



# 視 覚

## ～感性映像通信時代に向けて～

青木輝勝

東京大学先端化科学技術研究センター



五感情報のうち、視覚が人間に与える影響は最も大きく、ノンバーバル情報のうちの65～80%にも達すると言われている[1]。このため、視覚は他の感覚と比較して古くより多くの研究がなされている。しかしながら、今日なお多くの研究課題が山積しており、それぞれの分野で盛んに研究開発が進められている。

映像通信技術は、一般に、映像入力技術(イメージセンシング技術)、映像処理技術(映像符号化技術)、映像出力技術(映像表示技術)に大別できるが、本稿では、これらのうち、五感情報通信の立場、特に感性映像通信の観点から、映像符号化技術、映像表示技術について現状と今後の展望をまとめる。

### 1. 感性映像符号化技術

近年インターネット等を用いた映像配信サービスに盛んに行われるようになってきた。この背景としてISO標準のMPEG(Moving Picture Engineering Group)-1、2、4、ITU-T標準のH261、H263、H.26L、さらにはRealvideo等の映像符号化技術の進展、そしてADSL、CATVインターネット等のブロードバンドアクセス技術の一般化が大きな役割を果たしている。

映像符号化技術としてMPEGを例に挙げると、

- ・ MPEG-1：CDに代表される蓄積メディア用符号化技術。1/200の圧縮率を実現。
- ・ MPEG-2：デジタル放送、通信、DVD等の用途を想定した符号化技術。1/10～1/20の圧縮率。

- ・ MPEG-4：オブジェクト符号化を含む符号化方式。携帯端末、インターネットコンテンツ配信への用途。

の応用を前提とし、汎用符号化ながらある程度応用目標を明確化することによりそれぞれ大きな成功を収めているのは周知の通りである。しかし、これらの標準では、応用が異なるにも関わらずその基本技術としていずれもDCT(Discrete Cosine Transformation)技術およびDCT+MC(Motion Compensation)技術を用いている。これにより、映像コンテンツの扱いを大幅に容易化することに成功したが、その反面、MPEG-1誕生以降、多くの研究開発がなされているのにも関わらず最近10年間の進歩は「DCT+MC」の効率を大きく超える高能率、汎用映像符号化方式が出現していないのも事実である。この間、ISO、ITU-T等の国際標準化会合に提案されては消えた符号化方式は、それこそ枚挙にいとまがないが、残念ながら、いずれの方式も特定の映像に対してのみ効果的な高能率或いは高品質の符号化方式に過ぎなかった。

DCTやWaveletをはじめこれまでの符号化技術では、人間の視覚に影響が少ないと思われる高周波成分をカットして符号量圧縮を行うことにより、映像劣化が少なくかつ大幅な情報量圧縮を実現してきたが、今後の符号化を考えるにあたっては、これとは全く異なるアプローチが必要になると予想される。表1にANSIで規格化されている映像評価項目[2]を示すが、例えば"Edge Energy Difference"については、多くの場合、高周波成分のカッ

表1 ANSIT1.801.03の画像評価項目

	Comments
Motion energy difference	Added motion energy indicates error blocks, noise. Lost motion energy indicates jerkiness.
Repcated frames	Indicates jerkiness.
Edge energy diffrence	Added edge energy indicates edge noise, blockkiness and noise. Lost edge energy indicates blurriness.
Horizontal/Vertical (HV) to non-HV edge energy ratio	Indicates amount of blockiness.
Spatial frequencies difference	Added spatial frequencies indicates edge noise, blockiness and noise. Lost spatial frequencies indicates blurriness.

トによって生じる画像劣化であり、この部分を効率良く保持するためには新たな符号化アルゴリズムの開発が不可欠であるからである。

一方、上述してきたDCT、Wavelet等の「波形符号化技術」の他にも「対象物符号化(オブジェクト符号化)」の研究開発もこれまで盛んに行われている。この対象物符号化はさらに、

- (1)認識論符号化：画像オブジェクトの形状認識を利用した符号化方式
- (2)意味論符号化：画像オブジェクトの意味理解を利用した符号化方式

に分類されるが、特に(2)に関しては、画像理解の研究が広く実用レベルには達していないこともあり、汎用的な実用利用にはまだまだ多くの課題を残している。

さて、以上のような背景のもと、今後の映像符号化はどのような方向に向かってゆくのかを考えると、やはり最重要キーワードのひとつとして「感性」が挙げられるであろう。画像解像度についてはさらなる高精細化の方向に進んでおり、最終的には恐らく現在の4M画素から100Mクラスの画素数まで進化すると考えられ、また、これに呼応して、近年、心理学・生理学的側面からも研究が進み、その結果、人間が知覚できるか否かに関わらず高周波成分を含む映像や超高解像度の映像が人間の感性に与える影響が大きいことが報告されているからである。これらの状況から考えても、今後はこのような感性情報まで含めた符号化方式の開発が重要になると言えよう。

