCの2倍程度に抑えることできた。これによってリアルタイム処理に必要となる基本的なハードウェアの小規模化を実現した。

3.2.2 ハードウェア構成の最適化

スーパーハイビジョンエンコーダを現実的な装置サイズ

表4. FPGA実装時の機能分割

		回路規模	メモリ転送帯域
FPGA1	インター予測系	大	大
	イントラ予測系	中	小
FPGA2	インター予測生成	小	大
	変換・量子化系	中	小
FPGA3	エントロピー符号化系	小	中

で実現するために、ハードウェア構成も最適化している。HEVCの処理を機能的に分割し、基板あたり3石のFPGA(Field Programmable Gate Array)に実装している。回路規模とメモリ転送帯域がFPGAごとに平滑化されるように、各FPGAに実装する機能を表4のように配置している。

2.1節に示した基板間でのデータ転送では、データ転送タイミングを基板ごとにずらし、装置全体の負荷を分散させている。これによって、バス転送量やメモリ転送量の瞬間的なピークを下げ、部品点数や回路規模を削減している。

4. むすび

高い圧縮性能を持つHEVCの登場によって、今後新しい映像サービスが生み出される可能性がある。本稿で取り上げたスーパーハイビジョン放送はその一例であるが、それ以外に例えば、過去に敷設した狭帯域回線を用いた高品質な映像サービスや、可視光など伝送帯域に制限がある媒体による映像伝送等が挙げられる。HEVCは、HDTV以下の解像度でも高い圧縮性能を発揮する。伝送帯域の制約によってこれまで映像を伝送できなかった通信路でも、HEVCを使えば映像伝送を実現できる可能性がある。映像符号化技術が更に進歩し、圧縮率が上昇していくれば、世界中のありとあらゆる視覚情報を記録し、共有するようなサービスが生まれる可能性もある。今後の新しいサービスの創出と、それを実現する技術や製品の開発を通じて社会の発展に貢献していきたい。

参考文献

- Gary J. Sullivan, et al.: Overview of the High Efficiency Video Coding (HEVC) Standard, IEEE Trans. Circ. Syst. for Video Techn., **22**, No.12 (2012)
- K. Ugur, et al.: AHG4: Performance evaluation of HEVC on still picture coding, JCTVC-L0041 (2013)
- 山田悦久, ほか:新しい高性能映像符号化技術“HEVC”, 三菱電機技報, **87**, No.5, 289~292 (2013)
- High Efficiency Video Coding Test Model, Available (2013)
https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/HM-12.0/
- NHK放送研究所:高質感・空間再現メディアの実現
<http://www.nhk.or.jp/strl/vision/vision01/>

デジタルプリント品質向上への取組み

Towards Improved Digital Printing Quality

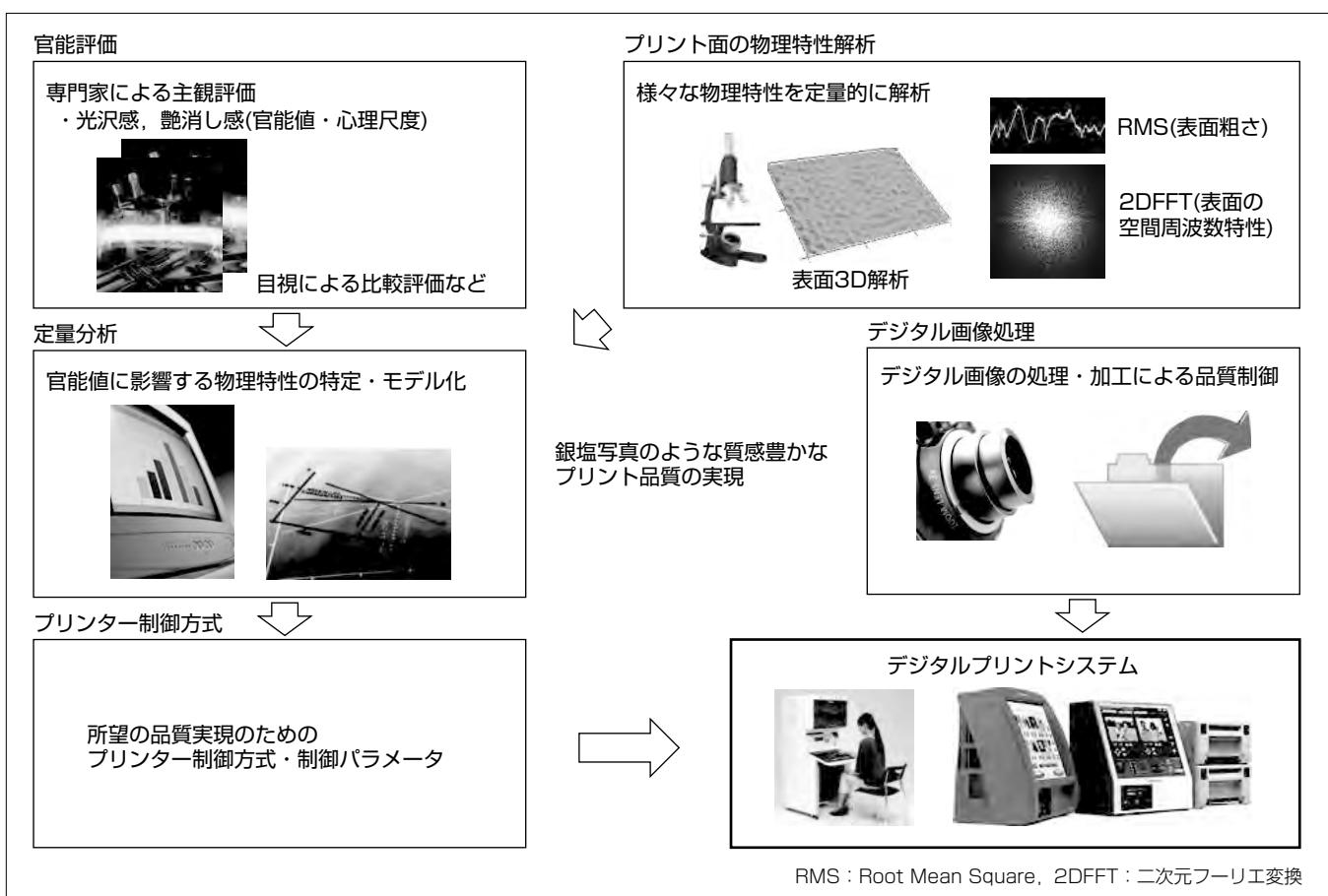
Momoyo Nagase, Ichiro Furuki, Shin Miura

要 旨

三菱電機では、コンビニエンスストアや家電量販店等でデジタル写真をセルフプリントするプリンターやプリントシステムを開発している。このデジタルプリントにも、高品質で、豊かな質感の再現が求められている。例えば、銀塩写真の“水面のような平滑な面質”すなわち光沢感、又は、表面が細かい粒状(凹凸)になっている面質による、“艶消し(マット)面質”的な再現などである。これらの要求に対して、当社は、①官能評価による心理的品質と物理特徴の関係分析による心理的品質の評価・モデル化、②詳細なプリント面質の分析に基づくプリント画質制御、③デジタル画像の処理による画質・質感の制御という総合的な検討を進めている。本稿ではこれらの取組みの一例について述べる。

品質の定量化では、光沢感に従来着目していた鏡面光沢に加え、表面構造パラメータが心理的な光沢感と強い関係があることを確認した。また、マット面質の再現では、同様に心理評価結果と、面の空間周波数の分析に基づき、中域周波数成分を強調するプリンターのサーマルヘッド制御によって質感の向上を可能とした。さらに、デジタル画像処理の一例として、画像そのものの光沢を強調する方式について述べる。画像から、拡散反射成分と鏡面反射成分を分離し、鏡面反射成分の強調によってデジタル画像の光沢感の改善ができる可能性を確認した。

これらの総合的な取組みによって、ユーザーの満足度が高い映像サービスを実現する。



銀塩写真のような質感豊かなプリント物を印刷可能なデジタルプリントシステム

官能評価によって得られる心理評価と、プリント面の物理特性解析結果を定量的に分析することによって、心理評価に影響する物理特性を特定する。結果をプリント表面処理パラメータに反映することによって銀塩写真のような豊かな質感をプリント面上で再現する。さらには、デジタル画像の信号処理によって、より高度でユーザーの要求に応える映像サービスを実現する。

