

# 枯れた技術に基づく 視線共有ビデオチャット グループウェアの研究開発

川路崇博・小泉真也

## ●要約

これまでも、ビデオ会議システムが開発されており、ビデオによる会議環境の実現に留まらず臨場感をも伝えようとする試みが行われている。しかし大規模かつ汎用性の低いシステムによりそれらは実現されている。そこで本稿では方向性を変え、「枯れた技術」を用いてより実用性の高いビデオチャットグループウェアの構築を狙った。

その実現においては、普及期を迎えたネットブック型 PC に着目し、ほぼ標準的に装備されている Web カメラを用いた双方向性メディアグループウェアの研究開発を行うことを検討する。Web カメラを用いたビデオチャット会議では、いわゆる「カメラ視線」ではなく、画面内に映し出されたチャット相手の顔を見て会話をを行う。そこで問題になるのは、視線の共有が行われなかったために、双方向性メディアとしてのネットブック型 PC の機能を生かし切れていない部分である。また遠隔であることを感じさせない臨場感が得られにくいという点も課題となる。そこで本稿では、ネットブック型 PC を用いたビデオチャットシステムにおける疑似的視線共有の手法を提案する。視線共有は PC に付属する Web カメラで撮影された顔画像に対して、視線の動きベクトルを元に映像修正を行うことにより疑似的に実現する。

## ●キーワード

枯れた技術の水平思考

視線共有

ビデオチャット

ネットブック型 PC

双方向性メディア

## 1. はじめに

### 1.1 ネットブック型 PC の誕生

ネットブック型 PC は非常に安価で小型のノートブック型 PC であり、Web ブラウジングやメールの送受信程度の比較的マシンパワーを必要としない用途においてシェアを伸ばしている。

2007年のアスーステック・コンピュータ (ASUS) による EeePC シリーズの発売を皮切りに、同じ台湾のメーカーであるエイサー (Acer) もネットブック型 PC の開発・発売に追随し、さらには米国ヒューレット・パッカード (Hewlett-Packard) ほか、日本国内大手メーカーである SONY、富士通なども続いて市場に参入している。

小型の PC は、UMPC (Ultra-Mobile PC) として、2006年 3月にマイクロソフトやインテルなどが発表した規格で、ノート型 PC と PDA (Personal Digital Assistant または Personal Data Assistance) の中間に位置づけられるものであった。

その後、モバイル用途の UMPC のタッチパネル機能を廃し、ハードディスクドライブの代わりにフラッシュメモリディスクを用いたネットブック型 PC が開発された。ULCPC (Ultra Low-cost PC) の爆発的な市場への広がりを受け、現在マイクロソフトが ULCPC 向けに、独自のライセンス発行を行っている。なお、ライセンスの発行には PC の基本スペックを下回る必要がある。主な基本スペックは以下の通りである。

- ・ OS : Windows XP ULCPC 向け Home Edition、または発展途上国向けの Starter Edition
- ・ ディスプレイ : 14.2インチ (1) 以下 (タッチパネル不可)
- ・ メモリ : 1GB 以下
- ・ CPU : Atom などの特定のシングルコア CPU
- ・ 内蔵ディスク容量 : HDD160GB 以下、または SSD (Flash Solid State Drive) 16GB以下 (2)

なお、OS については Ubuntu をはじめとした Linux ディストリビューションが提供されているなど、必ずしも Windows XP であるとは限らない。各メーカーともに「ネットブックは 2 台目のノート型コンピュータであり、既存のノート型 PC の市場を脅かすことはない」としながらも、その市場の反応を無視できないのが現状である。

### 1.2 ネットブック型 PC への着目

ネットブック型 PC が例え各ベンダーが主張するように 2 台目のノート型コンピュータであっても、その普及は無視できない。また、ネットブック型 PC の特徴として、Web ブラウジング、メールの送受信の基本機能の大前提として、インターネットへの接続があげられる。日本国内においては、e-Mobile が EeePC を抱き合わせて端末を販売しているなどをみても [メディアワールド]、インターネットへの接続が前提であることは明らかである。

ネットブック型 PC の販売台数は右肩上がりを示しており [DISPLAYSEARCH]、今後もその市場成長が見込まれる。

そこで我々は、独自の進化を遂げたと言っても過言ではないネットブック型 PC を利用したビデオ

会議システムの構築を試みる。ネットブック型 PC には、マイクロソフトの ULCPC 向けの OS ライセンスに抵触しない Web カメラが標準装備されていることが多い。むしろ、Web カメラが装備されていないネットブック型 PC はほぼ皆無といってもよい。

