

## 顔チャット™システム

"Kao Chat" System

井田 孝

IDA Takashi

竹島 秀則

TAKESHIMA Hidenori

堀 修

HORI Osamu

パソコンや携帯電話にカメラを接続することで顔画像を伝送し合い、互いの顔を見ながら会話できる環境が整いつつある。今回、実時間・全自動で参加者の姿を輪郭に沿って切り出して伝送し、それらを一つの画像に合成するシステムを開発した。これにより、チャットシステムにおいて合成画像を共有することで、参加者の一体感や親近感が得られるコミュニケーション空間を提供する。このシステムの処理量や伝送符号量は少なく、現状の普及型パソコンと公衆通信路で実用可能である。

Communication systems in which facial images are transmitted using a camera connected to a PC or cellular phone are becoming popular. Users of such systems are able to converse while viewing the faces of those with whom they are conversing.

Toshiba has developed an image composing system in which the images of participants are extracted along contours, transmitted, and automatically composed as an image in real time. By sharing the composed image via a chat system, an enriched communication experience characterized by feelings of unity and affinity is achieved. The proposed system is practical for use with current consumer PCs and networks because the required computational load is small and the required transmission bit rate is low.

## 1 まえがき

インターネットが家庭に普及し、デジタル画像のやり取りが頻繁に行われるようになってきた。そして、画像のように大量のデータを短時間に効率よく伝送できる様々な通信サービスが提供され始めた。例えば、ADSL(Asymmetric Digital Subscriber Line)では、上り数百 kbps、下り数 Mbps での伝送ができ、また、携帯電話でも 64 k ~ 384 kbps の通信ができるサービスが始まっている。このような状況から、今後、パソコンや携帯電話にカメラを接続することで顔画像を伝送し合い、互いの顔を見ながら会話を楽しむコミュニケーションが普及するものと思われる。その際、カメラで撮影したままの矩形(くけい)の画像を伝送・表示するのがもっとも簡便であるが、この場合は単に互いの表情を確認しながらの会話にとどまる。

一方で、互いがあたかもその場所にいるような三次元仮想空間を共有することで、臨場感や一体感のある会話を可能とするシステムの研究開発が進められている。しかし、これらでは専用のセンシング機器や、多くの処理量や伝送量を必要とし、一般家庭での実現は現状では困難である。

そこで、参加者の一体感が得られるコミュニケーションをパソコンだけで実現することを目指し、今回、顔チャット™と呼ぶシステムを開発した。これは、ネットワークにアクセスしたチャットへの参加者が、互いの姿が合成された動画像を実時間で共有しながら会話を楽しむものであり、奥行き情報

は持たないため三次元仮想空間にはやや及ばないものの、共有空間にいっしょにいるという一体感や親近感はいずれ得られる<sup>1)2)</sup>。

開発システムでは、筆者らが先に提案したフラクタル抽出法<sup>3)</sup>を用いて参加者のオブジェクト画像を切り出し、MPEG-4(Moving Picture Experts Group-phase 4)の任意形状オブジェクト符号化<sup>4)</sup>で切り出し画像を圧縮する。これらにより、ソフトウェアのみでこのシステムを実時間で動作させることができた。

## 2 顔チャット™システム

## 2.1 顔チャット™とは

インターネットにアクセスした参加者が、キーボードから入力するテキストをやり取りしながら会話を楽しむチャットが盛んである。このようなコミュニケーションシステムにおいて、参加者の姿を実時間で一枚の画像に合成し、その合成画像を互いに共有するものが顔チャット™である。

顔チャット™の画面のイメージを図1に示す。このように、参加者の動画像を共有空間となる合成用の背景(共有背景)に、輪郭に沿ってはり付ける。そして、テキストを吹き出しに表示するか、あるいは、音声を伝送することで会話をを行う。また、適当なアニメーションを合成してもよい。既存システム(図2)のように、参加者の画像をカメラで撮影したままの矩形で並べる場合と比較して、顔チャット™では、同じ空間に



図1. 顔チャット™の画面 参加者の姿を実時間で合成した画像を、互いに共有しながら会話を楽しむ。  
Desktop image of "Kao Chat" (kao : face)



図2. 既存システムの例 矩形の画像が並ぶだけであり、一体感や親近感が得られにくい。  
Desktop image of conventional system

溶け込んだような一体感や互いの親近感が増幅される。

このシステムを用いれば、共有背景を適当に制御することにより、互いの顔を見ながらネットゲームを楽しんだり、テレビのスポーツ中継を見ながら、ひいきのチームを仲間といっしょに応援することもできる。