1. まえがき

当社では、コンビニエンスストアや家電量販店等でデジタル写真をセルフプリントするプリンターやプリントシステムを開発している。デジタルカメラの高解像度化に伴い、プリンターにも銀塩(フィルム)写真並みの画質と質感の実現が求められるようになった。

当社のプリンターが採用している昇華型熱転写方式は、滑らかな階調表現が可能であり、画質そのものは銀塩写真と同等の品質が得られる転写方式である。しかし、その質感の再現では、まだ課題も多い。銀塩写真の質感を表す特徴の一つに、“水面のような平滑な面質”がある。このような平滑な面質を表現する指標が光沢感である。一般的に光沢のあるものには美しいものが多く、光沢が心理的な品質に与える影響は大きい。一方、銀塩写真の面質には、表面が細かい粒状(凹凸)になっている“艶消し(マット)面質”がある。この面質は、指紋が付きにくく、光による反射を抑えることができ、落ち着いた雰囲気のある情景やしっとりとしたやわらかな肌の表現に適している。ユーザーからは、このような面質、質感に対する要求も多い。

このような要求に対して、①官能評価による心理的品質と物理特徴の関係分析による心理的品質の評価・モデル化、②詳細なプリント面質の分析に基づくプリント画質制御、③デジタル画像の処理による画質・質感の制御という総合的な取組みを進めている。本稿では、光沢感とマット面質の再現を事例として、これらの取組みについて述べる。

2. プリント方式とプリント面の構造

プリント方式とプリント面の構造の違いを表1に示す。表中の表面断面は、実際のプリント表面の三次元形状測定結果を示す。銀塩写真とインクジェット方式では、プリントヘッドは印画紙に対して非接触で画像を形成する。したがって、プリント面質は印画紙の表面特性によって決まる。光沢仕上げには平滑な表面を持つ印画紙、マット仕上げには粒状面質の印画紙をそれぞれ用いる。これに対して、昇華型熱転写方式は、サーマルヘッドによってインクを熱し、加圧して画像を形成する。そのため、印画面にはサーマル

ヘッドに起因する凹凸が生じやすく、プリント面の光沢が低くなる傾向がある。また、画像形成時のプリント面には平滑性が要求されるため、粒状面質の印画紙を用いることができない。そこで、保護層であるオーバーコート層(OP)の転写エネルギーに強弱をつけてOP面を凹凸状に仕上げることでマット面質を実現している。

3. 光沢感の改善

3.1 光沢感と光沢度

一般に、光沢は鏡面光沢度として定量的に定義されるが、実際に観察者が知覚する心理的な光沢感と、物理量としての光沢度は、必ずしも一致しない。図1は、表面特性の異なる印画サンプルの光沢感に関する官能評価実験を行った結果と、物理量である光沢度の関係を示すものである。官能評価には表2に示すA～Gの7種類のプリントサンプルを用いた。Aは染料系インクジェット(写真光沢紙)、B～Fは記録メディア(インクシート、印画紙)やプリンター機種が異なる昇華型熱転写方式、Gは銀塩写真のプリントである。評価対象分野の専門家11人に7種類のサンプルを同時に提示し、サンプル表面の光沢感が高いと思う順に順位を付け、その順位を正規化することで順位間を定量的に表した。図1の左側は、官能評価結果を距離尺度に変換して表したものであり、上にいくほど心理的な尺度である光沢感

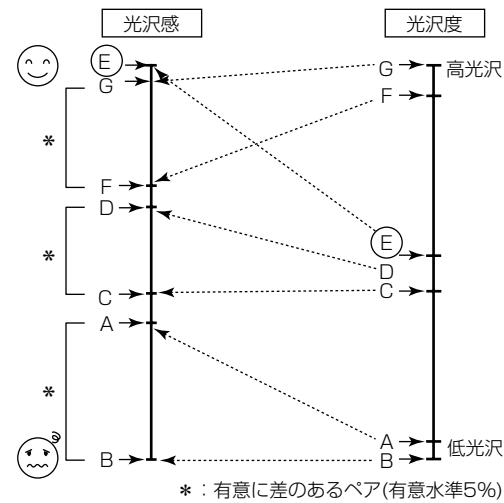


図1. 光沢感と光沢度

表1. プリント方式とプリント面の構造

	銀塩写真	インクジェット	昇華型熱転写
構造	露光(非接触) 床 鏡 青 保護層 発色層 支持体	インク吐出(非接触) マゼンタ シアン 受容層 支持体	加圧・加熱(接触) サーマルヘッド 保護層(OP) OPインク シアン 受像層 支持体 マゼンタ
表面断面	1μm 0μm 280μm 滑らかな面	1μm 0μm 280μm 表面に微細穴	1μm 0μm 280μm 表面凹凸

が高いことを示す。右側は定量的物理量である光沢度であり、その関係を点線で示した。図中、光沢感のAとC, DとF, GとEには有意差が見られず、有意差のある距離は“*”を付して記している。この図から明らかのように、光沢感と光沢度には大きな差異がある。最も官能評価値の高いEの光沢度は、G, Fよりも低く、プリント面の心理的な光沢感は、光沢度が高い方が良好であるとはいえない。

光沢度と光沢感の関係では、次の2点に大きな問題がある。

- (1) 光沢感が最も高いサンプルEの光沢度が低い。
- (2) 光沢感が最も低く、その次に低いサンプルAとの差が有意であったサンプルBの光沢度が、Aと同程度である。

3.2 光沢感とプリント面物理特性の関係

3.1節の結果を受け、心理的光沢感と光沢度以外の物理量の関係を分析した⁽¹⁾。心理的光沢感とプリント面の物理特性との相関を図2に示す。

評価に用いたプリントサンプル表面の物理特性は次のとおりである。光沢度は20°鏡面光沢度、Raはプリント表面の算術平均粗さ、du(dullness)は波長0.1mm以下の微細表面における光の散乱度、DOI(Distinctness of Image)は表面に映る像の鮮明さを示す写像鮮明性、Wa～Weは、プリント面の波長0.1mmから30mmまでの表面構造を表すパラメータで、自動車の塗装業界でよく用いられているオレンジピールの指標となるものである。

光沢感と高い相関が見られたのは、光沢度、Ra、du、Wb、Wcであった。これらの特性の中で、3.1節で述べた2つの問題点を説明できるパラメータは、duである可能性が高い。Eのduは全サンプル中最小であり、Bでは最大であるだけではなくサンプルAとの差異も大きい。

今後、他のサンプルの評価を含めたデータの収集とモデ

ル化の検討を進める予定であるが、従来の光沢度の向上にだけ着眼した改良に対し、光の散乱度などを指標に加えることで、より心理的光沢感の向上を図れる可能性が高いことが分かる。

4. マット面質の改善

昇華型熱転写プリンターにおけるマット仕上げ時の面質は、プリント面の凹凸パターン(マットパターン)によって決まる。これまで主観評価に基づく開発を行ってきたが、更なる改良のために詳細分析を行った。一般的に高く評価されているマットプリント面(基準)を定量的に分析し、その分析結果と、現状のマットパターンの比較によって改良を図った。この評価では二次元フーリエ変換を用いた評価手法を適用した。評価対象となるマットサンプル面は透明なOP層によって形成されているため、マット表面状態を単純に画像情報として取得することは困難である。そこで、レーザ顕微鏡によってマット面の凹凸高さを計測し、その高さ情報を二次元化することによってマットパターンを画像情報として取得し、マットパターンの周波数(空間周波数)特性を評価した。従来及び改善後のマットパターンの空間周波数特性を図3に示す。図の横軸は空間周波数の自然対数、縦軸はパワースペクトルの自然対数を示す。基準パターンと比較して、従来のマットパターンは中域周波数成分が低いことが分かる。サーマルヘッドの制御方法を改善し、中域周波数成分に相当する凹凸パターンを高密度化した。これによって、評価が高い基準パターンと同等の周波数特性を達成した。

