

## 特集 ■ ユビキタスと VR

## 拡張記憶インタフェース



河野恭之

奈良先端科学技術大学院大学

KONO YASUYUKI

## 1. はじめに

拡張記憶は、ユーザの日々の活動におけるコンテキストを計算機が保持しておき、その一部を再現してユーザの想起活動を促すことで、ユーザ自身の記憶領域が拡張したかのように感じられるような支援を行うという考え方である [1][2]。拡張記憶インタフェースの実現に際しては、計算機を身体に装着し常に動作させ、ユーザの日常活動を入力・記録すると共に、ユーザの置かれた様々な場所や時における状況を認識するために身体装着デバイスや環境中の様々なデバイスと連携するという意味で、ウェアラブルコンピューティングやユビキタスコンピューティングと親和性が高い。例えば Remembrance Agent<sup>1)</sup> は、ウェアラブル計算機環境を場所に関連して記録しておき、次にその場所に来た際に環境を再現することで、その場所で行っていた作業の継続を容易にする [1][3]。作業と関連づける場所情報の取得のために、環境に配置された IR ビーコンが用いられている。

ここ数年の計算機の小型化、マルチメディア処理能力の急速な進歩と環境整備に後押しされるように、画像・映像やオーディオなどのマルチメディア情報を扱うタイプの拡張記憶インタフェースの研究が増加傾向にあり、体験記録やその利用方法・技術に関するワークショップ等が積極的に開催されるようになってきている<sup>2)</sup>。これは、カメラやマイクなどのメディア入力デバイスをユーザが装着すればユーザが「見た」り「聞いた」りした一人称情報が獲得・蓄積でき、そのデータを再生することで比較的容易にユーザの体験を「再

現」できることが理由であると考えられる [15]。人間の五感のうち視覚と聴覚に対応する画像・映像及びオーディオ情報は、ホームビデオなどにより体験を記録するものとして広く認知されているのに対し、残りの三感(嗅覚, 味覚, 触覚)は他の感覚と独立したものではなく、多くの人にとっては視覚, 聴覚情報に付随したものとなっている。記録された視覚・聴覚に属する体験情報を再生すれば、それを閲覧した人が再生時のコンテキストにおいて再解釈を行い、含まれている多様な情報の中からエッセンスを抽出する。

体験の記録時, 及び体験記録再生時における能動的なユーザオペレーションの必要性により, 拡張記憶インタフェースの特性が異なると考えられる。表 1 にその分類を示し, 以下でそれぞれの特性を論じる。能動とは, 個々のユーザの体験を記録(を開始)する際, 又は再生(を開始)する際のトリガとしてユーザの能動的なオペレーションが必要なもの(能動)と, インタフェース側が概ね自動的に記録, 再生を行うもの(受動)である。記録時の「受動」は, システム動作時は常時記録を行うと共に検索のための自動インデックス付与を前提とするものが多い。再生時にユーザが「受動」的であることを可能にするためには, 記録されている特定の体験と現在の状況とを結びつけるための状況把握機能が必要となる。

表 1 拡張記憶インタフェースの分類

		記録時 (入力時)	
		能動	受動
再生時 (出力時)	能動	(a)	(b)
	受動	(c)	(d)

<sup>1)</sup> <http://www.remem.org/>

<sup>2)</sup> MSE2004(<http://www.ii.ist.i.kyoto-u.ac.jp/~sumi/pervasive04/>) や CARPE2004(<http://research.microsoft.com/CARPE2004/>) など

(a) システムへの取り込み(記録時)も、再生開始のトリガもユーザが能動的に与える。システムの支援は最小限となるが、ユーザ個々の目的・視点に応じて様々な使い方ができ、適用範囲が広い。Time Machine Computing[4]が広い意味でこのカテゴリに分類される。

(b) 常時獲得・記録されるユーザ体験を自動的に索引付けしておき、ユーザが明示的に与えたトリガに応じて体験記録の検索を行い提示する形式である。探し物や要約といったシステムのアプリケーションと親和性が高い[5][6]。

(c) 記録しておくべき体験(コンテンツ)をユーザが与えるとシステムがそれらを構造化し、その情報に基づき「提示するべき時」に「提示するべき体験」を提供する。リマインダ的なアプリケーションに向く。絵画などの対象物に関連づけられたビデオコンテンツを、対象物を見たときに再生するDyPERS<sup>3)</sup>など、AR/MRでよく見られるアプリケーションがこのカテゴリに分類される[7]。

(d) 記録も提示も自動的に行われる。ユーザの負担は最小限となるが、少ないユーザ入力から有用な出力を得る必要があるため、適用場面やアプリケーションは限定される。前述のRemembrance Agentがこのカテゴリとなる。

