

04

# 画像認識を用いて安心を提供する情報セラピー

安部伸治 内海 章 桑原教彰

ATR 知能ロボティクス研究所

## 少子高齢化社会の課題

我々は記憶障害や認知症を持つ人々のコミュニケーションを活性化し、介護者である家族を支援する情報セラピーインタフェースの実現を目指し、要素技術ならびに要素技術を組み合わせた支援システムの開発を行っている。

高齢者(とくに認知症者)は、記憶能力や言語能力、判断認知能力の低下などが原因で引きこもりがちになり、ひいては鬱を発症する場合が多い。認知症予備軍とも言われている団塊の世代(約700万人)の高齢化が進み2015年には250万人が認知症を発症すると言われており、確実に大きな社会問題となる。著者らは、老化や認知症などにより引きこもりがちな高齢者に対して社会参加手段を創出することを目的として、コミュニケーションを支援する仕組みの研究開発を行っている。基本的な枠組みとしては、映像処理に基づいて認知症者を見守りながら、必要に応じてAV機器やPCを制御して社会とのチャンネルを開くシステムを構築する。

## 情報セラピープロジェクト

### ◆情報セラピープロジェクトの問題意識

情報セラピープロジェクトは、情報通信研究機構の研究委託「軽度脳障害者のための情報セラピーインタフェースの研究開発」として平成15年9月に発足した。

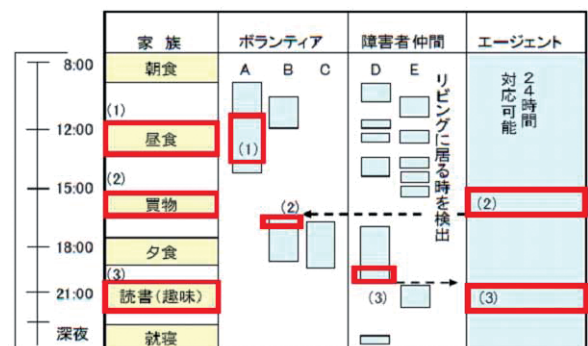
認知症をはじめとしてさまざまな原因で記憶能力や言語能力、判断認知が低下すると、日常生活では常にパートナーが必要となる。パートナーとしての役割は被介護者の家族が担うことが多く、介護の負担が問題になることが多い。ところが、介護家族の側にも日常生活上欠かせないタスクも多く、被介護者から目を離さざるを得ないこともしばしば発生する。家族にとって安心して被介護者を一人にできる時間を設けることはきわめて重要である。また、介護施設入居者や特に独居高齢者などの

場合には言語能力の低下なども影響して、外部との接触の機会も少なくなり、鬱を併発するケースも増えており社会問題化しつつある。

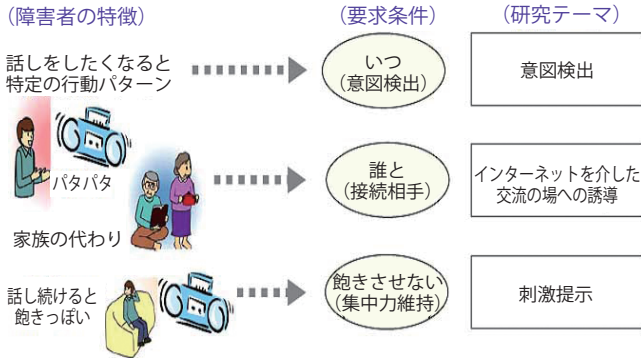
情報セラピープロジェクトは、このような課題とニーズによって生み出されたものであり、一時的に介護家族に代わる対話相手を見つけ出し、被介護者を遠隔で見守る仕組みづくりを目的とする。

図-1のように介護者である家族の目が被介護者から離れがちになる時間帯は、ネットワーク越しに対話できる相手を探し出し、ボランティアや障害者仲間とのチャンネルを開いて一時的な対話を実現する。対話相手が見つからない場合には仮想型ロボット(仮想エージェント)やこれに代わるコンテンツが対話相手を担う。