この中域周波数成分を強調するサーマルヘッド制御は、既にプリンターCP-K60DW-Sに搭載され、ユーザーから高い評価を得ている。

5. デジタルコンテンツの光沢部分検出・強調手法

プリント物の品質を向上又は、画質を変更する手法の1つとして、デジタルコンテンツそのものの加工がある。ここでは、3章で述べた光沢感の改善を画像処理によって

サンプル	A	B	C	D	E	F	G
20°鏡面光沢度	37.1	36.2	60.6	63.7	63.9	81.1	83.3
Ra(μm)	0.027	0.074	0.042	0.037	0.031	0.036	0.019
du	35.7	58.0	41.6	37.8	26.1	39.2	29.2
DOI	73.7	—	70.5	71.7	74.5	71.4	73.7
Wa(0.1～0.3mm)	31.4	—	35.8	38.3	31.8	38.5	38.8
Wb(0.3～1 mm)	45.2	—	48.2	52.5	58.9	51.0	60.0
Wc(1～3 mm)	39.4	—	40.1	43.9	54.0	41.4	50.3
Wd(3～10mm)	34.3	—	28.0	33.9	38.8	33.6	38.4
We(10～30mm)	25.0	—	21.2	24.6	40.3	24.5	27.2

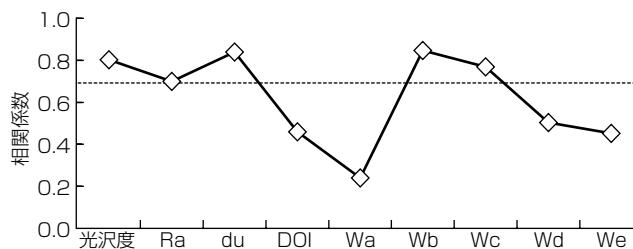


図2. 光沢感とプリント面の物理特性

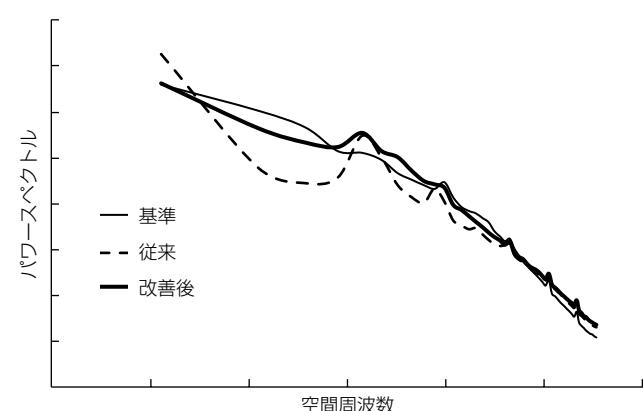


図3. マットパターンの空間周波数特性

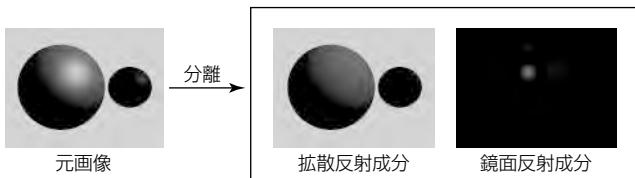


図4. 拡散反射成分と鏡面反射成分の分離結果

実現する手法の基礎検討について述べる。

5.1 光沢部分の検出手法

物体表面で反射した光が人の目に入ることによって、人は物体の形、色、質感を感じることができる。Shaferらは“反射光は拡散反射光と鏡面反射光の線形和で表される”という二色性反射モデルを提案した⁽²⁾。拡散反射光を I_d 、鏡面反射光を I_s とすると、反射光 I は式(1)で表される。

拡散反射とは、入射光が物体表面で様々な方向に反射することをいい、鏡面反射とは、入射光が入射面法線ベクトルに対して反対方向に反射することをいう。一般に、艶のあるものほど、鏡面反射成分の割合が高いとされている。

デジタルプリントのコンテンツの光沢部分を強調するために、まず、コンテンツからこの鏡面反射成分を抽出する。図4に拡散反射成分と鏡面反射成分の例を示す。

鏡面反射成分の色味は光源の色味と同一と近似でき(NIR(Neutral Interface Reflection)仮定)⁽³⁾、拡散反射成分は光源色と物体表面色との積で表現できる。例えば光源色を白色と仮定した場合、同一素材、同一色の物体では拡散反射成分の色味は一様となる。拡散反射成分に関しては、この性質を利用することで、同一物体色の拡散反射成分と鏡面反射成分とを分離することができる⁽⁴⁾。

これらの成分を正確に分離するためには、同一素材・色部分を高精度に識別する必要がある。コンテンツ中の光沢部分は一般に輝度が高く、カメラ感度の限界によって白飛びしてしまうことが多い。そのため、カメラ観測輝度の上限に達している部分に関しては拡散反射成分検出処理からは除外し、周囲画素の色相を参照することによって、精度良く拡散反射成分を抽出する。

5.2 光沢部分の強調手法

検出した鏡面反射成分を定数倍し、拡散反射成分に足し合わせることで、鏡面反射成分を強調した画像を生成する。光沢部検出・強調の処理フローを図5に示す。

鏡面反射成分を強調した画像と元画像に対して、画像評価対象分野の専門家4名による主観評価を行った。その結果、テストサンプルでは、鏡面反射成分を2倍にしたとき、自然性を損なうことなく、光沢感が向上することが分かった。多様な画像から鏡面反射成分を抽出する方法など課題はあるが、方法の可能性を確認することができた。

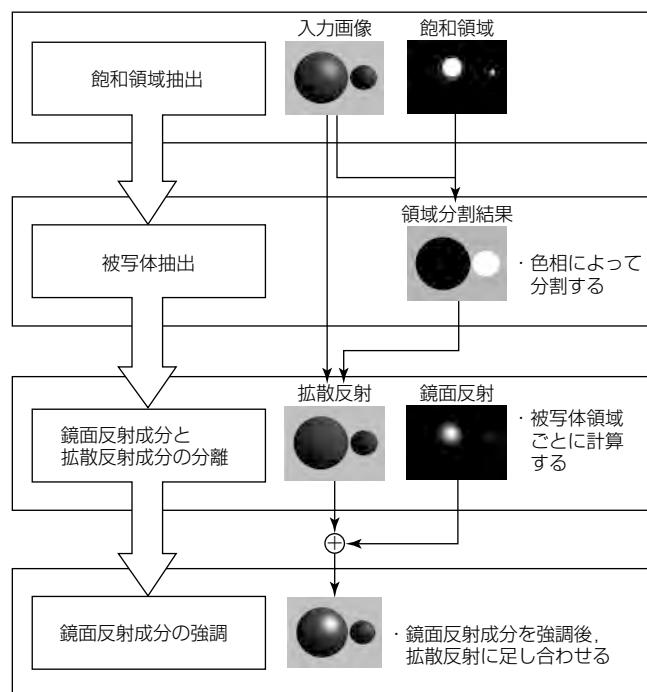


図5. 光沢部検出・強調の処理フロー

6. む す び

デジタルプリントをより質感豊かにするためのプリント面の面質改善検討について述べた。画像の品質、質感は、従来官能評価による比較が主に行われてきたが、この方法では改善のための制御手法を導くことができない。光沢感、マット面質で事例について述べたように、心理尺度と物理的特徴の分析を進め、将来的に画像品質、質感のモデル化を行い、これに基づくプリンター制御によって、高品質で多様な質感の実現を図る。さらには、画像処理技術によるコンテンツ自体の加工に関する研究開発も推進する。これらによって、ユーザーの嗜好(しこう)や要求に応じた、満足度が高い豊かな映像情報サービスを実現していく。

参 考 文 献

- (1) 古木一朗, ほか: デジタルプリント面における心理的光沢度に基づく評価, 電子情報通信学会総合大会講演論文集, **12** (2013)
 - (2) Shafer, S.A., et al.: Using Color to Separate Reflection Components, COLOR Research and Application, **10**, No.4, 210~218 (1985)
 - (3) Lee, H.C., et al.: Modeling light reflection for computer color vision, IEEE Trans. on PAMI, **12**, No.4, 402~409 (1990)
 - (4) 肥後智昭, ほか: 二色性反射モデルに基づくリアルタイム鏡面反射成分除去, 電子情報通信学会技術研究報告, PRMU, **106**, No.230, 77~84 (2006)