## 2. 研究の目的

普及段階に入ったネットブック型 PC は、主に外出先での Web ブラウジングやメールの送受信など、本来の目的で利用されることが多いが、我々が注目したのは、事実上標準的に装備されている Web カメラである。Skype や MSN メッセンジャーなどの IM (Instant Messenger) において、ビデオチャットは標準的に実装されている機能の一つである。

しかし、ビデオチャットを行う際に、視線の共有がしづらいという問題がある。これは、「カメラ目線」でビデオ会議を行わなければ、画面上の人物の視線と PC を操作する人物の視線が一致しないことに起因する。しかし、カメラ目線でビデオチャットを行うと、画面上の相手の顔を見ることができない。このことにより、せっかくのビデオチャット環境が生かされないのが現状である (図1)。

そこで、ネットブック型 PC の Web カメラから得られた顔画像を用いて動きベクトルを元に映像修正を行うことにより、疑似的な視線共有環境を構築することを目的とする。



図1 従来型ビデオチャット画像 (左) と、本来期待すべき視線共有画像 (右)

## 3. 関連研究

視線共有を実現したビデオ会議システムはすでに多数の先行研究がある。たとえば LERTRUSDACHAKUL らの研究では [LERTRUSDACHAKUL a] [LERTRUSDACHAKUL b]、視線共有にとどまらず、複数人との視線共有をも実現している。ただし eye contact によるこのシステムは設備が大規模であり、視線移動は手動による部分も存在している。また、大野 [大野] の実装システムは規模が大きいシステムを利用しながらも、遠距離間コミュニケーションにおいてユーザの視線位置と頭部の方向のみを伝送する。大野の研究においては、視線位置自体を機械的に相手に伝える機能がある。視線のトラッキングと補正、相手への提示に約 1 秒を要する。視線共有における知識移転の有効性は、Sellen [Sellen]、Velinotte ら [Velinotte] によって報告がなされているが、大野のシステムでリアルタイムでの伝達を行った際の検証は行われていない。

動画の動きの推測は映像修正における情報取得の重要な技術である。動きをパラメタ化する手法は、大きく分けて二種類に分類できる。一つは動きベクトルを用いる方法、そしてもう一つはカルマンフィルタを用いる手法である。動きベクトルを用いる手法は、時間的に連続するデジタル映像の中で物体の動きを画素のベクトルで表したものである〔金子〕〔3〕。また、画素のベクトルは平面的な動きを観測するため、アフィン変換を用いてグローバルな動きに対応する手法も提案されている〔金野〕。カルマンフィルタは、離散的な誤差のある観測から、物体の位置と速度のように時々刻々と時間変化する量を推定するために用いられ、複数のカメラで撮影した映像から物体の動きを推測する手法が提案されている〔伏尾〕。

## 4. プロトタイプ設計と想定される効果

### 4.1 プロトタイプの設計と実装

本研究では、動きの情報を動きベクトルで記録し、対話の映像にこの動き情報を適用することによって問題の解決を図る。

動きベクトルとは、オプティカルフローとも呼ばれ、フレーム中の各画素がどのような方向にどのような速度で動いているかをベクトルとして求めたものである〔Horn〕〔Nagel〕〔Adiv〕。

モニタを見ている動画フレームをフレームA、カメラを見ている動画フレームをフレームBとした場合、フレームAの像とフレームBの像の差分を考える。ここで、モニタを見ることによって対話を確立する場面において、対話中の動きの変化は少ないと仮定する。概してネットブック型PCはモニタの上方にカメラを備えているため、フレームA-B間の差異は視線あるいは顔の垂直移動が主であると推察できる。フレームAおよびBを学習し、B-A間の差分要素Cを作成すると、CをフレームAに適用してフレームBが得られる。以上より差分要素Cの取得を事前に行って、通信時の映像フレームA'にCを適用すれば、映像における疑似的な視線一致が実現すると考える。図2に画像処理の一連の手順を示す。