情報セラピープロジェクトが目指す仕組みは、障害者のための遠隔対話システムであるので、基本的には従来型のテレビ会議システムなどのような遠隔対話の仕組みをベースとしたものとなる。しかしながら、近年は使い勝手もきわめて単純化され利用しやすくなったとはいえ、まだ認知症などの障害者が簡単に利用できる仕組みにはなっていない。そこで、本プロジェクトでは映像処理技術やロボット技術を応用してこのような課題に挑む。



■図-1 情報セラピーの利用形態イメージ



■図-2 情報セラピープロジェクトのサブテーマ

## ◆情報セラピープロジェクトの要素技術

認知症などの脳障害者(被介護者)にとって、パソコンやAV機器の操作は遠隔対話のための大きな障壁となる。そこでまず、コミュニケーションに関する被介護者の意図を理解して適切なタイミングでコミュニティと接続させる仕組み「意図検出」が必要である。また、認知症者の特質として、少し時間が経つと他のことに注意が移りがちな傾向も強く、コミュニティ側とのコミュニケーションに注意を惹きつける手法「刺激提示」が必要である。さらに、これらの手法を利用者同士のコミュニティで適切に運用し、インターネットを介した交流の場へ誘導する手法「コミュニティ・プラットフォーム」を構築する必要がある。本章ではそれぞれのサブテーマ(図-2)について概要を述べる。

### サブテーマ1：意図検出

情報セラピープロジェクトにおける「意図検出」は、カメラやセンサを用いて被介護者を見守り、被介護者の状態を認識するための要素技術である。興味の対象を特定し、興味の持続・飽きなどを検出するほか、自己表現の機会のきわめて少ない認知症者を見守りながら外部とのコミュニケーションのきっかけづくりを支援する。本要素技術では、利用者が高齢あるいは脳障害を持った人々であることから、利用者(被介護者)に特殊なハードウェアを装着させたり、マーカーなどのようなものを装着させないことが基本設計思想である。また、日常生活の中に無理なく普及させるために、できる限り特殊な機材を用いないこと、あるいは利用する場合でもできる限り安価で入手しやすいものとするのも重要である。そこで、本要素技術では基本的にはWebカメラなどのような安価なカメラを利用した画像・映像処理技術をベースとした手法を採る。本要素技術は、被介護者の室内における生活の場が視野に入るよう複数台のカメラを設置したアンコンシャス型見守りロボットの概念に相当する生活空間を想定したものである。

### サブテーマ2：刺激提示

刺激提示は、被介護者に対する情報提示部分であり、被介護者をできる限りモニタに惹きつける役割を担う。認知症者の多くはひとつのことに長時間集中力を維持させることが難しく、介護者が安心して被介護者をひとりにできない原因となっている。そこで、サブテーマ1の意図検出と組み合わせて、飽きが検出されたときの注意の喚起ならびに集中力維持に効果的なコンテンツが必要である。本サブテーマでは、高齢者や障害を持った利用者が長時間集中力を維持するためのコンテンツならびにコンテンツを容易に制作する仕組みを構築する。

### サブテーマ3：コミュニティ・プラットフォーム

コミュニティ・プラットフォームは、ネットワークを介して被介護者の生活の場から外部へと開かれる対話のためのチャンネルの役割を担う要素技術である。

できる限り既存のネットワークや市販端末を利用して、特殊なサービスアプリケーションはその都度サーバからダウンロードさせる方式とする。また、高齢者や障害を持った利用者を想定しているため、できる限り難しい端末操作から利用者(特に被介護者)を解放する設計思想とする。

基本構成はネットワークとサーバ、端末群からなり、サーバは端末間のコンテンツの同期共有と、介護者端末による被介護者端末の制御を担う。

## 意図検出

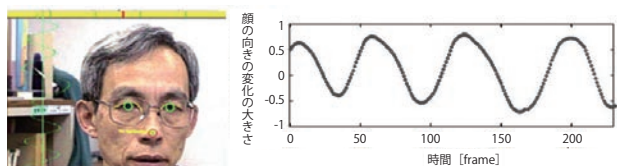
本章では、映像処理をベースに認知症者を見守り、認知症者のさまざまな意図を検出する手法について述べる。検出する意図については、興味の対象の特定、興味の持続・飽きの検出、行動意図の検出に基づく行動支援などを対象とする。