映像同定技術“Video Signature”とその応用技術

工藤大樹*
鶴崎真理子*
阿倍博信**

Outline and Application of Video Identification Technology "Video Signature"

Daiki Kudo, Mariko Tsurusaki, Hironobu Abe

要 旨

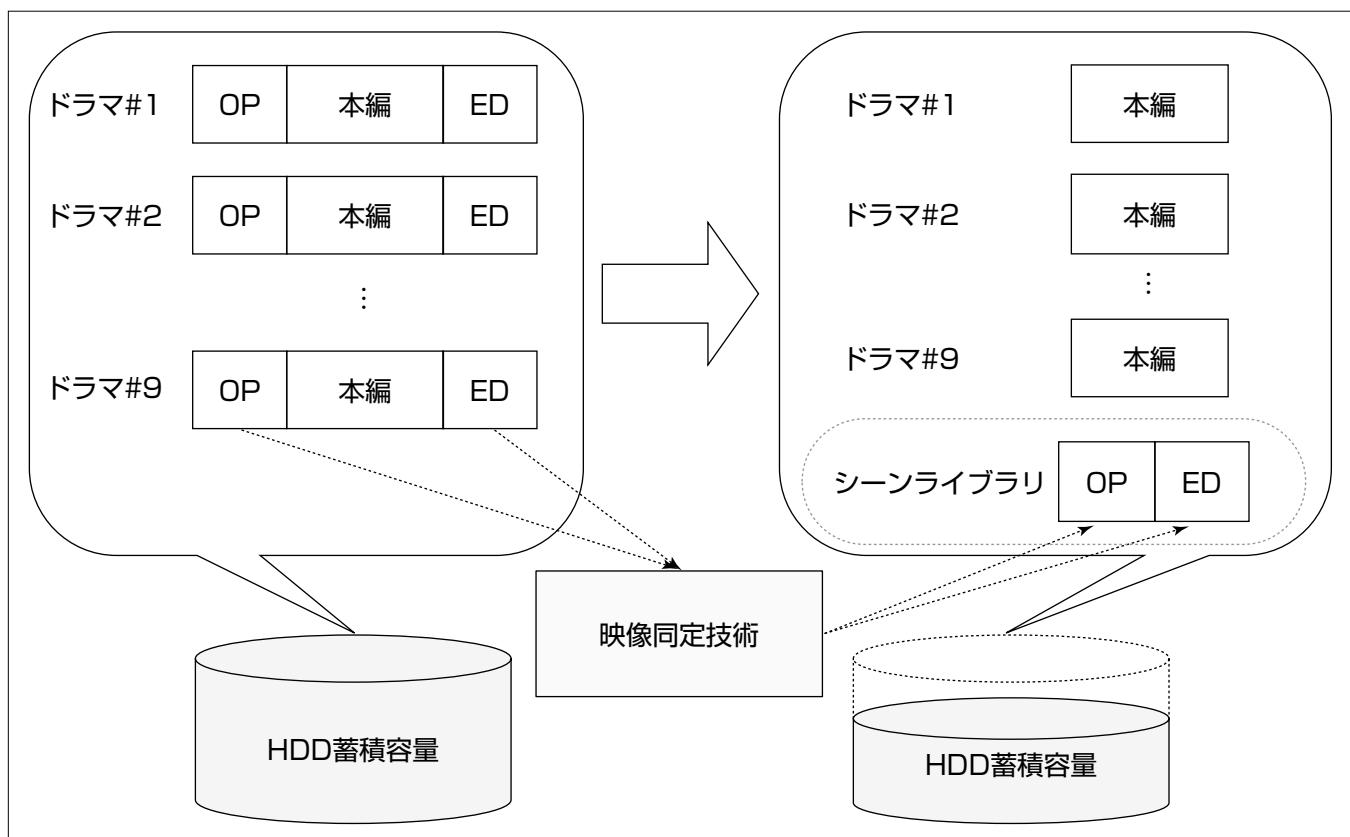
大量の画像から所望の画像を検索する“画像検索技術”には、用途・目的に応じて様々な方式が存在する。その一つに同一の映像を検索する“映像同定技術”がある。既に、各種の加工に対しても頑健に同一性の検証が可能な手法が確立しているため、インターネット上に不正に共有された映像の発見などに利用できる。本稿では、この映像同定技術の新たな応用として、HDD(Hard Disk Drive)レコーダー向けの蓄積効率向上について述べる。

一般に、レコーダーに蓄積された映像には、ドラマのオープニングなど同一映像が複数記録されることが多い。これを、映像同定技術によって検出し、シーンライブラリとしてまとめる。元の映像からは、検出された同一映像は削除し、再生方法を示したメタデータを付与し、再生時はこの

メタデータを利用して、シーンライブラリから映像を読み出し、再生する。これによって効率的な映像蓄積が実現できる。

この方式では、処理時間と、映像の編集処理が問題となる。そこで、前者に対しては、音声情報を利用し、処理区間の候補を絞り込むことで高速化を行った。後者に対しては、映像符号化の単位(Group Of Picture: GOP)に基づく検出処理を行うことで、映像の削除、接続のための付加的な復号・再符号化処理を不要としている。

今後、同一映像だけでなく、類似映像を利用した方式についても検討し、単に効率的な蓄積にとどまらず、映像情報の構造化・情報活用にも応用する予定である。



映像同一部分を利用したHDD蓄積効率向上

同一ドラマをシリーズ(#1～#9)で録画したHDDレコーダーを想定している。ドラマのオープニング(OP)とエンディング(ED)は同一映像であることから、これらをシーンライブラリとして蓄積し、各ドラマは本編だけを蓄積している。これによって、HDD全体での蓄積効率が向上する。再生時はメタデータを利用することで、オープニングとエンディングを含めて再生する。

1. まえがき

撮像から蓄積・配信にいたるまで、画像(静止画)・映像(動画)処理技術は急速に発達し、ユーザーは大量の画像・映像を保持し、気軽に楽しめるようになった。

これに伴い、所望の画像を探す画像検索技術の研究も盛んに行われている。画像検索には、幅広い用途・目的に合わせた様々な技術が存在する。例えば、キーワードを使ったインターネット上の画像検索⁽¹⁾のほか、顔や人物の検出⁽²⁾、構図を利用した映像検索技術⁽³⁾、同一の画像・映像を検出する画像同定技術⁽⁴⁾等が挙げられる。

ユーザーが保持している画像に目を向けると、デジタルカメラやスマートフォンで自ら撮影したものと、他者が制作したコンテンツの2つに大別できる。このうち、後者の代表的なものとして、家庭用HDDレコーダー(Hard Disk Driveを用いる映像蓄積装置、以下“HDDレコーダー”という。)内の録画番組が挙げられる。

HDDレコーダーは、当初、アナログ放送を録画するものから始まり、その後、高画質のデジタル放送サポートや搭載チューナー数の増加による複数番組同時録画などによって、録画データ量は飛躍的に増大している。大容量のHDDを持つ機器でさえ、蓄積容量が不足する場合も多く、外部HDDを接続し利用するユーザーも少なくない。今後も、より高解像度の4K/8K放送が予定され、HDD容量はますます逼迫(ひっぱく)するものと思われる。

この解決方法の一つが、放送映像をより高効率の符号化方式(例えば、H.264)に変換、圧縮する方法があり、既にHDDレコーダーでの実用化もされている。これは、録画番組単体での蓄積効率向上を図る方法である。

これに対し、本稿で述べる技術は、HDDに蓄積される各映像の特徴量を抽出し、その映像間の同一性を利用して蓄積効率向上を図る手法である。

2. 画像・映像同定技術

2.1 概要

画像・映像の同一性を判定する技術が研究開発され、標準化が行われた。これが、画像・映像同定技術(Visual Signature)⁽⁴⁾であり、静止画向けのImage Signature⁽⁵⁾と、映像向けのVideo Signature⁽⁶⁾の2種類からなる。

