

マルチメディア通信に向けた将来ビジョンと技術課題

高野陸男

NTTヒューマンインタフェース研究所

マルチメディア通信の実現に向けて将来ビジョンと技術課題を述べる。まず、メディア符号化とネットワーク性能との関係を示し、高圧縮符号化技術とメディア構造化技術が、経済化の観点から重要であることを述べる。構造化の例として、ISDNカラオケ、キャッチビジョンについて説明する。次に、シームレス化技術が人間の生産活動支援の観点から重要であることを述べる。例として、クリアボード、多視線一致撮像表示システム、音像定位通信システムについて説明する。最後に、マルチメディア通信により変革がもたらされることが期待される実作業として、協調作業、遠隔モニタリング、情報提供サービスについて考察する。

FUTURE VISION AND TECHNOLOGY TOWARD MULTIMEDIA COMMUNICATION

Rikuo Takano

NTT Human Interface Laboratories

1-2356 Take, Yokosuka 238-03, JAPAN

This paper discusses future vision and technology toward multimedia communication. First, the relation between media coding and network is shown. High efficiency coding techniques and structure techniques are important for economical multimedia communication systems. As the examples of systems using structured media, the Karaoke system through ISDN and the event driven visual communication system are shown. Next, seamless multimedia communication systems such as the ClearBoard, the multiple-eye-contact teleconference system and the stereophonic sound image localization system are shown. Finally, cooperative work, monitoring and information providing are discussed as the fields innovated by the multimedia communication systems.

1. まえがき

コンピュータ分野を中心にメディアのデジタル化が進み、複数メディアを同一のコンピュータで処理できるようになってきた。これに呼応してマルチメディアシステムの開発が活発化している。

一方、通信分野でも、ISDNを中心としたネットワークのデジタル化が進められている。その結果、デジタルという同じ土俵上で、コンピュータとネットワークを融合することが可能になり、音声、映像等の複数の表現メディアをデジタルネットワークを通して通信するマルチメディア通信への期待が高まっている[1]。

さらに、映像符号化の世界では、映像ダウンサイジング現象があらわれ映像が扱いやすいメディアになってきた。例えば、7年前を1とすると、映像コーデックの大きさ、価格、通信速度の比を掛け合わせたものは一万分の一になっている。このように、マルチメディア通信は現実味を帯びたものになっている。

本稿では、まず、音声、画像について、その品質とその通信に必要とされる通信速度との関係を示し、どのネットワークでどのようなメディアを扱えるかを議論する。そして、経済化の観点からサービスごとに本質的に必要なメディアの種類と品質を検討することが重要であることを述べる。

次に、高圧縮符号化に加えて、メディアの構造化が経済化に重要であることを述べる。すなわち、一つのメディアの構造に着目すると、サービスごとに必要とされる品質を満たす構造が存在し、その構造で通信することにより、高速・低価格のマルチメディア通信を実現できることを示す。

次に、人間の生産性を向上させるという観点から、メディアのシームレス化が重要であることを述べる。シームレス化により、人間と人間のスムーズなコミュニケーションが実現できる。

最後に、実作業に変革もたらすという観点から、マルチメディア通信を考察して、協調作業の支援、正確な情報伝達、音声情報中心のサービスについて述べる。

2. ネットワークとメディアの符号化

マルチメディア通信において、ネットワーク性能とメディア符号化技術によって、通信できるメディアの品質が変わる。主なメディア符号化に要求される通信速度は表1に示すとおりであるので、主なネットワークで扱えるメディアは表2のようになる。従って、多様なメディアを通信するためには、高速・広帯域のネットワークであるB-ISDNが必要となる。

多様なメディアを通信するためのB-ISDNのキ

表1 主なメディアの符号化

メディア	符号化名	品質	通信速度	方式
動 画	H.261	狭帯域テレビ電話/会議 QCIF(144x176) CIF(288x352)	64~1920Kbps	DCT,MCフレーム間予測
	MPEGI	VTR等の蓄積メディア	1~1.5Mbps	DCT,MCフレーム間双方向予測
	MPEGII	テレビ放送品質 NTSC(480x720) HDTV品質 (1035x1920)	5Mbps以上	審議中(1993年未報告案予定)
静止画	JPEG	テレビ印刷品質	64Kbps	適応型DCT
音 声	G.711	3.4KHzアナログ電話	64Kbps	非圧縮
	G.722	7KHzAM放送	48/58/64 Kbps	SB-ADPCM
	G.721	3.4KHzアナログ電話	32Kbps	ADPCM
	G.728	3.4KHzアナログ電話	16Kbps	LD-CELP
	MPEG/AUDIO	48KHz音響	64K~192Kbps/ チャンネル	NUSIC AM