以下本稿では、上述の特性に合わせて筆者らが取り組んでいるいくつかの拡張記憶インタフェースを紹介する。

## 2. Residual Memory

Residual Memoryは、常時獲得されるユーザ視点画像と既に蓄積された未編集の長大な体験映像記録との類似性を判定し、類似性の高い映像区間を「ユーザの現在地で過去に記録された映像」であると判断して検索結果とする、表1(d)のタイプの拡張記憶インタフェースである。この際に(1)頭部装着カメラは激しく運動するユーザの頭部動作の影響を受けるため、カメラ近辺に取り付けた加速度センサによりユーザの頭部運動を検出すると共に、(2)ユーザ自身の手や走行中の自動車などは場所の類似性に無関係のため、映像中の運動領域を検出し、それらの影響をキャンセルした形で類似度計算を行う方式を提案している[8]。

このような検索方法を提供し、検索された映像区間を提示することで、システムは2種類の利用形態をユーザに提供することができる。一つはユーザが来たことのある場所に意識することなく再び訪れた際に、その場所での拡張記憶要素を提示してその場所での思い出を想起させるという形態である。もう一つは、特定の場所で発生

した出来事を思い出すために、ユーザがその場所に行くことで拡張記憶要素を参照して内容を確認するという備忘録的、すなわち表1(c)的な使い方である。あらかじめ整理されていない体験映像記録に対して、リアルタイムに得られている頭部装着カメラからの画像をクエリとして高速に検索できる手段を提供することは、体験時及びその後の整理の際に重要視していなかった体験から情報を抽出するために必要と考えられる。我々は拡張記憶の高速な検索手段として、頭部装着カメラ映像の時間的、空間的遷移特性を利用したHySIM (Hybrid-Space Image Matching)を提案している[9]。

頭部装着カメラはユーザと共に空間移動するため、検索クエリ画像と検索対象の体験映像記録の間には、頭部動作により類似度の増減を繰り返しながらゆるやかにトレンドが変化する時間的遷移特性、おおよそ同じ空間で記録された映像が体験映像記録中には複数存在すると共に類似画像が短い時間内に密集するという空間的遷移特性の二つがある。

我々は、二つの特性が共に時系列・連続的に記録された一つの映像記録を介して関係を持つことから、この関係性を連結したままでこれらの特性を考慮すれば、大規模映像記録からの高速な類似視点映像の検索を実現できると考えた。アイデアとしては、時間的特性による探索を行える空間と空間的特性による探索を行える空間の異なる空間を二つ構成し、これらの空間を行き来しながら探索するものである。

## 3. I'm Here!

現代人の日常生活環境には数多くの物が存在している。人はそれらを利用することで、日常生活における様々なタスクを遂行している。前準備としての物探しタスクを遂行することに手間取り、物を利用したタスクを行うまでに時間を浪費する場面は、人の日常生活で頻繁に発生する。一般のビジネスマンが物探しタスクのために浪費する時間は1年間に150時間であると言われる[10]。コンピュータの支援によってユーザが物探しタスクを効率的に行えるようになれば、ユーザは浪費していた時間を他の有意義なタスクに費やし、ユーザの日常生活がより豊かになると期待できる。

我々は表1(b)のタイプの拡張記憶インタフェースとして、これまで物探しタスク支援を行うウェアラブルシステムI'm Here!を提案し構築してきた[11]。物探しタスクには、自分自身が最後に置いた物を対象とする場合と、自分が置いた後で他人が移動させた物を対象とする場合

<sup>3)</sup> [hecp://research.microsoft.com/~nuria/dypers/dypers.htm](http://hecp://research.microsoft.com/~nuria/dypers/dypers.htm)

がある。我々は前者の物探しタスクに関する支援に着目した。この場合、物探しが失敗する原因は、物を最後に置いた場所を思い出すことができないという記憶活動エラーにある。人の視点に近い位置から撮影した視野映像を拡張記憶として蓄積・検索し、人の記憶活動を補強するというアプローチは、日常生活における物探しタスクを支援するサービスに適している。

I'm Here! が支援の対象とする物(対象物)は、人によって把持され移動される物(把持物体)である。日常生活において、把持物体の数は多い。人はそれらの物体を把持して頻繁に移動するので、把持物体は物探しの対象となることが多い。

I'm Here! は、ユーザの視野映像が撮影された時点の対象物の固有名を視野映像にインデックス情報として付加し、インデックス付きの視野映像を視野映像データベースとして蓄積する。物探しタスクを支援する場面において、I'm Here! は対象物の固有名をキーとして視野映像データベースを検索し、対象物をユーザが最後に把持した時点の視野映像をユーザに提示する。ユーザは映像中の周辺情報から、対象物の置き場所を推測できる。