## 2.2 サーバとクライアントの構成

顔チャット™のシステム構成は、まず、図3(a)に示したクライアント分散型が考えられる。クライアントでは、カメラから取り込んだ画像から顔画像を切り出し、圧縮の後にサーバに送信する。サーバでは、受信した圧縮データを他のクライアントに送信する。各クライアントでは、受信画像を再生し、自分の画像とともに共有背景に合成する。このシステムであれば、サーバは、単にデータを分配するだけであり少ない処理量で済む。そのため、クライアントとして用いるパソコンのうちの1台がサーバを兼ねることもできる。

また、図3(b)に示したサーバ集中型では、サーバで画像を再生し、合成画像を生成する。合成画像は各クライアント

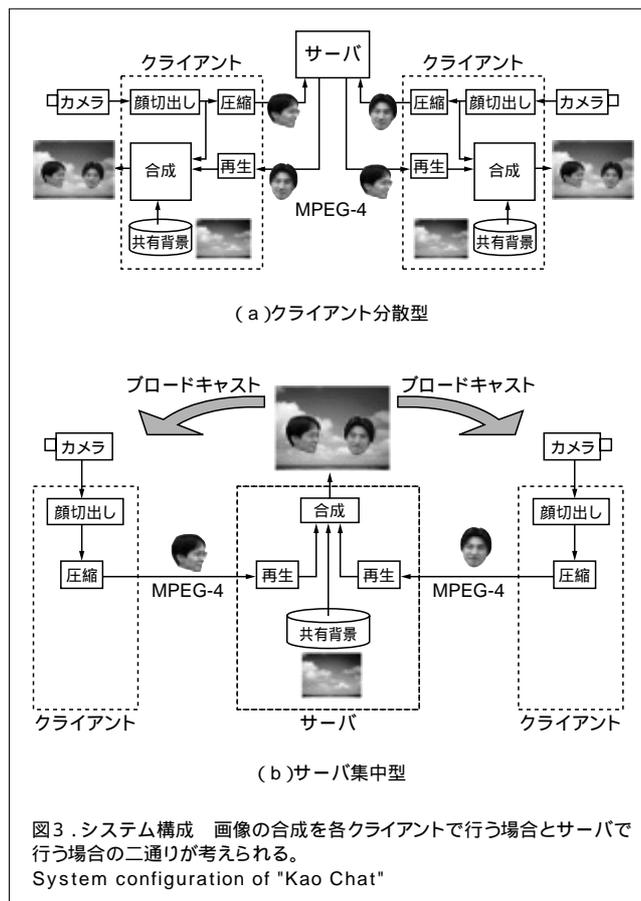


図3. システム構成 画像の合成を各クライアントで行う場合とサーバで行う場合の二通りが考えられる。  
System configuration of "Kao Chat"

に配信する。この場合には、サーバを別に用意する必要があるが、クライアントの処理は軽くなり、特に携帯端末には好つごうである。

## 2.3 顔チャット™システムにおける技術課題

顔チャット™を実現するうえでポイントとなる画像処理技術の課題は、次の二つである。

- (1) 実時間でのオブジェクト画像の切り出し 実時間かつ全自動で利用者の姿を輪郭に沿って切り出す必要がある。ここで、切り出すとは具体的には、人物のアルファマスクを生成することである。アルファマスクは、原画像におけるオブジェクト領域を表すために、オブジェクト領域に1、背景領域に0の画素値を配した2値画像である。
- (2) 少ない符号量での切り出し画像の伝送 アルファマスクと画像データを利用する通信サービスで伝送可能な符号量に圧縮する必要がある。

## 3 開発システム

今回は個人どうしが直接接続する場合を想定し、高い処理速度が求められるサーバを必要としない、クライアント分散型のシステム構成とした。

前述した技術課題(1)に関しては、クロマキー法(単一色の背景にオブジェクトを置き、その色以外をオブジェクトと判定する手法)を用いるのが一案である。これによれば高い精度で切出しが行える。しかし、一般の家庭や屋外でこのような背景を設定するのは困難である。そこで、任意の背景からオブジェクトを切り出せるように、背景差分法に加えてフラクタル抽出法を用いることで、正確に切り出す方法を開発した。詳細は次章で述べる。

技術課題(2)に対しては、MPEG-4の任意形状オブジェクト符号化(この特集のp.10参照)を用いた。この符号化方式では、フレーム間予測を用いるので、顔画像のように動きが比較的小さい画像では、特に高い符号化効率を得られる。

このシステムの信号処理はソフトウェアで実装され、パソコンとカメラ以外の専用ハードウェアは必要としない。

#### 4 オブジェクト画像切出し手法の詳細

オブジェクト画像の切出しは2段階で行われる。まず、背景差分法を用いてオブジェクトの概略形状を取得する。次に、概略形状を元にして、フラクタル輪郭抽出法を用いて、輪郭を正確に抽出する。このシステムを用いた予備実験として、人物の輪郭形状を抽出した結果を図4に示し、この図を用いながら手法を説明する。