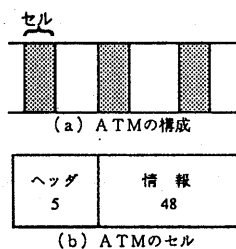


図1 ATMの構成

表2 ネットワークとメディアの関係

通信手段	通信速度	用途	代表的通信メディア
電話	100~10kbps	パソコン通信	文字 図形 データ
パケット交換 N-ISDN	10k~1Mbps	テレビ電話 テレビ会議	静止画 低品質動画 高速データ NTSC動画 印刷情報
LAN	1M~100Mbps	映像中継	
B-ISDN ギガビットLAN 光ファイバ伝送	100Mbps以上	高速系間通信	HDTV動画 大容量データ

ーとなる通信モードが非同期転送モード

(ATM:Asynchronous Transfer Mode)である[2]。ATMは情報を固定長のセルと呼ばれる単位に分割して送る。一つのセルは図1に示すように送り先を表すヘッダ(5バイト)と送る情報(48バイト)からなる。従って、通信速度の違いは、単位時間あたりのセル到着数の違いとして現われ、各通信速度に固有なハードウェアを必要としないため、多様なメディアを統一的に扱うことができる。

B-ISDNを用いれば、HDTV品質の画像と48KHz音響品質の音声を通信できるが、サービスによってはN-ISDNやLANで十分なものもある。従って、ユーザーの満足度を見据え、本質的に必要なメディアの種類と品質を検討することにより、オーバーメディア、オーバークオリティを回避し、高圧縮符号化により経済化をはかることが重要である。

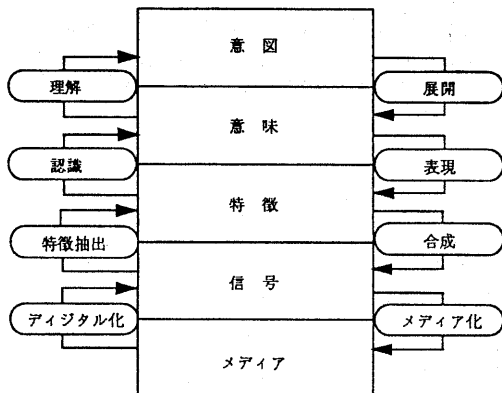


図2 メディアの構造化

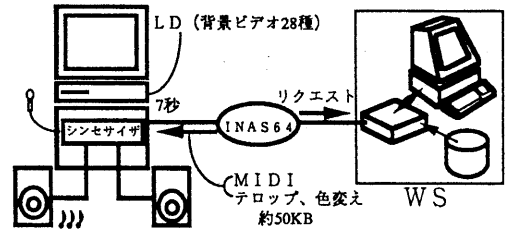


図3 ISDNカラオケ

3. メディアの構造化

経済化を実現する別の手段としてメディアの構造化がある。メディアの構造化とは、メディアをデジタル化した信号に変換したのち、さらに、図2に示すように、特徴抽出、認識、理解等の処理により、特徴、意味、意図等の高次な構造に変換しておくことである。構造化には様々なレベルがあり、これらの処理は、必ずしも自動処理でなくてもよい。以下では構造化の例を二つ示す。

(1) ISDNカラオケ[3]

タイトーが開発したX2000カラオケ通信システムも構造化によってN-ISDNで十分な性能を達成したシステムである。このシステムでは、図3に示すようにカラオケの基本背景動画は端末側に蓄積しておき、MIDI形式の音楽データ、テロップ(テキスト)とテロップの色変え情報の約50KBをN-ISDNを通してワークステーションから端末に送る。MIDIは楽譜情報を楽器間や楽器とパソコン間でやりとりするためのきまりである。MIDI形式で送られた音楽データはシンセサイザ等で様々な楽器の演奏として再生することができる。このように、音楽データを直接送るかわりに、意味レベルの表現に構造化して送ることにより、5分間の曲を演奏する場合ワークステーションと端末を5分間接続する必要がなく、7秒ですむ。このように、高速で低価格のカラオケサービスが実現されている。