### ◆集中度モニタリングのための人物動作解析

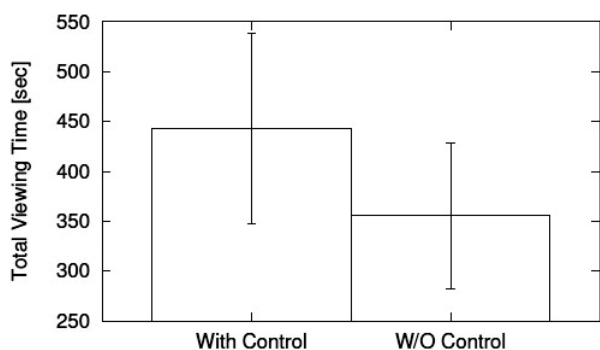
近年のAV機器や遠隔会議システムは、従来に比べ格段に使いやすくなった。しかし、残念ながらまだ高齢者や脳障害を持った方が簡単に利用できるものとはなっていない。本節では、映像処理に基づく被介護者のTVモニタへの集中の度合いの検出手法とAV機器制御への応用について述べる。以下、集中度推定の手がかりとなる顔の向きの抽出、体動(拍動作)の抽出について詳しく述べる。

#### 顔の向きの検出

顔の向きを検出することにより、TVモニタに対して集中しているか、あるいは飽きによりモニタから注意が逸れているかを検出する。さらに、飽きを検出することによりAV機器のチャンネルを制御してコンテンツを切り替え、注意を引き戻すアプリケーションに応用した。本手法は、健常者が日常行うチャンネルのザッピングに相当



■ 図-3 顔の向きを検出



■ 図-4 顔の向きを検出によるコンテンツへの集中の効果

する行為を、アンコンシャス型ロボットに代行させるものである。

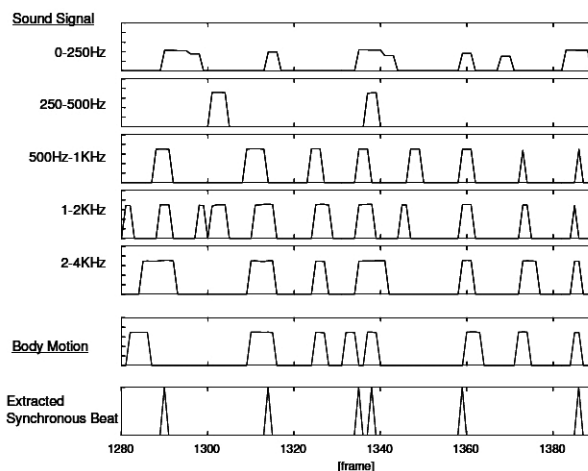
顔を撮影したビデオ画像から目、鼻先位置を実時間(30frame/sec)で検出し、顔の向きの変化を推定する。顔の検出は6分割矩形フィルタ(SSR-filter)による顔候補抽出とSVMによる顔判定処理による。首を左右に振った際の、顔の向きの変化の大きさを図-3に示す。図の縦軸の値は0が正面を向いている時であり、±1は検出限界をあらわしている。このときの検出画像例を図-3左に示す。

我々は、本手法を用いて認知症者のTVモニターへの集中度を検出し、一定時間以上モニターから顔の向きが逸れるのを検出すると、TVのチャンネルを変更し認知症者が長時間集中できる環境を構築した<sup>4)</sup>。

図-4に本手法によるコンテンツ切り替えの効果を示す。ここでは、テレビ視聴中に顔の向きが一定時間テレビから逸れている場合にコンテンツを切り替える制御を行い、制御を行わない場合と総視聴時間を健常者6名について比較した。図中左は切り替え制御ありを右は切り替え制御なしの結果を示している。切り替え制御ありの場合に有意に視聴時間が長くなっており、顔の向きに応じたコンテンツ切り替えの効果が確認された。

### 体動の検出

認知症者の特質として、症状が進行して言語能力が大幅に低下した場合でも、音楽を楽しむ能力は多くの場合残されており、知っている曲目や馴染みやすいメロディ



■ 図-5 拍動作の検出例

に出会うと、比較的重度の認知症者でも体動によるリズム表現を示す場合が多い。このように、顔の向きがTVモニターから逸れていても、コンテンツに集中している場合がある。BGMに対してリズムを刻んでいる場合のような“拍動作”が検出される場合には、顔の向きに関係なくコンテンツに対して集中していると判断してよい。