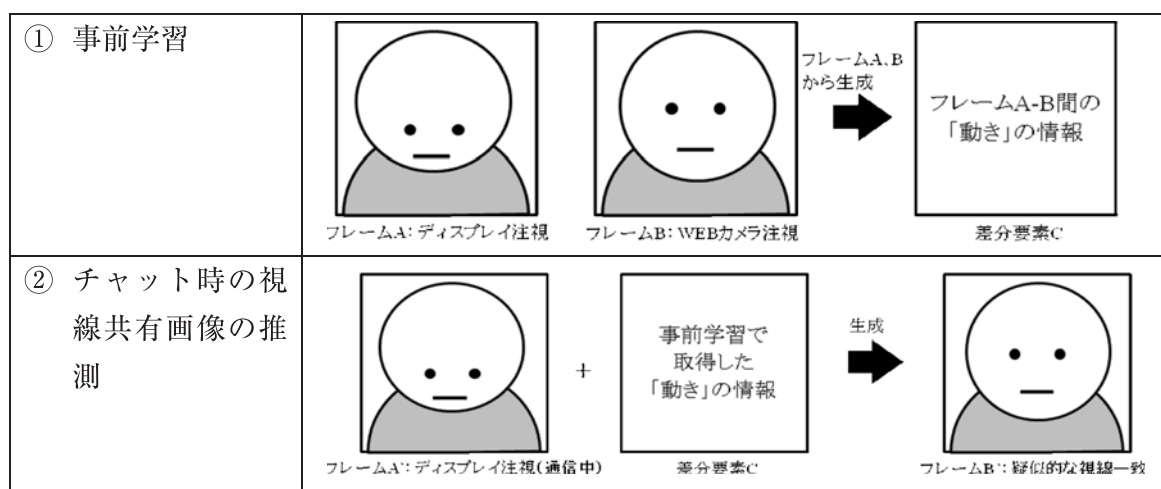


図2 提案手法の手順

### 動きベクトルの決定方法

動きベクトルを求める方法は、大きく分けて次の2つである。ひとつは「ブロックマッチング法」であり、もうひとつは「勾配法」である。

ブロックマッチング法では、画像中のある大きさの領域をテンプレートのブロックとし、次フレームの画像中を全探索し、前フレームの評価するブロックと次フレームの注目ブロックとの差分評価関数の値を最小とする点を対応点として決定する。ブロックマッチング法の利点は、直感的に理解しやすく実装がしやすいこと、雑音に強いこと、そして輝度値の急激な変化に対してもフローの誤差が少ないことである。反面、欠点としては、全探索を行うため計算時間が膨大となること、そして拡大・縮小、回転運動に弱いことが挙げられる。

勾配法では、動く物体を示す画素の明るさは移動後も変化しないという仮定のもとで、時空間微分と動きベクトルとの関係式を求め、この式を利用して対象の動きを推定する。勾配法の利点は、対応点検索をおこなう必要がないことである。反面、欠点として弛緩法などのエネルギー最小化のための前処理を必要とすること、輝度値が急激に変化するところではフローの誤差が激しいこと、そして雑音に弱いことが挙げられる。

本研究において想定する場面では、眼や顔といった動きの対象の移動量が小さいこと、そして通信前に学習データの取得のための動きの観測を行うことから、雑音に強いブロックマッチング法を採用する。

以下にブロックマッチング法の手順を示す。はじめに、フレームAの任意の画素  $p(x,y)$  について近傍  $n*n$  画素の矩形範囲を定める。この矩形範囲を画素  $p(x,y)$  のテンプレートと呼び  $P(x,y)$  と表す。次に、フレームBの全画素に対して  $P(x,y)$  と同じパターンを持つ矩形範囲を探索する。このとき  $P(x,y)$  に対して最も距離の小さいフレームBの矩形範囲を  $Q(x,y)$  と表し、 $Q(x,y)$  の中心画素を  $q(x,y)$  で表すとき、 $p(x,y)$  におけるベクトルは座標  $p(x,y) - q(x,y)$  間の縦横それぞれの移動距離として求めることができる。図3にブロックマッチングに基づくフレーム間の対応点決定を示す。また図4に実際の画像から取得した動きベクトルを示す。図4(a)はネットブックPCのモニタを注視した画像を示し、図4(b)はネットブックPCのカメラを注視した画像を示す。図4(c)は、図4(a) - 図4(b)間の動きベクトルである。