上述した対象物符号化技術の延長で言えば、上述した

(1)認識論符号化→(2)意味論符号化の流れを基礎に、新たに、

(3)感性符号化：画像オブジェクトの形状認識、意味理解をもとに人間の感性に影響を与える情報を完全保持する符号化方式

と定義することができよう。

例えば「サッカー中継」を取り上げると、「認識論符号化」は、各オブジェクト(選手、観客、ゴール、ボール等)の形状に着目して符号化を行うことに相当し、「意味論符号化」は、選手オブジェクト、ボールオブジェクトは品質重視、観客オブジェクト、ゴールオブジェクトは効率重視の符号化を行うことに相当する。一方「感性符号化」では、選手オブジェクトのうち、躍動感を感じさせる部分を抽出し、特にその部分を完全保持した符号化を行うということになる。

このような感性符号化技術を開発するためには、現在の意味論符号化技術のさらなる進展とともに、映像が人間に与える心理学・生理学的特性のより正確な解明が必要不可欠であり、その実現は決して易しいものではないが、符号化技術の最終ゴールがこのような感性符号化になることは恐らく間違いないであろう。

2. 感性映像表示技術

映像表示技術は、用途に応じて非常に様々な技術が存在するが、感性映像通信との関連で言えば、

- (1)高精細ディスプレイ
- (2)3Dディスプレイ

## (3)その他のディスプレイ (CAVE、HMD、視線一致型)

などが挙げられる。

(1)の高精細ディスプレイに関しては、例えば3840×2070 (HDTV品質の4倍の解像度)のディスプレイ開発が報告されている[3]。このディスプレイでは、複数台のプロジェクタから投射した画像を画素単位でスクリーン上で合成する手法(重畳投射型)を開発している。具体的には、プロジェクタに使用しているTFT-LCDパネルが開口部と遮光部から構成されているという特徴を利用して、スクリーン上で遮光部に相当する投射領域に他のプロジェクタの開口部からの投射領域を精度良く重ねることによって超高精細化が実現されている。

また、(2)の3Dディスプレイに関しては、一般にメガネ装着型、メガネ不要型に分類され、メガネ不要型では特にパララックスバリア方式、レンチキュラレンズ方式がよく知られている。3Dディスプレイの原理は両眼視差に基づき立体感を表現することであるが、これらの方式では、一般に大画面化が困難であり、この点が技術的には今後最も重要な課題となっている[5]。また、近年ホログラフィに関する研究[6]も盛んに行われている。従来の3Dディスプレイが両眼視差のみを使用していたために自然な3D像を表示できず長時間の使用が困難であったのに対し、ホログラフィでは両眼視差に加え、眼の輻輳と調節、運動視差等を支援しており、より自然な立体感を得ることができる。現在のところホログラフィ技術はデバイス開発等の面でまだ不十分ではあるが、原理的に理想的な3Dディスプレイであり今後の実用化研究に期待が集まっている。

一方、(3)に関しては、代表的な高臨場感ディスプレイとして、IPD (Immersive Projection Display)、HMD (Head Mounted Display) がよく知られている。

IPDは、主に視野の広がり(視野角)、奥行き感(立体感)などの支援を目的に開発された没入型ディスプレイであり、国内では5面ディスプレイ型のCABIN[6]がよく知られている。

また、IPDと同様の目的を実現する技術としてHMDも開発されている。HMDはもともと1968年に頭に装着する3次元ディスプレイとして発表され、その後VR技術の発達とともに没入感を要求されるVRの視覚ディスプレイとして様々なシステムが研究開発されてきた。現在のHMDはおおむね両眼視差を利用して立体視を実現する

ものが主流で、これまで島津製作所、オリンパス、Virtual Research社、Kaiser Electro Optics社等で開発が行われているが、一方、立体視を行わず、両眼に同一映像を表示して見かけ上大画面映像を実現するHMDもいくつか市販されており、オリンパスのEye-Trek、ソニーのGlasstronなどがよく知られている。

また、ディスプレイの利用を人間同士のコミュニケーションを前提とした場合には、視線一致をはじめとする従来のディスプレイとは異なる様々な要求条件が生じる。具体的には、高臨場感ビデオ会議システムの分野を中心として、これまで