本項では、認知症者の“拍動作”を検出する手法について述べる。

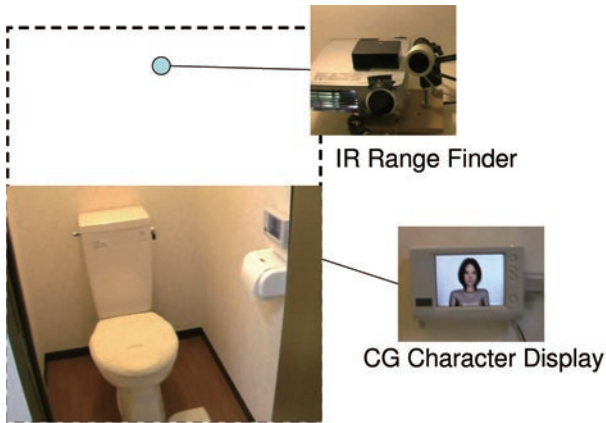
まず、フレーム間差分処理により体動パターンを検出する。ここでは時刻 $t$ における動きの大きさを差分により得られる移動画素の画素数 $N_t$ で表す。ユーザが手足などで拍子をとる“拍動作”は画素数 $N_t$ が極小となる時刻として検出できる。

$$\text{beat} = \begin{cases} 1 & \left( \frac{dN_{t-1}}{dt} < 0 \text{ and } \frac{dN_t}{dt} \simeq 0 \right) \\ 0 & (\text{otherwise}). \end{cases}$$

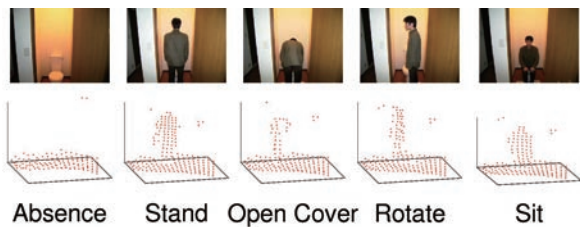
図-5は、この方法で検出された拍動作の一例を示している。本手法により手・足・首などによる拍動作が検出可能である<sup>5)</sup>。

### ◆人物動作認識によるトイレでの行動指示

比較的症状の重い認知症者では、日常のタスク遂行が困難で大きな介護負担の原因になる場合がある。特にトイレでの手順誤りはとりわけ大きな問題で、介護コストを大きく押し上げる要因の1つである。そこで我々は人物の姿勢変化を検出して適切な次の行動を指示する手法



■図-6 トイレにおける行動指示システム



■図-7 レンジファインダによる姿勢検出結果

を開発した(図-6)。

トイレでの手順は比較的単純なタスクからなり、タスクを構成する各々の状態は、人物の姿勢と一意に対応付けることができるため、比較的容易に効果的なアプリケーションを構築することが可能である。

トイレはプライバシーが特に重要視される環境であることから本システムは、通常のビデオ撮影ではなく、赤外線ドットパターンを照射し、ドットパターンの変化から人物の姿勢を推定する手法を採用した(図-7)。

推定された人物の姿勢とタスクのマッピングを行い、手順テーブルと照合して正しい手順をCGキャラクターが指示する。

## ◆視線の検出

日常生活における人の注意・興味の対象を特定し、認知症者や脳障害者の注意・興味の変化に応じた適切な情報提示・コンテンツの提供を行うための手がかりとして、人の視線情報を非接触で計測する手法の開発を行っている<sup>8)</sup>。

本手法は、利用者に負担を強くない非装着型(利用者に装置やマークなどの装着を必要としない方式)で、かつ、特殊な装置構成を必要としない単眼カメラ方式として開発した。

本手法ではまず、カメラからの入力画像列に対して集中度モニタリングで用いた6分割矩形フィルタによる顔候補探索とSVM(サポートベクトルマシン)による識別



■図-8 視線推定結果

処理を適用し、画像中の顔位置(両目位置)を特定する。次に得られた両目位置から虹彩中心、鼻位置を推定する。次に、抽出された鼻、目(虹彩)の位置を利用し、文献9)の特徴点抽出法を用いて顔領域内の画像特徴点を検出する。さらにこれらの特徴点と虹彩中心の関係から顔特徴点と眼球中心の相対関係、眼球半径、虹彩半径などのパラメータを推定する。視線推定過程では、推定された眼球中心位置と虹彩中心位置から視線方向を決定する。図-8に本手法を用いた視線推定の結果を示す。