##### 4.1 背景差分法による概略形状の取得

クライアントソフトウェアの起動時に、参加者はカメラの前

を空けておく。これにより、部屋の画像が背景画像(図4(a))としてパソコンに保存される。

オブジェクト画像を切り出す際には、カメラから1フレームずつキャプチャされる入力画像(図4(b))と背景画像の差を求め、その差が大きい場合には、そこにオブジェクトがあると判定する。ただ、この判定を1画素ずつ行うと撮影時のノイズにより、背景部分をオブジェクトに、あるいはオブジェクト部分を背景に誤判定することが多くなってしまふ。そこで、複数画素をまとめたブロックごとに画素値の平均を求め、その平均値ごとに判定する。これにより、誤判定が少ないオブジェクト領域が得られ、これを概略形状とする。図4(c)にオブジェクト領域以外を灰色に塗りつぶした結果を示す。図を見ると、ブロック単位で背景差分法を用いるために、概略形状の輪郭は、がたがたしたものとなるのがわかる。特に向かって右側の耳から首にかけて、背景が多くオブジェクト領域に混入した。しかし、この段階では輪郭にずれがあってもかまわない。このずれは、後段の輪郭抽出によって正しい形状に補正される。

なお、背景差分法は、携帯端末などで、カメラが振動する場合には用いることができない。そのような場合には、分離度法などを用いれば概略形状を取得することができる<sup>(1)</sup>。

##### 4.2 フラクタル輪郭抽出法

フラクタル輪郭抽出法<sup>(3)</sup>は、アルファマスクの輪郭がオブジェクトの輪郭とずれている場合に、それと一致させるものである。

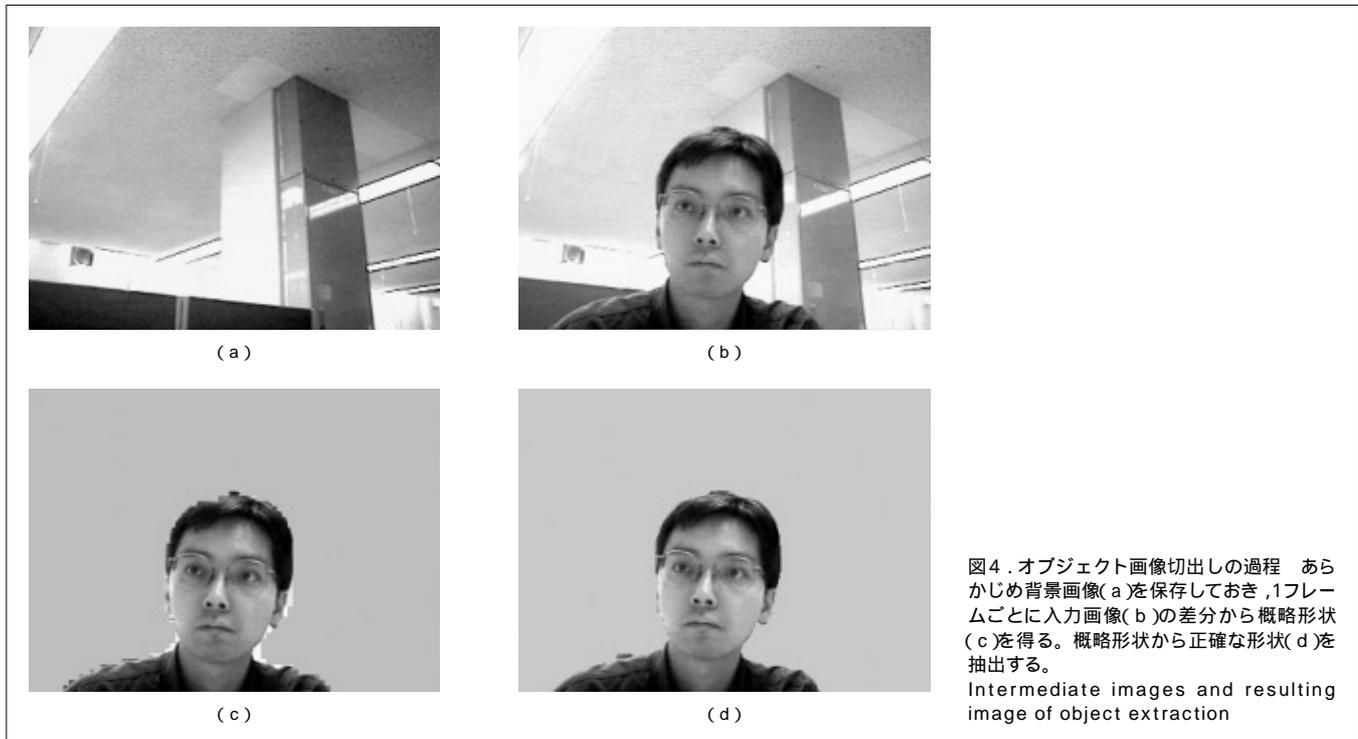


図4. オブジェクト画像切出しの過程 あらかじめ背景画像(a)を保存しておき、1フレームごとに入力画像(b)の差分から概略形状(c)を得る。概略形状から正確な形状(d)を抽出する。  
Intermediate images and resulting image of object extraction