画像・映像同定について、静止画像の場合を例に述べる(図1)。女性の画像(図中の“検索キー画像”)に対して、“類似”的意味を、“女性の画像”とするのであれば、図中の全ての画像が検索結果として出力されるべきである。しかし、画像同定は、画像コンテンツの同一性を判定することが目的であり、回転・モノクロ化・左右反転等の画像処理を施し、画素のパターンとしては全く異なる画像でも、同一として、検索する技術である。

映像データに対する同定を目的としたVideo Signatureは、この画像同定の技術を動画に応用した技術で、図2で示しているように、フレームごと、すなわち個々の静止画の特徴量と、フレーム90枚で構成される一定時間長(セグメント)ごとの特徴量から同一性を検出する手法である。個々のフレーム特徴量に対しては詳細なマッチングを、セグメント特徴量に対しては、大まかなマッチングを行うことで、フレームの加工やフレームレートの変化等、映像加工に対しても誤りが少なく、頑健性の高い同定が可能である⁽⁴⁾。



図1. 同一の画像

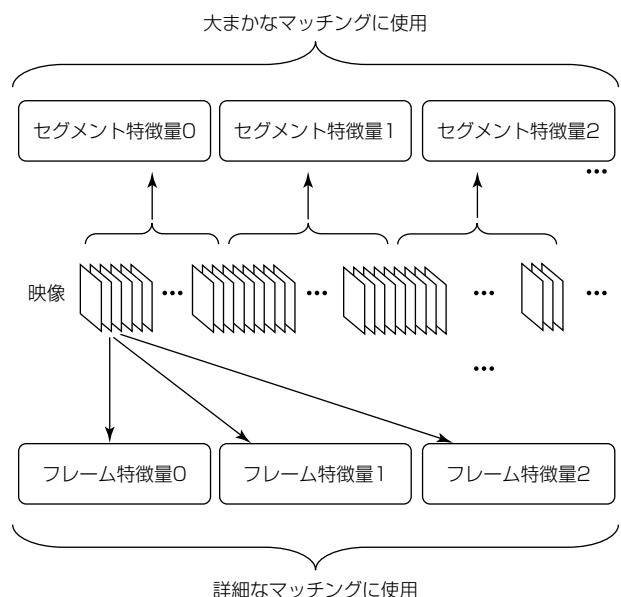


図2. Video Signatureで使用する特徴量

2.2 Video Signatureの応用

Video Signatureは、インターネット上の動画共有サイトなどの映像コンテンツの不正流通の検出を目的に考案されたものであり、その有効性が確認されている。

3. 映像同定を応用した蓄積効率向上

3.1 概 要

HDDレコーダに蓄積されている番組は、ユーザーの嗜好(しこう)に基づき、シリーズ番組(ドラマ・アニメ等)をまとめて記録しているケースが少なくない。これらシリーズ番組のオープニングとエンディングや、スポーツ番組におけるリプレイ等、レコーダ内には同一のシーンが多数存在する。

映像同定技術を利用すれば、レコーダ内で同一の映像部分を特定することが可能になり、レコーダ全体の蓄積容量を削減することが可能になる。

図3に例を示す。映像A／Bから同一映像部分(①, ②)を特定する。同一映像部分は、“シーンライブラリ”として別途保存し、元の映像からは削除する。これをレコーダ内に存在する全ての映像に対して行うことによって、レコーダ内に複数存在する同一部分はシーンライブラリに保存される一つを除き削除でき、蓄積容量の削減ができる。

なお、編集済み映像には、再生用のメタデータを付与する。この記述に従い、シーンライブラリを参照、再生することで、元の映像と同じ映像の再生が可能である。

3.2 手 順

3.1節で述べた蓄積効率向上の方法を実現するため、具体的には次の手順で行う。

- (1) レコーダ内に記録されている映像のVideo Signature特微量を抽出する。
- (2) 抽出した特微量を利用して、レコーダ内の各映像の同一部分を特定する。
- (3) 特定した同一映像部分を抜き出してシーンライブラリを作成する。
- (4) 各映像から同一映像部分を削除して編集済み映像を作成する。

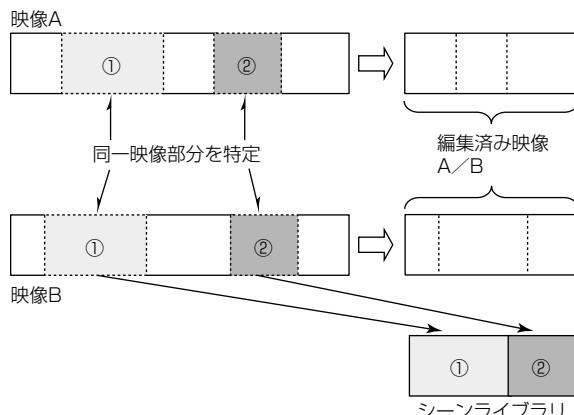


図3. 同一映像部分を利用した蓄積容量削減

成する。同時にそれぞれの再生用メタデータを作成する。

この手順は、大きくは前半(1), (2)の特微量抽出・同一映像部分特定と、後半(3), (4)のシーンライブラリ及び編集済み映像の作成(編集)に分かれる。それぞれの処理高速化について述べる。

3.3 特微量抽出速度の改善

Video Signatureを使用した映像同定を行う場合、映像中の全フレームについて特微量を抽出するため、処理量が大きなものとなる。そこで、音声情報を利用し、同一映像区間候補を限定し、処理量を削減する方式を考える。

図4で示すように、映像コンテンツでは、シーンが切り替わる箇所に無音区間が存在することが多い。そこで無音区間を判定し、その位置から映像コンテンツ内の同一シーンを含む区間の候補を抽出する。特微量抽出は、この区間に對してだけ行う。これによって、特微量抽出だけではなく、同一性の検証のためのマッチング処理についても大きな処理量削減が可能となる。

実際のテレビ番組を用い評価実験を行った。音声情報による特微量抽出区間の限定を行ったところ、特微量抽出処理時間は通常のVideo Signatureを使用したときの38%に削減することができた。さらに、マッチング処理を含めた全体の処理時間は、45%に削減となった(図5)。

今後の実用化に向けては、更なる処理量削減が必要である。これは、同一映像区間候補の抽出精度を高めることで実現できる。声のパワーの一定しきい値判断による現状の無音区間判定は誤検出も多く、処理量削減効果が十分ではない。この精度向上に加え、シーンの切り替わりを特定する無音以外の情報の活用も可能である。例えば、EPG(電子番組表)の時刻情報を用いれば、番組自体の開始・終了時刻を特定できる。今後これらを検討するとともに、サーバで精度よく同一映像区間を検出、配信するクラウドサービスの可能性も検討する予定である。

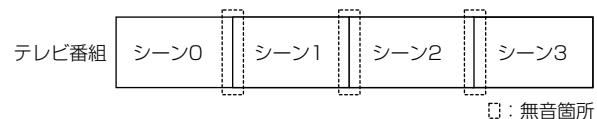


図4. 無音区間利用法

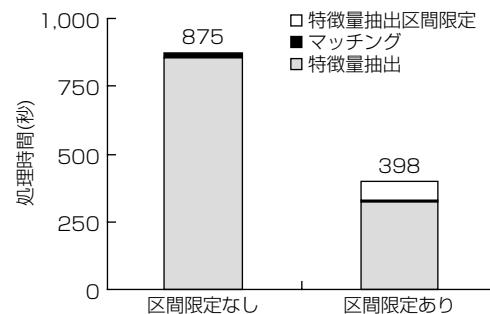


図5. 特微量抽出処理時間の削減

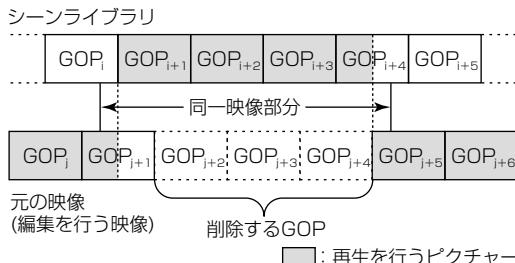


図6. GOP単位での編集

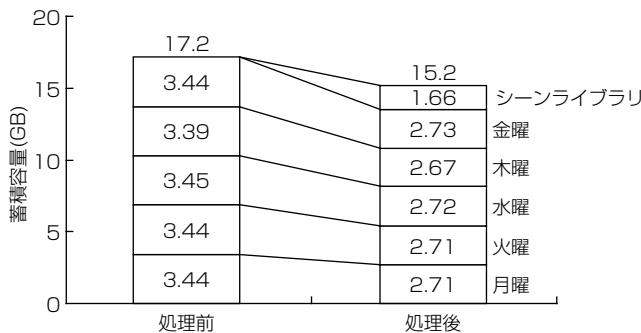


図7. 蓄積容量の変化

3.4 編集方法

同一映像部分の削除、シーンライブラリから読み出した映像との接続では、映像符号化技術との整合性が問題となる。H.264に代表される映像符号化技術では、GOPと呼ばれる単位で、画像間の相関を利用した圧縮・符号化が行われる。そのため、映像符号は、前後のフレームに依存した情報であり、編集処理を行うには、一度復号し、編集を行い、再符号化を行う必要がある。この処理量も大きなものとなる。