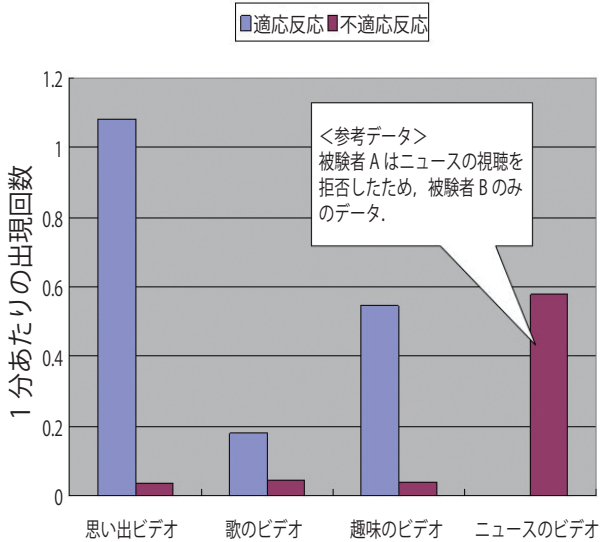
## 刺激提示

介護者の生活に最も制約を課す原因として、1つのことに対する認知症者の集中力が維持できずに問題行動を誘発し、安心してひとりにすることができないことが挙げられる。そこで、意図検出(サブテーマ1)と組み合わせ、認知症者に対して必要に応じて注意を喚起し、集中力を維持するために効果的なコンテンツが必要である。認知症者の多くは近年の記憶に比べて昔の記憶は比較的多くが残っており、これを利用することで長時間集中力が維持される傾向がある。本章ではこのような効果を引き出すコンテンツとその応用システムについて述べる。

## ◆思い出ビデオ

認知症者の問題行動抑制のために心理的な安定を引き出す目的で、思い出ビデオを用いる手法が提案され、その有効性が臨床現場で報告されている<sup>1)</sup>。思い出ビデオは、デジタル化した昔の写真アルバムから作成したスライドショーに映像エフェクトを施した上でBGMやナレーションを付加し、視聴者にとって魅力的なコンテンツとして編集したものである。本コンテンツは個人の思い出に即した内容であるので、通常のTV番組や従来回想法に用いられてきた一般的なコンテンツに比較してより集中度が維持できるという特色がある。認知症者数名によるコンテンツ比較評価実験(図-9)でも、他のコンテンツに比較して好意的な反応や発話数が多いことが観測された。本研究では、思い出ビデオを認知症者の発話促進に役立てるとともに、患者を思い出ビデオに集中させ、介護者が介護の手を離す時間帯を設ける方法として用いる。