- (1)視線一致
- (2)実物大映像
- (3)ゲイズ・アウェアネス
- (4)画面ポインティング

等の要求条件が挙げられている。既存ビデオ会議システムでは通常図1に示すように、ディスプレイ上部にカメラが配置されることが多いが、これではカメラの撮影方向と会議参加者の視線方向が一致しないため、会議参加者同士の視線が一致できない。また、(3)ゲイズ・アウェアネスとは、3名以上の会議において会議参加者の視線方向(誰を見ているのか)がお互いにわかる状態を指す語であるが、これも既存のほとんどのシステムでは実現できていない。例えば、図2に示すビデオ会議画面の場合、一見すると会議参加者D氏がC氏を見ているように見えるが、実際にはD氏が単にディスプレイから目を離して横を向いているに過ぎない。また、(4)画面ポインティングに関しては、ビデオ会議に特化した議論であるが、自端末のディスプレイに表示された会議資料を指先でポインティングした時に相手にそれが伝わるか否かということである。現在のビデオ会議システムではほとんどの場合、図3に示すように会議相手が画面をポインティングしても指先がカメラの視野角に入らないためポインティング位置を相手に伝えることができないため、何らかの方策が必要である。

以上のように対面環境と同様の自然なコミュニケーションを実現するためにはこれら4要求条件の支援は極めて重要である。

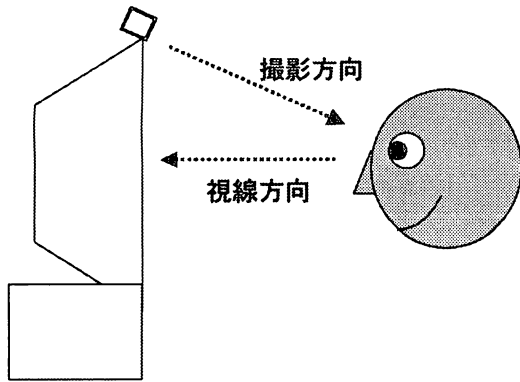


図1 視線一致の困難性

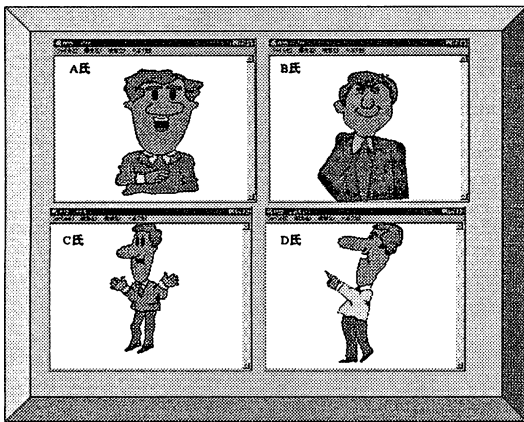


図2 ゲイズ・アウェアネスの困難性

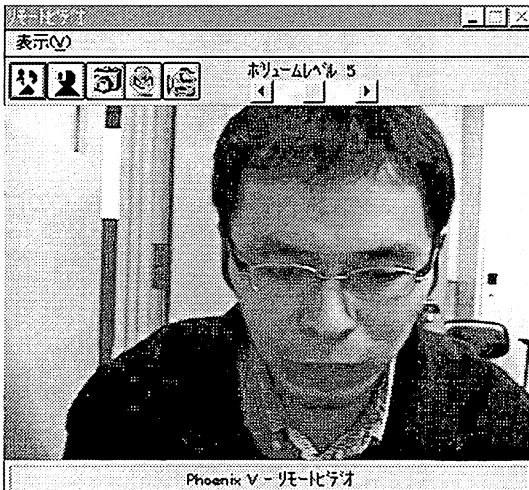


図3 画面ポインティングの困難性

これまで(1)視線一致の実現手法としては、ハーフミラーを用いたシステムが最も広く知られている。しかし、ハーフミラーを用いると、カメラへの光量が不足したり、画面が奥まって見えるなどの問題点があり、残念ながら十分な臨場感は得られない。

また、ハーフミラー型以外では、透明／散乱を交互に

繰り返す液晶スクリーンを用いたシステム[7]、視線を検出し仮想空間上での視線一致を実現するシステム[8]、小型ディスプレイを複数用いディスプレイ1つあたり会議参加者1人を映し出すシステム[9]などの研究例がある。しかし、これらの手法を用いた場合、視線一致と実物大映像の双方を同時に実現することは一般に非常に困難である。