図 1 は I'm Here! のハードウェア構成である。I'm Here! が視野映像データベースを構築するためには、ユーザの視野映像を撮影する機能と、対象物を認識する機能が必要である。そのために、I'm Here! は頭部装着型カメラを用いてユーザの視野映像を撮影し、視野映像に映る対象物の見た目の画像に基づいて対象物を認識する。対象物を認識するための対象物データベースと視野映像データベースは、ウェアラブルコンピュータに蓄積さ

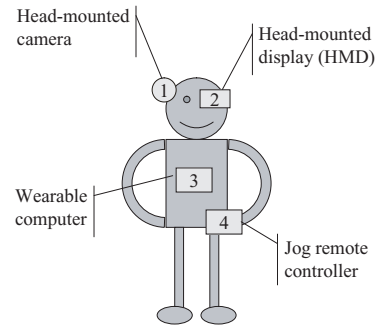


図 1 I'm Here! のハードウェア構成

れる。ユーザは、ジョグリモートコントローラを用いて I'm Here! に指示を送り、頭部装着型ディスプレイによって I'm Here! から提示される情報を受け取る。

図 2 に I'm Here! の機能フェーズを示す。I'm Here! の機能フェーズは、登録(Registration)、インデクシング(Indexing)、検索(Retrieval)のフェーズから構成されている。

**登録(Registration)：**ユーザはこのフェーズで対象物を I'm Here! に登録する(図 2 A)。この時点でユーザに課される操作は、対象物を把持して目の前で回転させるという簡単なものであり、ユーザが対象物を利用するタスクのついでに気軽に行える。その後、ユーザは対象物に固有名を入力する(図 2 B)。A において登録された対象物の画像的な特徴は、B で入力された固有名と関連付けられ、I'm Here! の対象物 DB に格納される。

**インデクシング(Indexing)：**I'm Here! は、登録のフェーズで構築された対象物 DB を利用し、日常生活を送るユーザの視野映像に対して自動的にインデックスを付加

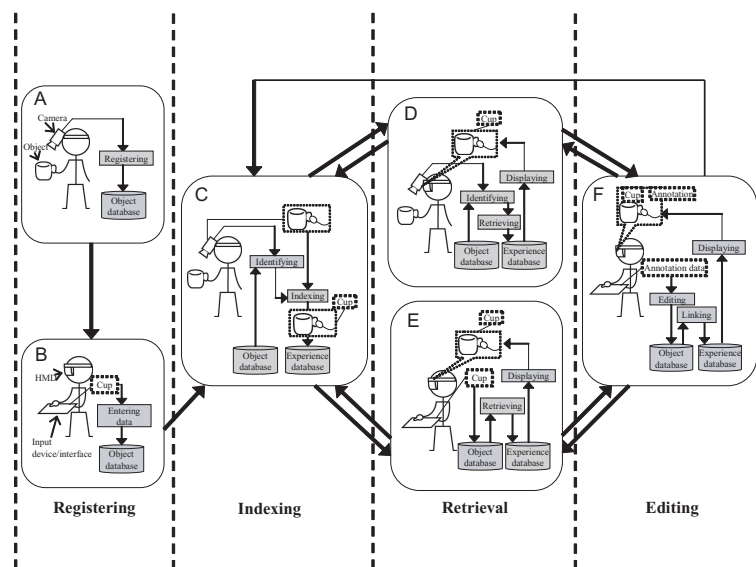


図 2 I'm Here! の機能フェーズ

して視野映像 DB を構築する。ユーザは、このフェーズでは I'm Here! に指示をする必要はない。

**検索 (Retrieval):**ユーザは検索のフェーズを起動し、I'm Here! に登録された対象物の一覧から検索対象となる対象物の固有名を指定する。I'm Here! は対象物の固有名を検索キーとして視野映像 DB のインデックス情報を検索し、最後に対象物が把持された時点の視野映像をユーザに対して提示する。

I'm Here! がインデクシングのフェーズにおいて視野映像データベースを逐次構築するには、ユーザが把持した対象物の画像を実時間で入手し、物体認識を行う必要がある。我々は、背景領域から対象物の画像を実時間で抽出する機能を有する頭部装着型カメラ ObjectCam2 を開発した [12]。図 3 は装着状態の ObjectCam2 である。