感情表出



発話時間

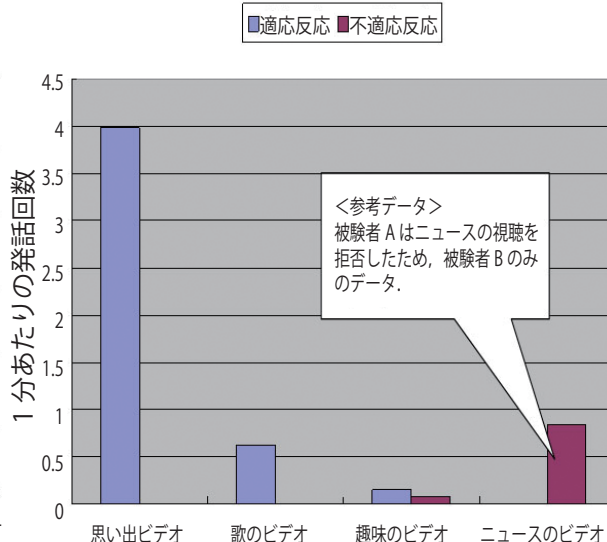


図-9 思い出ビデオの効果



図-10 思い出ビデオ制作支援ツール

エフェクト、トランジションが使用される条件	エフェクト、トランジション	使用される情報
写真中の2つのリージョン間の距離が閾値以下のとき	Zoom, Pan	写真中の人物領域
写真中の2つのリージョン間の距離が閾値以上のとき	RegionOverlap, Transition	
連続する2枚の写真がモノクロからカラーに変わるとき	Mono2Color	写真の色調
視聴者に思い出深い被写体に対応するリージョンにだけエフェクトを適用するとき	Zoom, Pan, RegionOverlap, Transition	写真中の人物情報
連続する2枚の写真の年代差が閾値以下のとき	PhotoOverlap, Transition	撮影日付
連続する2枚の写真の年代差が閾値以上のとき	RotateTransition	
シナリオが一貫しているとき	PhotoOverlap, Transition	撮影日付, 撮影された出来事
連続する2枚の写真の年代差が閾値以下で、かつシナリオ転換点であるとき	RotateTransition	

表-1 思い出ビデオ制作支援ツールに実装したレンダリングオントロジ

◆思い出ビデオ制作支援ツール

ビデオコンテンツは制作コストがかさむため、本プロジェクトでは思い出ビデオ制作支援ツール(図-10)を構築してコスト削減を図っている。

認知症者は以下の知見に示されるような視覚的認知に関する特性を有している。

<知見1> 認知症者は一般に画面全体に対する視覚的な注意が行き届かないことが臨床上しばしば観察される<sup>2)</sup>。

<知見2> 視覚的保持という、前の画像の刺激が残存することで視覚的な認知が混乱する現象が観察される<sup>3)</sup>。

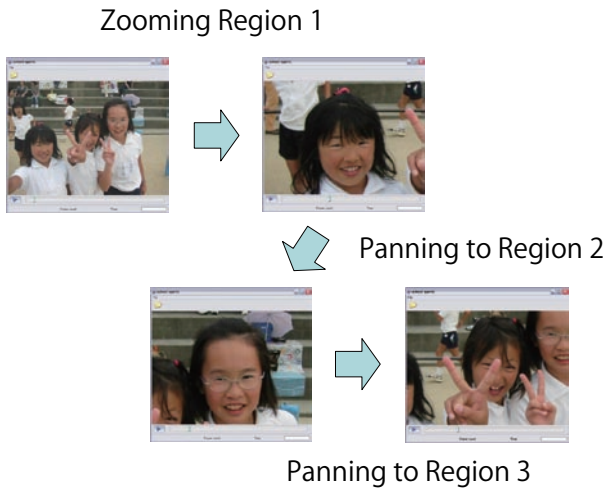
これらの知見を熟知したビデオクリエイター3名によるヒアリングにより、下記を始めとする多くのビデオ制作

ノウハウを抽出した。

<<ヒアリングによるノウハウ(抜粋)>>

- ①知見1から、写真の中で注意してもらいたい人物へのパン・ズームが視覚的認知に有効であると考えKenBurns効果を利用した。また、パンを使用する際には、時間軸を意識して写真の左から右へと一定方向に流すようにした。
- ②知見2を意識して、写真中の次の人物への切り替わりに、パンの代わりにフェードアウト・インを使用する場面がある。

これらのノウハウを元に、表-1のようなレンダリングオントロジにまとめ、それを実装した思い出ビデオ



■ 図-11 思い出ビデオの再生例

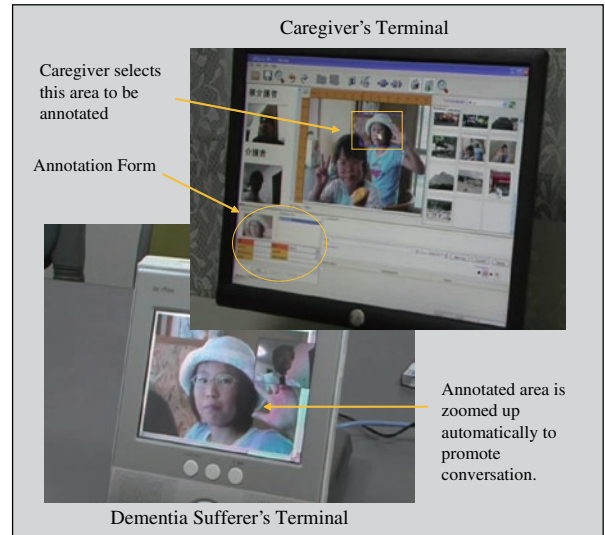
作成支援ツールを構築した<sup>6)</sup>。図-10のインタフェースにより写真にアノテーション（インデクス付与）を行うことで、レンダリングオントロジに沿った思い出ビデオが生成される。図-10のインタフェースでは、撮影日付やデスクリプションなどの写真全体に対するインデクス（Dublin Core）を入力するほか、写真の人物領域を指定して人物名や人物間の関係など（FOAF）を入力する。図-11は図-10の3つの領域 Region1～Region3が指定されたときに（この場合、領域間距離は閾値以下）、表-1のレンダリングオントロジを適用して Region1にZoomしPanでRegion3まで順次移動するようなビデオが生成された例である。なおFOAFの値は写真素材の検索などに使用する。