この手法では、アルファマスクの輪郭に沿って自己相似モデリングと呼ばれる解析を行う。そして、解析結果によって決まる縮小写像の反復によって、アルファマスクの輪郭はオブジェクトの輪郭に収束する。この手法は、従来は抽出が困難であった複雑な形状の輪郭線も正確に抽出できるのが特長である。

図4(d)に示したように、概略形状のずれが補正され、ほぼ正確な輪郭が抽出された。

## 5 実験

4台のパソコンをクライアントとして用い、うち1台をサーバに兼用して、LANを介した接続実験を行った。パソコンのMPUのクロックは、各々1.7GHz、1GHz、700MHz、600MHzであり、この順序で接続台数を2台、3台、4台と切り換えた。キャプチャ画像サイズは320×240画素、MPEG-4の目標符号量は256 kbpsとした。

1.7 GHzのパソコンでの切出しフレームレートは、接続台数が2台の場合で6.7フレーム/秒、3台で6.1フレーム/秒、4台で5.5フレーム/秒であった。今回の実験では、高いフレームレートを得るためのボトルネックは、符号量ではなく演算量であった。したがって、より高速のパソコンを用いればフレームレートは増えると思われる。また、画像がキャプチャされてから他のパソコンに表示されるまでの遅延は、およそ0.5秒であった。実験中の画面のようすを図5に示す。

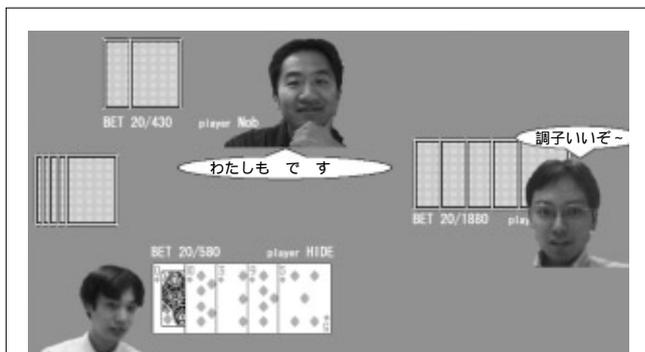


図5．開発システムの画面 参加者の姿が、約6フレーム/秒の動画で表示された。各参加者が入力したテキストは、その人の吹き出しに表示される。

Desktop image of developed software

なお、画像サイズを小さくすれば、更に少ない符号量での伝送が可能であり、例えば、160×120画素の場合、32 kbps程度でも劣化の目立たない画像を送ることができる。

## 6 あとがき

顔チャット™を提案し、そのシステム構成と実現方法を示した。伝送実験により、パソコンのMPUのみを用いて、5.5フレーム/秒のフレームレート(4台接続の場合)で実時間伝送が可能であることを確認した。

例えば、4台を接続した場合の伝送量は、およそ、上り256 kbps、下り768 kbpsであり、現行のADSLサービスによって、リアルタイムに動画を共有することが十分可能である。

このシステムは、背景によらずに正確にオブジェクト画像を切り出せるので、一般の家庭で用いることができる。

## 文献

- (1) 井田 孝,ほか：“リアルタイム顔画像切り出しとMPEG-4を利用した動画共有型チャットシステムの実現”. 知能情報メディアシンポジウム論文集. 東京, 2000-12, p.13 - 18.
- (2) 竹島秀則,ほか：“参加者の切り出し画像を実時間で共有するチャットシステムの開発”. 電子情報通信学会総合大会講演論文集. 東京, 2002-03, p.391 - 392.
- (3) Ida, T.; Sambonsugi, Y. Self-affine mapping system and its application to object contour extraction. IEEE Trans. Image Processing. 9, 11, 2000, p.1926 - 1936.



井田 孝 IDA Takashi, D.Eng.

研究開発センター マルチメディアラボラトリー 研究主務, 工博。画像符号化と画像構造化アルゴリズムの研究・開発に従事。IEEE, 電子情報通信学会会員。Multimedia Lab.



竹島 秀則 TAKESHIMA Hidenori

研究開発センター マルチメディアラボラトリー。画像構造化アルゴリズムの研究・開発に従事。電子情報通信学会会員。Multimedia Lab.



堀 修 HORI Osamu, D.Eng.

研究開発センター マルチメディアラボラトリー 主任研究員, 工博。映像処理技術及びマルチメディア関連技術の研究・開発に従事。IEEE, 情報処理学会, 電子情報通信学会会員。Multimedia Lab.