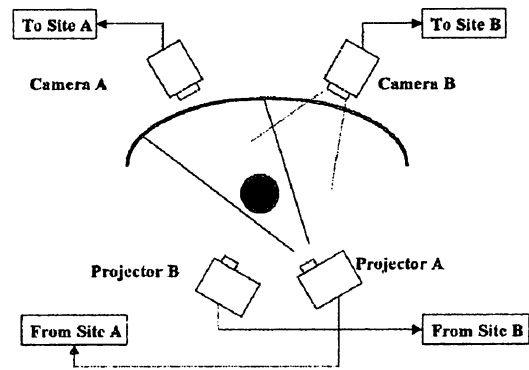


図4 MAJICシステムの原理

また、(1)視線一致、(2)実物大映像を同時に満たすシステムとして、特殊フィルムを用いたMAJICシステム[10]がある。MAJICシステムは、さらに(3)ゲイズ・アウェアネスまで支援できるシステムであるが、原理的にディスプレイに近づくと自分の影がディスプレイに映ってしまうため、画面ポインティングについては支援できない。また、光量の不足による画像品質の劣化が見られることも臨場感を損なう一要因となっている。

一方、特殊なホログラムスクリーンを用いた文殊の知恵システム[11]がある。これは世界ではじめて(1)視線一致、(2)実物大映像、(3)ゲイズ・アウェアネス、(4)画面ポインティングをすべて満足するシステムとして現在注目を集めている(図5)。

さらに、文殊の知恵システムが映像品質、端末サイズ、通信帯域等の点で実験レベルであったのに対し、上記の4要求条件をすべて満足しつつ、実用化に向けて映像品質、端末サイズ、通信帯域等の点で大幅な改善が図られたシステムとしてCCC (Collaborative Content Creation) プロジェクト[12]の以心伝心システムがある。本プロジェクトはその名の通り、遠隔協調3DCG制作を前提としたビデオ会議システムであり、約45日間にわたりプロのCGクリエイターによる実証実験が行われた結果、その臨場感、使用感(操作性)について高い評価が得られ

ている。



図5 文殊の知恵システム

### 3. まとめ

本稿では、五感情報通信の中の「視覚」関連技術について、特に感性映像通信の立場から、映像符号化技術、映像表示技術について現状と今後の展望をまとめた。「感性」の定義はいまだに明確ではなく、従って感性映像通信技術も今後様々な展開が考えられるが、少なくとも「人間」を無視した映像通信であってはならないことには異論がないであろう。本稿で述べた人間の感性に影響を与える符号化方式、人間同士の自然なコミュニケーションを実現する映像表示方式以外にも「人間」に重点をおいた様々な映像通信技術が存在するが、これらを含めた感性映像通信技術全体のさらなる発展が図られ、私達の日常生活が「感性」面でより豊かになることを心より期待する次第である。

#### 文献

- [1] Birdwhistell, "Kinesics and Context", Univ. of Pennsylvania Press, 1970.
- [2] ANSI T1.801.03, "Digital Transport of One-way Video Signals - Parameters for Objective Performance Assistance", 1996.
- [3] 中沢憲二、上平員丈、酒井重信、"超高精細大画面表示システム"、映像情報メディア学会誌、Vol.52、No.7、1998.

- [4] 松村進、"裸眼立体ディスプレイ"、映像情報メディア学会誌 vol.54, No.3, 2000.
- [5] 高橋秀也、志水英二、"電子ホログラフィ"、映像情報メディア学会誌 vol.54, No.3, 2000.
- [6] <http://www.ihl.t.u-tokyo.ac.jp/Projects/CABIN/index-j.html>
- [7] S. Shiwa and K. Nakazawa, "Eye Contact Display Technologies for Visual Telecommunications", NTT REVIEW, Vol.5, No.2, 1993.
- [8] H. Takemura and F. Kishino, "Cooperative Work Environment Using Visual Workspace", Proc. of CSCW '92, 1992.
- [9] A. J. Sellen, "Speech Patterns in Video Mediated Conversations, Proc. of CHI '92, 1992.
- [10] Ken-ichi Kameyama, Koichi Ohthomi and Yukio Fukai, Interactive volume scanning 3-D display with an optical relay system and multidimensional input devices, SPIE - The International Society for Optical Engineering, Vol. 1915, 1993.
- [11] Terumasa AOKI, Kustarto WIDOYO, Nobuki SAKAMOTO and Hiroshi Yasuda, "Next-Generation Videoconference System with Sense of Reality", International Conference on Computer Communications '99 (ICCC'99), 1999.
- [12] <http://www.sohbi.co.jp/c3project/>または <http://www.c3project.com>

#### 【略歴】

青木輝勝 (AOKI Terukatsu)

平5 東大・工・電子工学卒。平10 同大大学院博士課程了。現在、東大先端研助手。工博。高臨場感ビデオ会議システム、高能率画像符号化方式、コンテンツ保護技術、高速大容量IPルーティング交換方式等に関する研究に従事。平6 (財)電気電子情報通信振興財団猪瀬学術奨励賞、1999 IEEE Globecom Student Award (共著)、平12 (財)電気通信普及財団論文賞 (共著) 受賞。