(2) キャッチビジョン[4]

キャッチビジョンは、図4に示すようにTVカメラから入力された映像の内容を画像認識技術で

抽出して、人や車の出現などの意味のあるシーンを
含むフレームのみを離れた場所にあるモニタ画
面に送信するものである。画像認識技術により動
画像データを意味情報に構造化することにより、
必要な情報を保存したまま、データ量を大幅に削
減して低価格で、遠隔地の交通、人流の監視、調
査を行うことができる。

以上の例からわかるように、本質的な情報だけ
を通信し、付帯的な情報は低品質で通信するか端
末に蓄積しておくことにより、経済的でユーザー
の要求性能を満足するマルチメディア通信システ
ムを構築することができる。

4. マルチメディア通信のシームレス化

マルチメディアの目指すものは何であろうか？
コンピュータやネットワークの目指したものは人
間の生産性の向上であるが、コンピュータやネッ
トワークにマルチメディアをのせることにより、
コンピュータやネットワークが人間にとって使い
やすいものになり生産性の向上を加速できると考
えられる。

生産性を向上させるためには、メディアのシ
ームレス化による人間と人間のスムーズなコミュニ
ケーションの実現が必要であると考えられる。一言で
シームレス化といっても、様々なレベルがある。
ここでは、まず、メディアがマルチであることを
人間に意識させないようにさせる意味でのシーム
レス化により、離れた場所にいる人間が連続した
空間でつながっているように視覚的に感じられる
ことを実現した例として、クリアボードと多視線
一致撮像表示システムとを説明する。また、聴覚

的に連続空間を感じられることを実現した例とし
て音像定位通信システムを説明する。

(1) クリアボード[5]

クリアボードは、マルチメディア通信を用い
て、設計などの考える作業を、離れた場所にいる
人間が一緒にできるようにする実験装置である。
そのポイントは”双方の描画情報の重ね合わせ”
の実現にある。

一言で言えば、ホワイトボードの機能とTV電
話の機能を組み合わせたものである。スクリーン
として、通常のディスプレイのかわりに透明なガ
ラス板を用い、それをハーフミラーにしておく。
これにより、図5に示すように、相手の描いた情
報と相手の顔画像を、自分のスクリーンに映し出
すことができるとともに、自分の描いた情報も自
分のスクリーンに描くことができる。従って、あ
たかも同じホワイトボードに描いているかのよう
に、打ち合わせや討論を行える。

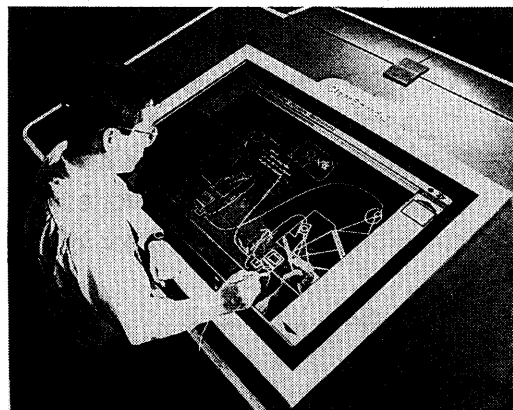


図5 クリアボードの外観

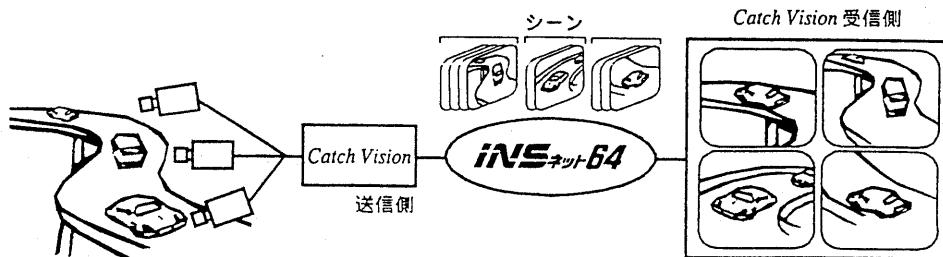


図4 キャッチビジョン

この装置の原理は図6に示す。スクリーンの上にあるカメラが、描いた情報を撮影する。このカメラは、同時にスクリーン表面のハーフミラーに反射した人間の顔も撮影する。これらの映像は相手に送られて、スクリーンの後ろにあるビデオプロジェクターからスクリーンに投影される。