そこで、符号化方式に着目し、復号・再符号化を不要とするGOP単位での編集手法を提案する。

図6では、シーンライブラリと元の映像の同一映像部分はGOPの区切りと一致していない。シーンライブラリではGOP_iの途中～GOP_{i+4}の途中まで、一方で編集対象の映像では、GOP_{j+1}の途中～GOP_{j+5}の途中までが同一映像部分である。

このとき、元の映像はGOP_{j+2}～GOP_{j+4}までを削除し、同一シーンの開始GOPであるGOP_{j+1}と終了GOPであるGOP_{j+5}を残す。つまり、同一映像部分に完全に内包されるGOPだけを削除し、シーンライブラリ側は、同一映像部分を完全に外包するGOPで構成する。これによって、編集によってGOPが変化することなく、映像の再符号化処理は不要となる。ピクチャー単位で編集する場合に比べて、蓄積する映像にはオーバラップ区間が生じるが、蓄積容量は2GOP増えるだけであり、大きな影響はない。

なお、再生用のメタデータには、ピクチャー単位の再生情報が含まれる。

3.5 削減効果

この手法を平日5日分の同一番組（合計5時間、平均8.0MbpsのHD映像）に対して適用し、効果を検証した。

その結果、図7に示すように11.6%（17.2GBから15.2GB）の蓄積容量削減効果が得られた⁽⁸⁾。

この手法は、レコーダ内に多くの映像を録画しているほど同一シーンも増加し、効果が大きくなると考えられる。より長期間、番組を録画することで、更なる圧縮効率向上が期待できる。

4. むすび

映像検索技術の一つである映像同定技術“Video Signature”を利用したHDDレコーダにおける蓄積容量削減方法を述べ、11.6%の削減効果があったことを示した。今後、EPGなどの情報を活用すれば、更に効率的な検出も可能となる。

我々は、同一映像部分ではなく、類似の映像シーンを活用する蓄積効率向上についても検討している⁽⁸⁾。さらに、昨年、次世代符号化規格であり圧縮効率が高いH.265が標準化された⁽¹⁰⁾。これらを組み合わせ、HDDを最大限に活用可能な方式を開発する予定である。この技術は、単に蓄積容量の削減にとどまらず、映像情報に対する付帯情報の付与、構造化という観点でも期待が高い。監視映像への適用などへの適用も検討し、映像情報の有効活用を図る。

参考文献

- (1) Google画像検索
<http://www.google.co.jp/imghp>
- (2) 山内悠嗣, ほか: 画像からの統計的学習法に基づく人検出, 電子情報通信学会論文誌.D, **J96-D**, No.9, 2017~2040 (2013)
- (3) 望月貴裕: 画像解析技術を利用した映像検索, NHK技研 R&D, No.128, 18~25 (2011)
- (4) 工藤大樹, ほか: 画像同定技術: Video Signature, 三菱電機技報, **85**, No.11, 657~660 (2011)
- (5) ISO/IEC 15938-3:2002/Amd 3: 2009
- (6) ISO/IEC 15938-3:2002/Amd 4: 2010
- (7) M.Tsurusaki, et al.: An Efficient Recording Method for Video Recorders Using the Video Signature Tools, Proc. of 2014, 528~529 (2014)
- (8) 工藤大樹, ほか: 映像同定を利用したレコーダ蓄積容量削減と再生方法, P2-09, PCSJ2014
- (9) T. Semitsu, et al.: Video Re-compression Using Inter-scene Correlation among Multiple Video Streams, Proc. of IWAIT2014, 231~236 (2014)
- (10) ISO/IEC 23008-2

タッチパネル感度向上技術

佐々木雄一*
根岸博康*

High Sensitivity Detection Method for Projected Capacitive Touch Panels

Yuichi Sasaki, Hiroyasu Negishi

要旨

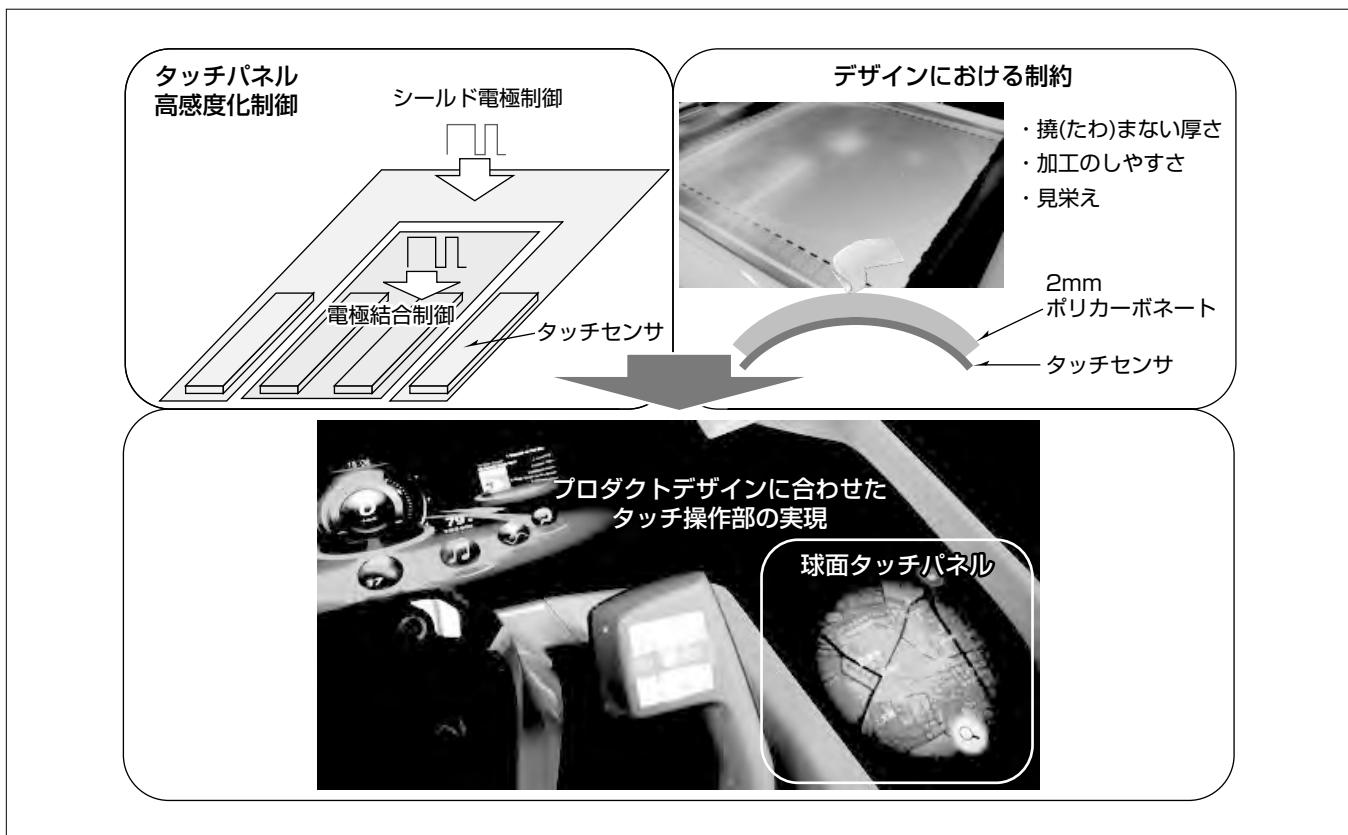
タッチパネルインターフェースは、ユーザーに直感的な操作を提供するため、スマートフォン、タブレット端末にとどまらず、家電、車載機等で広く用いられるようになってきている。今後より広い製品分野に適用するために、デザイン面では、曲面構造でのタッチ操作の実現が求められている。