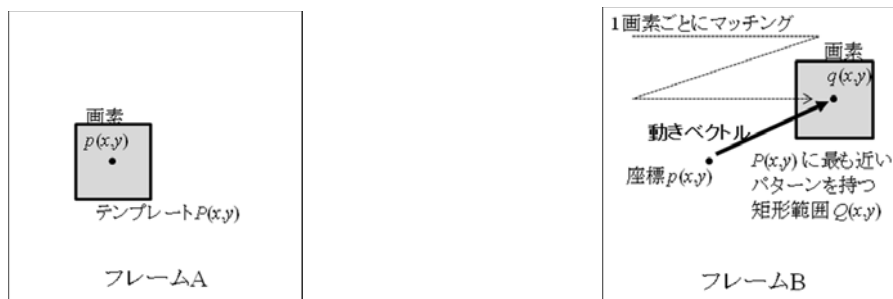


図3 ブロックマッチングに基づくフレーム間の対応点決定





図4 実際の画像から取得した動きベクトル  
(撮影 EeePC 130万画素、速度ベクトルを計算する画素間の距離：5の場合)

#### 4.2 本システムで想定される効果

本システムを、後述する今後の展開まで研究開発を行った場合、次のような効果が想定される。

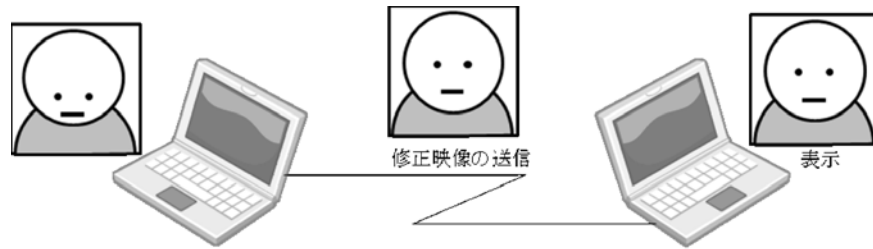
- ・ ネットブック型 PC の付加機能である Web カメラを利用した“より”ネットブック型 PC “らしい”利用の幅を広げることができる。特にファミリーユース、マンツーマン形式での会議での利用シーンにおいて効果を発揮する。
- ・ 従来の会議（対面/同室型）を対面/遠隔型グループウェア（本研究での提案システム）とすることにより、会議環境に大規模な設備を準備せず、安価に構築することができる。
- ・ 視線の共有により、暗黙的な知識流通支援が期待できる。

### 5. 今後の展開

今後我々は、静止画ではなく動画での疑似的視線共有環境を構築する。さらにネットブック型 PC の名が示すとおり、実装された本システムを用いてネット越しでの会議を行い、実際に一カ所に集まった会議との比較を行う。さらに、ステレオマイクを用いた会議参加者の位置関係を割り出し、会議室の雰囲気も再現できる環境の会議支援グループウェアの開発をも視野に入れる。

次に、システム構築に置いて検討すべき点を示す。検討事項は次の2点である。ひとつは、通信中の映像に動きベクトルの情報を適用するタイミングである。もうひとつは通信中の映像について、視線移動に基づく動きベクトルの再学習の必要性を検討することである。通信中の映像に動きベクトルの情報を適用するタイミングは、図4に示すように送信前または受信時の2つが想定できる。送信前の適用ではカメラで撮影した映像に動きベクトルを適用して送信する。この方法では純粋に映像を送信するため単一の通信プロトコルのもとでシステムを構築できるが映像修正による遅れが懸念される。受信時の適用では、予め動きベクトルを送信しておき、通信時に受信した映像をその動きベクトルで修正して表示する。この方法では映像送信に遅れは生じないが、システム的设计において通信プロトコルを複合的に考慮する必要があること、および受信端末の映像処理の遅れが懸念される。

(a) 送信前に修正



(b) 受信時に修正



図5 動きベクトルを適用するタイミング

動きベクトルの再学習は、事前に学習した動きベクトルが、画像修整情報としてどの程度の時間耐え得るかを検討する必要がある。また、通信中に動きベクトルの再取得が必要である場合、その手法はブロックマッチング法とするか勾配法とするかの検討も必要であると考えられる。

●註

- (1) 10.2インチ以下から修正された。
- (2) HDD 80GB 以下から修正された。
- (3) 動きベクトルに関する研究は数多く行われているため、本稿ではたとえば金子らの研究をあげる。

●参考文献

[Adiv] G. Adiv : *Determining three-dimensional motion and structure from optical flow generated by several moving objects*, IEEE Trans. On PAMI, Vol.7, No.4, pp-384-401, 1985.