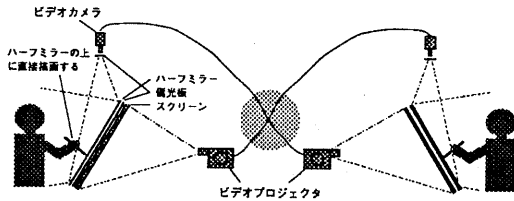


図6 クリアボードの原理

(2) 多視線一致撮像表示システム[6]

複数人が出席する会議において、出席者ごとに見える風景は異なる。すなわち、ある出席者から見るとアイコンタクトがとれている出席者も、別の出席者から見ると、別の方向を向いているように見える。これが物理的に連続した空間における会議の見え方である。この連続空間の見え方を通信会議システムにおいて実現するのが多視線一致撮像表示システムである。

図7は2人対2人の通信を上方から見た図である。それぞれの出席者に対向して設置したカメラ

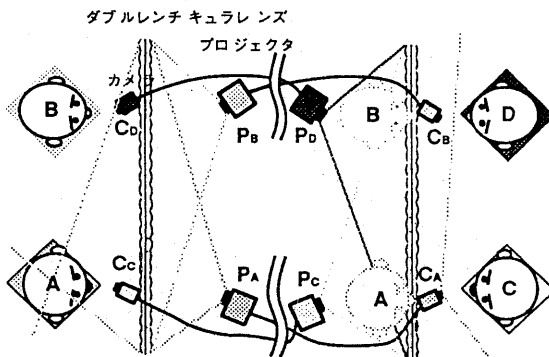


図7 多視線一致の原理

からの撮映像は、通信回線を介して通信相手のプロジェクターからスクリーン上に投影される。スクリーンはダブルレンチキュラーレンズを用いて構成されているので、スクリーンの背後に設置されたプロジェクターからの投影光は出席者の位置へ集光するように作用する。すなわち、出席者Aが出席者Dを向いて話しているとき、出席者Aの画像はカメラC_Dから撮像されプロジェクターP_Dに投影され、出席者Dから出席者Aが自分の方向を向いているように見える。出席者Cからは、カメラC_Cが捉えた出席者Aの横を向いた画像が見えるので、出席者Aは出席者Dを向いているように見える。図8は実際に観察される画像である。



(a)左着席者の観察する画像



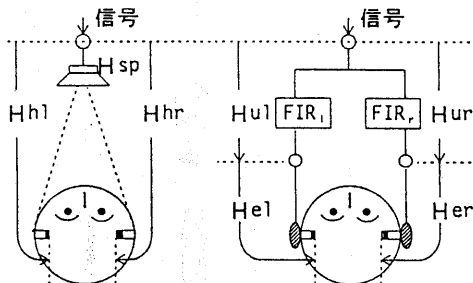
(b)右着席者の観察する画像

図8 観察される画像

(3) 音像定位通信システム[7]

このシステムは、多地点を通信回線で結び、ステレオイヤホンを用いて、離れたところにいる話

者の音像を近くに定位して、あたかも同一会議室で議論しているように聴こえるようにしたものである。これは、図9に示すように聴こえる音声の伝達関数（頭部伝達関数）が音像を定位する伝達関数（頭外定位伝達関数）と外耳道伝達関数の積であることを用いて、頭外定位伝達関数を話者ごとに異なった位置に音像が定位するように変化させることにより実現されている。



■：小型マイク，FIR：デジタルフィルタ，⊙：イヤホン
 Hh：頭部伝達関数，Hu：頭外定位伝達関数
 He：外耳道伝達関数，Hsp：拡声器伝達関数

図9 音像定位の原理

以上のように、離れた場所にいる人間の間をマルチメディア通信を用いて、感覚的に連続した空間でつなぎ、人間と人間のコミュニケーションをスムーズにする技術が開発されつつある。

5. 実作業の変革に向けたマルチメディア通信

前章で述べたように、マルチメディアの目指すものは生産性の向上であると考えられる。従って、生産現場等の実作業に役立つマルチメディア通信システムを開発して、実作業に変革をもたらすことが重要である。ここでは、マルチメディア通信システムが変革をもたらすことが期待される三つの実作業について述べる。