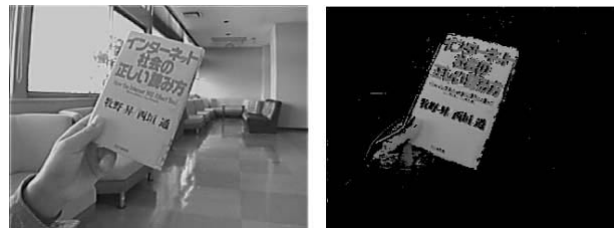
図 3 ObjectCam2 装着外観

ObjectCam2 は、カメラ前から前方の環境へ向けて赤外光を照射し、その反射光を計測する。ObjectCam2 は赤外光源を明滅させ、消灯時にはカラー画像を、点灯時には反射してくる赤外光を含むカラー画像を交互に撮影する。そして、ObjectCam2 は消灯時と点灯時の画像の明度を差分し、照射された赤外光の反射光のみを観測した画像 (反射赤外画像) を得る。

I'm Here! は、ObjectCam2 を用いることでユーザ視点からのカラー画像と反射赤外画像を共に 30fps で得ることができる。I'm Here! はそれらの画像を利用し、ユーザが把持している対象物の画像を抽出する。反射光の輝度が距離の二乗に反比例するという原理に基づいて反射赤外画像を二値化してマスクとして用いることで、I'm Here! はカメラ、すなわちユーザの顔面からの距離が近い領域 (近接領域) の画像を得ることができる (図 4)。近接領域画像からユーザの手領域を肌色に基づいて排除することで、I'm Here! はユーザが把持している対象物のカラー画像・反射赤外画像を得る。I'm Here! はこれらの画像を用いて、対象物の認識を行う。

#### 4. Ubiquitous Memories

Ubiquitous Memories は、ユーザが記憶を整理するために拡張記憶の各要素を実世界に遍在する物理的対象と関連付ける機能を提供する、表 1(a) のタイプの拡張記憶インタフェースである [13]。実世界物体は、人にとっ



(a) カラー画像 (b) 抽出結果

図 4 ObjectCam2 による対象物画像抽出

て体験を思い出すトリガとなる機能を備えている [14]。Ubiquitous Memories では対象物に「触れて」拡張記憶要素を「貼り付ける」というメタファを、この関連付けの操作に採用している。このメタファを実現するために、実世界対象には RFID タグが埋め込まれている / 貼り付けられているとし、ユーザは手首に小型 RFID タグリーダを装着しているとす。ユーザはまた、拡張記憶要素を「貼り付ける」「コピーする」といった個々のコマンドに対応するオペレーションタグと呼ばれる特別な RFID タグをいくつか身に着けている。

実世界対象を用いた拡張記憶整理操作の基本手順は下記の通りである。ある時、ユーザがある場の状況その場に存在する実世界対象 X に貼り付けようとしたとする。その場合、ユーザはまず貼り付け操作に対応するオペレーションタグ (自分の身体) に触れる。次にユーザは実世界対象 X に触れると X に埋め込まれ / 貼り付けられている RFID タグの ID がリーダから読み込まれる。ユーザの頭部装着カメラから得られた映像が拡張記憶としてタグ ID と関連付けられ、実世界対象 X を管理するサーバに無線 LAN 等のネットワーク経由でその拡張記憶が転送され蓄積される。

実世界対象 X に対応付けられた拡張記憶に参照しようとするユーザは、X に単純に触れると、同様にタグリーダが X のタグ ID を読み込む。システムは X を管理するサーバに拡張記憶要素の転送を要求し、ダウンロードされた拡張記憶要素がユーザの HMD に再生される。上記のような実世界対象に「触れる」という単純な操作構成により、拡張記憶の整理と参照が可能となっている。

検索された拡張記憶要素の再生中に上記の貼り付け操作が行われると、再生中の拡張記憶要素が貼り付け操作の対象となる。これにより、ユーザはある実世界対象と一旦関連付けた拡張記憶要素を再整理することが可能となる。また、本システムのユーザは他のユーザが貼り付けた拡張記憶要素を参照することができ、これにより拡張記憶の共有が実現されている。

Ubiquitous Memories は既に述べたように、記録・再生のどちらの時点においてもユーザの能動的な操作を要する表1(a)のタイプのインタフェースである。このため、ユーザの体験と体験に関連づけられる対象物の選択と意味付けはユーザに任されており、例えばユーザにとって重要なイベントの体験を長期間関連付けておくなどのアルバム的な使用方法[14]から、一時的なメモ代わりの使い方まで種々のアプリケーションが考えられる。