例えば球面形状のタッチパネルでは、加工の容易性から、表面材料として、ポリカーボネート(以下“ポリカ”という。)を用いることが望ましいが、硬さを考慮すると、その厚さは、2 mm以上が求められる。しかし、2 mm厚のポリカ越しに静電容量方式のタッチパネルで、指のタッチを感度良く検出することは、これまで困難であった。

そこで感度向上のため、電極結合制御、重畠検出制御、シールド電極制御を開発した。電極結合制御は電極を束ねることで指の接触による実効面積を増やし、重畠検出制御は電極間の重畠によってセンサ境界での感度低下を防止し、シールド電極制御は指以外による寄生容量を抑制する技術である。

10インチ、直径350mm、表面を2 mm厚ポリカでカバーした球面タッチパネルを試作して評価した。その結果、従来の検出方式に比べ、提案方式のタッチ検出信号はS/N比で平均3.0dB以上の改善があり、感度の向上を確認した。

今後は手袋越しでも利用可能となるよう感度向上を図るとともに、曲面を生かしたインターフェースの実現を目指す。



タッチパネルの感度向上技術の応用

タッチパネルの感度向上技術を適用することで、センサと指の間隔が離れている場合でも、感度良く指の接触を検出できる。タッチパネルの表面の形状の自由度が格段に向上するため、デザイナーが思い描くプロダクトデザインを実現することが可能となる。車載機での車のデザインに合わせた曲面タッチパネルディスプレイ、各種家電の意匠を損なうことのないユーザーインターフェース等、様々な用途に適用する予定である。

1. まえがき

静電容量型方式によるタッチパネルインターフェースは、スマートフォンやタブレット端末だけでなく、家電や車載製品の操作性・デザイン性を向上させる技術として利用が広がっている。操作面では、表示されたアイコンなどの直接的操作で分かりやすいことに加え、軽いタッチで反応するためジェスチャインターフェースに適しており、デザイン面では、縁がないフラット構造の操作部を実現できる。今後、更に広い製品分野にタッチパネルインターフェースを適用するために、デザイン面からの要求として、プロダクトデザインに合わせた様々な曲面構造でタッチ操作を実現する技術が求められている。

これに対し、三菱電機は球面加工した静電容量方式のタッチパネルの開発、試作を進めている。球面形状のタッチパネルは、加工しやすくするために上面部を形成する材料としてポリカを用いることが望ましい。表面の硬さを確保するには、2 mm以上の厚さとする必要があるが、2 mm厚のポリカ越しで、十分に感度の高い指の接触検出は困難であった。

そこで、ポリカで表面加工された球面など曲面タッチパネルでも、感度良く指の接触を検出するため、電極結合制御、重畳検出制御、シールド電極制御を開発した。

2. 投影容量型タッチパネルの検出原理

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種類の検出方式がある⁽¹⁾。この研究で用いる球面形状のタッチパネルは、投影容量型の検出方式に基づいている。球面タッチパネルのセンサ電極は、曲げに強く、抵抗値が均一で透明な導電膜を用いている。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(図1)。以降では、この形状を1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接触すると、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、指の接触を検知することができる。ただし、球面タッチパネルは上部に

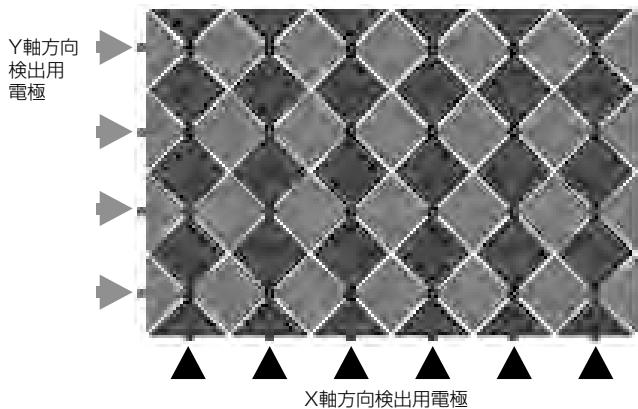
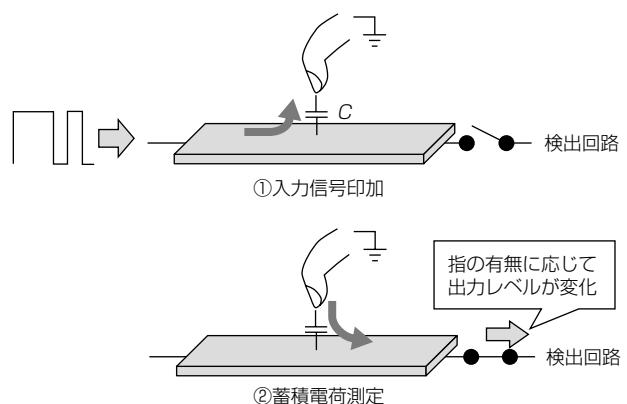


図1. 投影容量型タッチパネルのセンサ電極構成



入力信号印加時に電荷がCに蓄積され、これを測定することで、指の有無を判定可能

図2. 指の静電容量の検出方法

2 mm厚のポリカがあるため、得られる検出値が小さく、指の接触を判断する上で十分でない。

3. 球面タッチパネルの高感度化

試作した球面タッチパネルとその高感度化制御部分についての構成を図3に示す。このシステムは、厚さ2 mmのポリカ越しで十分な指の検出値を取得するために、高感度化処理(電極結合制御)、ノイズ低減処理(シールド電極制御)、指位置に依存しない感度安定化処理(重畳検出制御)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Cは、2つの間の面積Sに比例し、距離d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、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。図4のように1本ずつ検出を行う場合、指と電極間の接触面積がSで指に発生する静電容量はC₁あったのに対し、2本結合の場合S'(>S)となり、面積が増える分だけ指に発生する静電容量C₂は大きくなる。これによって、タッチパネル