第1の実作業は、人間と人間のコミュニケーションが重要な役割をはたす協調作業（コラボレーション）である。例えば、複数人からなるチームによる新製品の設計やソフトウェアの作成などが

ある。協調活動を支援するマルチメディア通信システムの例を二つ説明する。

(1) LAN用映像通信会議システム[8]

このシステムは、通常のワークステーションやパソコンに図10に示すLAN用映像通信モジュールを接続することにより、データ、映像、音声の通信を可能にしたものである。これにより、図11に示すように、LANを介して、ワークステーション、パソコン間のウィンドウで映像・音声の会議ができる。また、同時に、別のウィンドウでは、プログラミング、ワードプロセッシングや作図ができるので、共同編集、共同設計、共同作成などの共同作業を容易に行える。

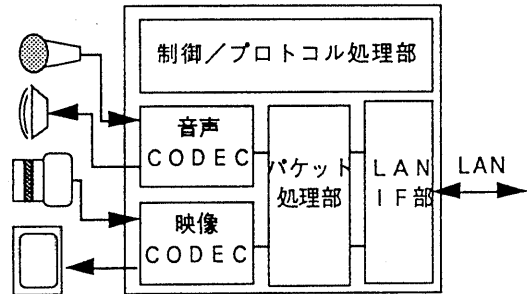


図10 LAN用映像通信モジュールの構成



図11 LAN用映像通信システムの画面

(2) チームワークステーション[9]

このシステムは、自分の机の上の映像を半透明状に合成して画面上に表示することで、遠方にいる相手とあたかも同じ机で作業するような環境を実現している。書き込みや手による指示などを、自分の机の上にある印刷物はもちろん、相手の机の上にある印刷物にも行える。図12に示すように、机の上の映像を撮像するカメラに加えて、人間の顔を撮るカメラを設置することにより、顔画像を表示することもできる。この図では、生徒が書いている習字を先生がリアルタイムに添削している場面である。

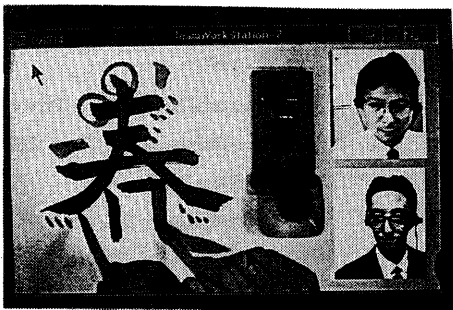


図12 チームワークステーションの画面

第2の実作業は、遠隔モニタリングである。この実作業では、必要な情報を正確に伝えることが要求される。例えば、最近、専門医の不足のため、専門医のいない病院といる病院を映像通信システムでつなぎ、専門医のいない病院でも遠隔診断を受けられるようにしたシステムが報告されている。そのとき送らなければならない情報の本質は、送られた映像の忠実性であると思う。例えば、通信システムを通すことにより色が変色すれば誤った診断がくだされることが予想される。そこで、カラー映像を正確に伝える技術が必要になる。このための技術として、入出力機器の色補正技術[10]や照明の影響を取り除いた物体色推定技術[11]が研究されている。

第3の実作業は、音声情報を中心とした情報提供サービスである。なぜ、音声情報が中心である

かといえば、データ量の少なさによる経済性もさることながら、意味的な情報量は映像より音声の方が多いためである。まさに、百見は一聞にしかずである。このように、音声情報を主体に考えると、一気に現実味を帯びる情報提供サービスも少くない。例えば、ネットワークキャスティング [12]というサービスについて考えてみよう。このサービスは、現在の情報提供メディアの空白になっている情報提供メディアを提供しようとするサービスである。図13に示すように、情報提供メディアを情報の配布時間と配布対象者数により整理すると、配布時間が数時間から数十分で配布対象者数が数百人から数万人までの、情報提供メディアが存在しない。そこで、B-ISDNを用いて、オンデマンドで動画像を配布する娯楽向きのメディアとしてTVネットワークキャスティングを、オンデマンドでテキストを中心とした情報提供するメディアとしてマルチメディアネットワークキャスティングが提案されている。マルチメディア通信の中心は音声情報であるという観点から、このサービスを見直すと、音声を中心とした娯楽、情報提供向きのネットワークキャスティング（これを”ネットワークキャスティング-N”と呼ぶ）が考えられる。このサービスは、N-ISDNで行うことができる。また、最近、ラジオを用いたパーソナリティによる通信販売が盛んに行われていることから、これをネットワークキャスティング-Nにのせれば、物販サービスの分野も拓かれる可