[DISPLAYSEARCH] DISPLAYSEARCH, Mini-Note (Netbook) Shipments to Double Y/Y to More Than 30M Units in 2009, While Notebook Shipments Flatten : [http://www.displaysearch.com/cps/rde/xchg/displaysearch/hs.xsl/071309\\_mini\\_note\\_netbook\\_shipments\\_to\\_double\\_y\\_y\\_to\\_more\\_than\\_30m\\_units\\_in\\_2009.asp](http://www.displaysearch.com/cps/rde/xchg/displaysearch/hs.xsl/071309_mini_note_netbook_shipments_to_double_y_y_to_more_than_30m_units_in_2009.asp) (2009年7月16日閲覧)

[伏尾] 伏尾光平、守田了：スケート中継映像からの競技者の動きの推定，電子情報通信学会技術研究報告・PRMU，パターン認識・メディア理解，104(524)，pp.107-112，2004。

[Horn] B. K. P. Horn and B. G. Schunck, : *Determining optical flow*, Artificial Intelligence, Vol. 17, pp185-803, 1981.

[金子] 金子豊、鹿喰善明、田中豊：画像の局所的特徴を考慮した固有値こう配法による動き推定、電子情報通信学会論文誌D、Vol.J85-D2, No.2, pp.262-271, 2002。

[金野] 金野彰：フローベクトルを用いたニュース速報テロップの除去，岩手県立大学大学院ソフトウェア情報学研究科 修士論文，2006。

[LERTRUSDACHAKUL a] LERTRUSDACHAKUL Thitiporn, TAGUCHI Akinori, AOKI Terumasa, YASUDA Hiroshi, *A Study on Gaze Communications and Synchronous Pointing of Multiparty Videoconference for IshinDenshin System*, 電子情報通信学会技術研究報告，電子情報通信学会，pp.177-180, 2003.

[LERTRUSDACHAKUL b] LERTRUSDACHAKUL Thitiporn, TAGUCHI Akinori, AOKI Terumasa, YASUDA Hiroshi :

*Transparent Gaze Communications for Multiparty Videoconference System*, IEICE TRANS. INF. & SYST., VOL.E87-D, NO.6, pp.1328-1337, JUN 2004.

[メディアワールド] 株式会社メディアワールド、イー・モバイル100円パソコンセット：

<http://emobile-agent.com/100enpc/> (2009年9月14日閲覧)

[Nagel] H. H. Nagel : *Displacement vectors derived from second-order intensity variations in image sequence*, Computer Vision, Graphics and Image Processing, vol. 21, pp. 85-117, 1983.

[大野] 大野健彦：視線共有に基づく遠隔地間コミュニケーション，電子情報通信学会技術研究報告，電子情報通信学会，pp.55-60，2005．

[Sellen] Sellen, A. J.: *Remote Conversations: The Effects of Mediating Talk With Tecnology*, Human-Computer Intraction, Vol. 10, pp. 401-444. , 1995.

[Velinotte] Velinotte, E. S., Olson, J., Olson, G .M. and Fu, X.: *Video helps remote word : speakers who needs negotiate common ground benefit from seeing each other*, CHI '99: Proceedings of the SIGCHI conference on Human factors in computing systems, ACM Press, pp. 302-309, 1999.

## ● 英文タイトル

Research and Development of Video Chat Groupware with Shared Sight based on the Withered Technology

## ● 要約

Numbers of the video conference system has been developed until today and such systems try to deliver the conference environment along with the realistic sensation. However, these are constructed with large scale system that lacks of general versatility. This study aimed to build the video chat groupware for practical use by using the withered technology.

To develop the system, we focused on the netbook(Mini-PC) that is widely spread, and discuss the development of the interactive media groupware using web camera that is included in the netbook as a standard equipment.

Users of video chat conference with web camera conduct the conversations by looking at face of the person in the screen, not looking at the camera. This causes the problems that users do not share the eye line, and it results the dysfunction of system as the interactive media. In addition, users hardly receive realistic sensation without realizing distance between them. In this paper, the simple video chat system with shared eye line is proposed. Images used in this system are processed with optical-flow method to correct the eye line and the result enables users to share the eye line.

## ● Keywords

Lateral thinking with Withered technology

gaze awareness

video chat

netbook

interactive media