## 5. まとめ

本稿では、ユーザの過去に属する体験記録を蓄積してそのユーザの記憶を外化し、それを場面に応じて検索し提示することで体験想起を支援する拡張記憶インタフェースについて、インタフェースを設計する際の観点を整理し、我々の取り組みを紹介した。人の記憶活動は多岐にわたっており、本稿で紹介した以外にも種々の適用場面を想定した拡張記憶インタフェースの提案が待たれる。しかしながら、既に述べたようにユーザが受動的にサービスを受けられるようにしようとするほど、開発した技術の適用範囲が狭くなる傾向にある。我々は現在、一般化可能な想起パターンの洗い出しのために、既に蓄積された体験記録データを体験の種類と想起パターンの観点から分類・分析を行っている。これにより、より適用範囲が広くかつユーザ負担の少ない拡張記憶インタフェースのデザイン、開発を目指している。

## 参考文献

- [1] B. Rhodes. The Wearable Remembrance Agent: a System for Augmented Memory. Proc. 1st International Symposium on Wearable Computers (ISWC97), pp.123-128 (1997)
- [2] M. Lamming, M. Flynn. Forget-me-not: Intimate Computing in Support of Human Memory. Proc. FRIENDS21: International Symposium on Next Generation Human Interface, pp.125-128 (1994)
- [3] B. Rhodes. Using Physical Context for Just-in-Time Information Retrieval. IEEE Trans. Computers, Vol.52, No.8, pp.1011-1014 (2003)
- [4] 暦本. Time-Machine Computing: 時間指向ユーザインタフェースの提案, ソフトウェア学会 WISS99 論文集, 近代科学社 (1999)
- [5] 長崎, 戸田, 川嶋: 日常生活における行動記録映像の構造化, 信学技報 PRMU, pp.109-114 (Dec.2002)
- [6] K. Aizawa, T. Hori, S. Kawasaki, T. Ishikawa. Capture and efficient retrieval of life log. Proc. Pervasive 2004 Workshop on Memory and Sharing of Experiences, pp.15-20 (Apr.2004)
- [7] B. Schiele, T. Jebara, N. Oliver: Sensory Augmented Computing: Wearing the Museum's Guide. IEEE Micro Journal (2001)
- [8] 河村, 河野, 木戸出: 運動情報を考慮した類似映像の検索～着用指向情報パートナーにおけるユーザの記憶想起支援に向けて～. 人工知能学会全国大会論文集, 3E1-02 (2001)
- [9] T. Kawamura, N. Ukita, Y. Kono, M. Kidode. HySIM: A Hybrid-space Image Matching Method for High Speed Location-based Video Retrieval on a Wearable Computer. Proc. MVA2002, pp.94-97 (Dec.2002)
- [10] L. Davenport. Order from Chaos. Three Rivers Press, NY, 2001. (平石訳. 気がつくとも機がぐちゃぐちゃになっているあなたへ. 草思社 (2002))
- [11] 上岡, 河村, 河野, 木戸出: I'm Here!: 物探しを効率化するウェアラブルシステム, ヒューマンインタフェース学会論文誌, Vol.6, No.3, pp.19-30 (2004)
- [12] 上岡, 河村, 馬場, 吉村, 河野, 木戸出: 日常生活における記憶活動支援のためのウェアラブルカラーカメラ. 情報処理学会インタラクティブ2005 (Feb.2005)
- [13] 河村, 福原, 村田, 武田, 河野, 木戸出: 対象物に『触れる』行為と記憶の遍在化による日常記憶支援. 電子情報通信学会論文誌, Vol.J-88-D-I (2005)
- [14] E. v. d. Hoven, B. Eggen. Tangible Computing in Everyday Life: Extending Current Frameworks for Tangible User Interfaces with Personal Objects. Proc. EUSAI 2004, Netherlands (2004)
- [15] J. Gemmell, L. Williams, K. Wood, G. Bell, R. Lueder. Passive Capture and Ensuing Issues for a Personal Lifetime Store. Proc. CARPE2004, pp.48-55 (2004) <http://research.microsoft.com/barc/mediapresence/MyLifeBits.aspx>

## 【略歴】

河野恭之 (KONO Yasuyuki)

奈良先端科学技術大学院大学 情報科学研究科 助教授  
1989年大阪大学基礎工学部情報工学科卒。1994年同大学院基礎工学研究科博士課程了。同年(株)東芝入社。同社関西研究所研究主務などを経て、2000年より現職。知的CAI, マルチモーダル理解, 音声対話インタフェース, ウェアラブルインタフェースの研究に従事。情報処理学会, 人工知能学会, ヒューマンインタフェース学会, 電子情報通信学会, 認知科学会, IEEE-CS, ACM 各会員。博士(工学)。