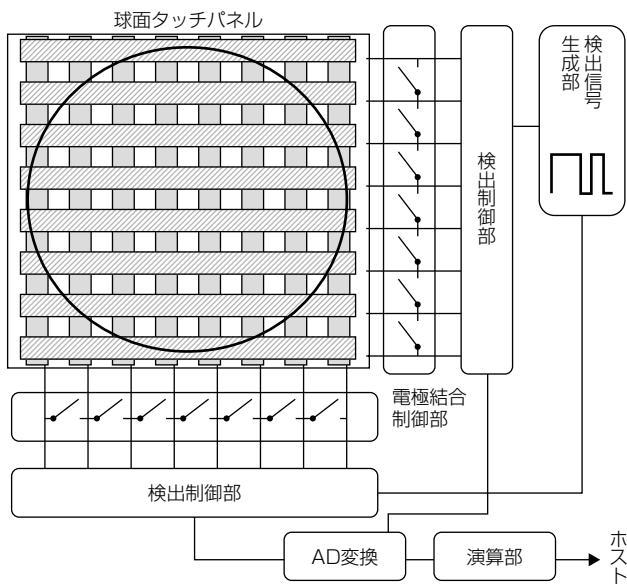


図3. 構成

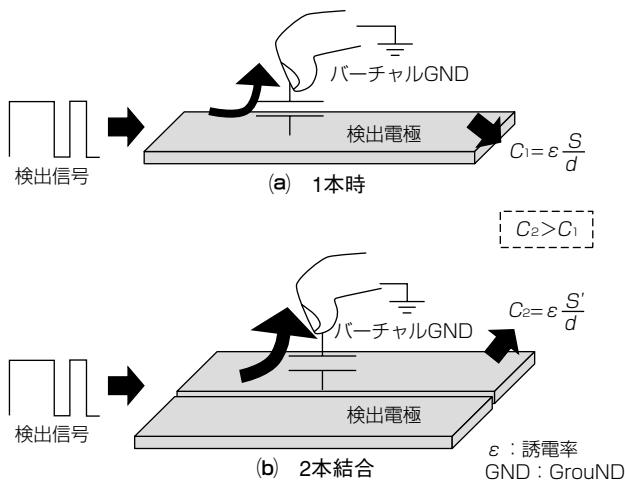


図4. 電極結合制御

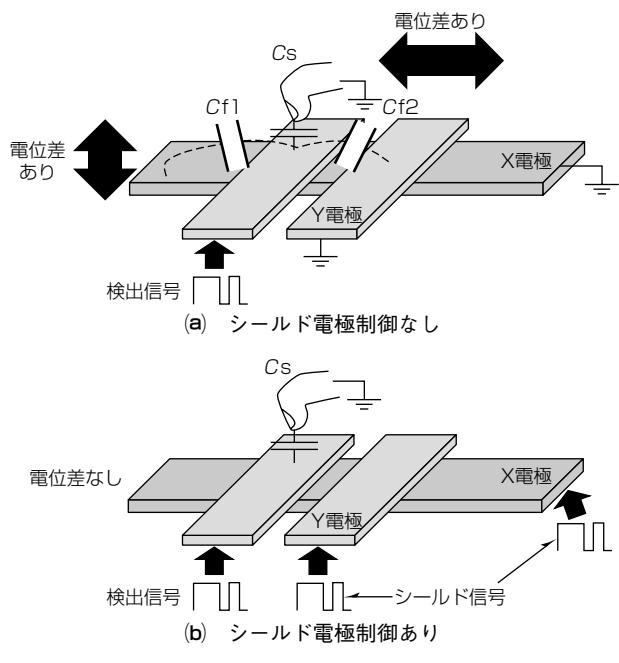


図5. シールド電極制御

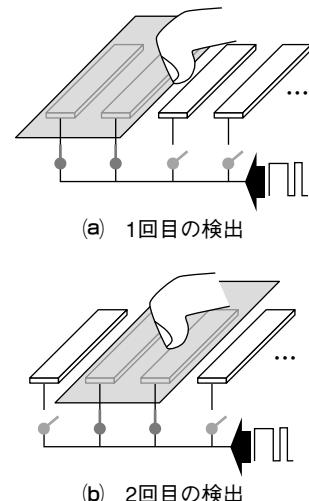


図6. 重畠検出制御

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。

3.2 シールド電極制御

タッチパネルに複数の電極がある場合、指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 C_{f1} , C_{f2})。この場合、検出電極から寄生容量 C_{f1} , C_{f2} を介して周囲の電極に信号が流れ込むため、本来計測したい指の接触による静電容量 C_s 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、Y軸方向検出用電極を検出電極とした場合、それと隣接する電極と直交するX軸方向検出用の電極に対し、検出電極に流れている検出信号と同じ振幅・位相を持つシールド信号を印加する。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 重畠検出制御

電極結合したタッチパネルの検出処理では、電極の境界部分で感度が落ちるという問題点があった。そこで、電極結合制御によって前回結合した電極の一部を次回の検出電極と結合し重畠させて検出制御を行う重畠検出制御を行った。図6のように束ねた電極を1センサずつずらすことによって、センサの境界部分に対する接触でも感度を落とすことなく検出でき、電極を束ねることによるセンサ解像度の低下も防ぐことができる。

4. 感度評価実験

電極結合制御、及び、シールド電極制御による感度向上効果を確認するため、図7のタッチ位置に対し、次の条件下S/N比を計測した。

- (1) タッチパネルは上部に厚さ2mmのポリカを貼り付け

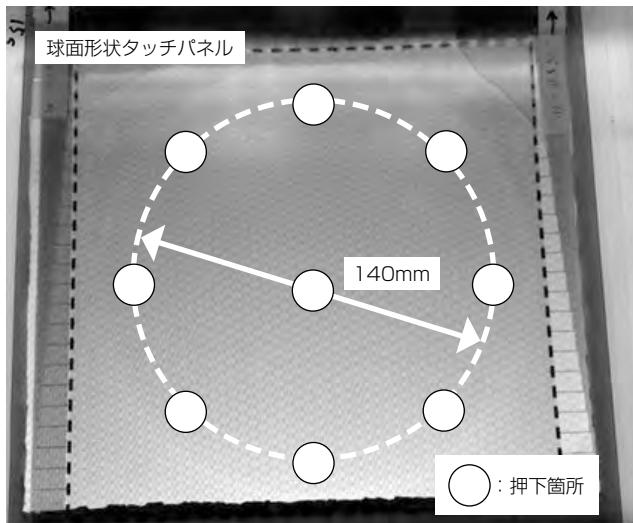


図7. 感度評価位置

た10インチ、直径350mmの球面形状である。タッチパネルを構成する電極は、ダイヤモンド型で幅7.8mmであり、図7のX(縦方向)に24本、Y(横方向)に24本並べられている。

- (2) 静電検出ICはPSoC^(注1)のCY8C24894を使用し、検出パラメータは最も感度が取れるものに設定する⁽²⁾。
- (3) 検出値は、指タッチした時のセンサ値からベースライン値(指タッチなし時のセンサ値)を減算した値である。
- (4) 評価では人差し指を使用する。
- (5) 容量検出では次の3つの方式を適用する。

方式1：電極1本使い&対向電極に対するシールド電極制御

方式2：2本結合&対向／隣接電極に対するシールド電極制御と重畠検出制御

方式3：3本結合&対向電極に対するシールド電極制御と重畠検出制御

- (6) S/N比の計算方法は次のとおりである。

① 図7の箇所で100サンプル分の検出値を取得し、その平均値をSとした。

② 各センサについてセンサ出力値100サンプル取得。その分散値 σ のうち、最大のもの($= \sigma_{\text{worst}}$)を求め、この3倍をNとした。

③ ①と②で求めたSとNから式(1)でS/N比を求めた。

$$\text{S/N比} = 20 \log \left(\frac{S}{N} \right) \quad \dots \dots \dots \text{(1)}$$

今回の提案方式について、シールド電極制御については、事前実験で有効性を確認していたため、対向電極に対するシールド電極制御を実施した方式1をベースラインとし、電極結合制御と、重畠検出制御の有効性を確認することが

表1. 感度評価結果

		従来方式			提案方式		
		方式1	方式2	方式3	方式1	方式2	方式3
S/N比平均値	X(dB)	5.8	9.9	11.5			
	Y(dB)	7.1	7.6	10.1			
S/N比最小値	X(dB)	0.9	5.9	10.1			
	Y(dB)	4.1	1.1	8.0			

実験の主目的である。さらに、電極結合制御と、シールド電極制御の感度向上効果の比較を方式2と3によって行う。

実験結果を表1に示す。表1でS/N比平均値は9箇所の平均値、S/N比最小値は最も感度の低かった箇所のS/N比である。

表1から、方式1に比べ、方式2と3は感度が向上していることが分かり、電極結合制御と、重畠検出制御の有効性を確認した。

また、方式2と3の比較では、方式2適用時にYのS/N最小値が著しく低くなかった。原因を分析したところ、タッチパネルの感度むらによって、特定のタッチ箇所で検出レンジをオーバーしたことが原因であることが分かった。今後、詳細評価を行い、各制御と感度の関係を明確化し、更なる改良を図るとともに、このような感度むらを補正するため、不感領域に対する結合本数など、検出パラメータの個別調整を検討する予定である。

(注1) PSoCは、Cypress Semiconductor Corp.の登録商標である。

5. む す び

球面形状のタッチパネルに対し、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、試作品の主観評価では2mmのポリカ越しでも十分に高い感度が得られることを確認した。感度評価実験の結果からも、従来の電極を1つずつ用いて検出する方式に比べ、S/N比が平均3.0dB以上改善されることが分かり、方式の有効性を確認した。

今後は手袋でも入力可能となるよう感度を向上させていくとともに、タッチパネルの形状を活か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。

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) Cypress' CapSense Sigma-Delta Algorithm, Cypress Semiconductor White Paper No.001-41925 (2007)
<http://www.cypress.com/?docID=3629>