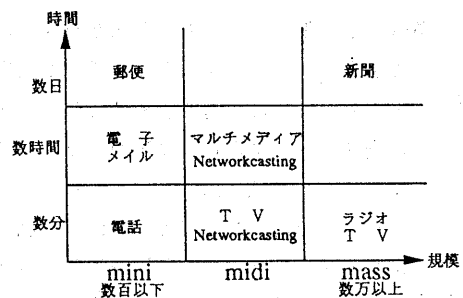


図13 情報の配布

能性がある。

6. むすび

マルチメディア通信の将来ビジョンと技術課題について展望した。すなわち、技術的には、メディアの構造化技術、マルチメディア通信のシームレス化技術がキーテクノロジーであり、それにより、2次、3次産業を中心とした実作業の変革に寄与する必要があることを述べた。

ここで、取り上げなかったキーテクノロジーとして、メディア間変換技術、セキュリティ技術がある。メディア間変換技術とは、音声をテキストに変換する技術やその逆の変換技術等である。一般に、メディア間変換のためには高度な認識技術が必要とされるために、実用レベルに到達しているメディア間変換技術は少ない。しかし、メディア間変換技術は、どこでも、だれでもがマルチメディア通信の恩恵を受けることを実現するためには不可欠な技術である。セキュリティ技術として、通信の内容を暗号化して盗まれないようにする技術に加えて、通信される各種情報のコピー防止機能が重要になる。この機能をマルチメディア通信が持つことにより、安全に、安心して情報を享受することができ、ひいては豊富な情報を扱うことができることになろう。

通信システムとして、電話、ファクシミリ、TV会議等が開発され、人間の生産活動は確実に効率化された。しかし、反面、これらのシステムにより、さら精査な議論のための出張が増えている。通信は交通量を増やしているのである。マルチメディア通信では、離れている人間があたかも同じ場所にいるようにコミュニケーションできる環境を実現して、できるだけ通信で済ませられるようにしたい。交通を代替する通信に近づきたいものである。なぜなら、交通を減らすことは省エネルギー化になり地球温暖化を減速し大気汚染を低減することになるからである。地球環境保護は21世紀の我々の課題であり通信の果たす役割も大きいと考える。

参考文献

- [1] W. Partridge: 無線通信、HOBBIT、ペン入力を統合するパーソナル情報機器の市場形成シナリオ、液晶ディスプレイ・セミナー、日経BP社、1993.
- [2] 濃沼、高橋: ATMノードシステム技術、NTT R&D, Vol.42, No.3, pp.275-282, 1993.
- [3] 渡辺博則: 花開くか、ISDN利用のマルチメディア情報サービス、日経コミュニケーション、pp.72-79, 1993.1.18.
- [4] 安達、小池、石井: 映像抽出通信の検討、1993年電子情報通信学会春季大会、D-240.
- [5] H. Ishii, M. Kobayashi and J. Grudin: Integration of interpersonal space and shared workspace: clearboard design and experiments, in Proc. ACM CSCW'92, pp.33-42, 1992.
- [6] K. Nakazawa, S. Shiwa and S. Ichinose: Private display method for teleconferences, in Proc. JAPAN DISPLAY'92, pp.395-398, 1992.
- [7] 島田、林: 両耳イヤホンを用いた頭外音像定位ステレオ方式、1992電子上方通信学会春季大会、SA-7-5.
- [8] 田尻、阪谷、秦泉寺、久保、西村、吉村: LAN用映像通信会議システムPMTC/LAN、TV学会技術報告、IDY93-51, 1993.
- [9] H. Ishii and N. Miyake: Toward an open shared workspace: computer and video fusion approach of teamworkstation, Comm. of ACM, Vol.34, No.12, pp.37-50, 1991.
- [10] M. Inoue and M. Kosugi: Enhancement of natural color reproduction on CRTs, SID'92 Digest, Vol.23, pp.198-201, 1992.
- [11] 川村、本郷、鎧沢: ハイライトを持つ画像からの光源色の推定、TV学会技術報告、VAI92-38, 1992.
- [12] 釜江尚彦: VI&Pを目指すヒューマンインタフェース技術の動向、NTT R&D, Vol.41, No.2, pp.141-148, 1992.