## コミュニティ・プラットフォーム

### ◆ 思い出ビデオ共有&対話システム

普及型TV電話端末（フレッツフォン）に思い出ビデオを配信し、介護者端末からの制御によってコンテンツを共有しながら対話を実現するシステム（図-12）を構築した<sup>7)</sup>。

思い出ビデオの制作にはアノテーション情報が必要であり、基本的には写真の所有者に対するインタビューが必要である。しかしながら、物理的な移動を伴うインタビューは、思い出ビデオの制作コストを押し上げる要因になってしまうため、これを遠隔で行うための仕組みとして開発を行った。介護者端末（PC）でアノテーション作業を行うと、被介護者端末に同じコンテンツが表示され、被介護者端末上でのズーム・パンなどの操作を介護者端末から制御できる。フレッツフォンにはWebブラウザが搭載されており、介護者端末からサーバを介して被介護者端末へ、静止画像とSMILで記述したレン



■ 図-12 思い出ビデオ共有&対話システム



■ 図-13 遠隔傾聴トライアルサービス

ダリングコマンドならびにJavaScript形式の簡易SMIL Playerを送信して思い出ビデオを再生する。本ツールにより、遠隔で写真を共有しながらインタビューとアノテーション作業を実施することができる。

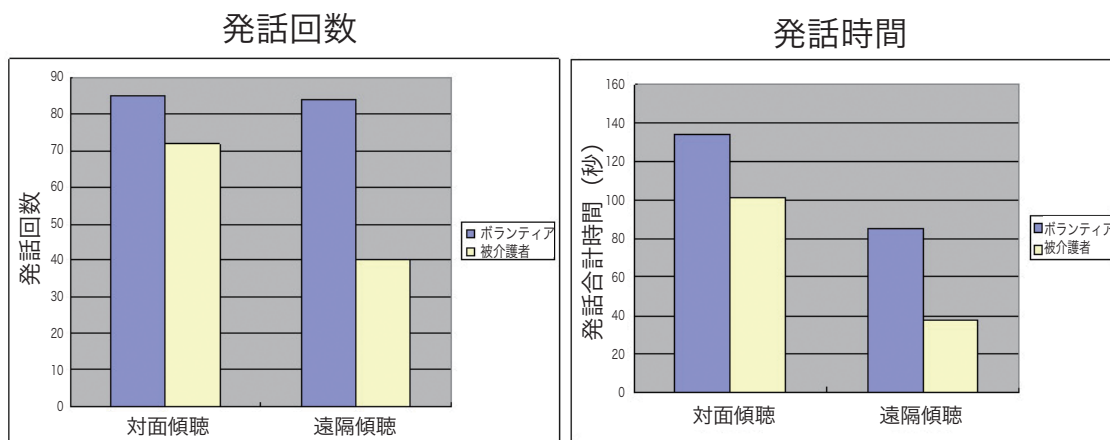
### ◆ 遠隔傾聴トライアルサービス

思い出ビデオや写真素材をPCとフレッツフォンあるいはフレッツフォン同士で共有し、遠隔で傾聴活動を実施する環境（図-13）を構築し、トライアルサービスを実施した。実施期間は2006年4月一杯である。

本トライアルには、(株)ベストライフ（有料介護老人ホーム）、NPO法人ホールファミリーケア協会（傾聴ボランティア）、ATR（技術提供）、NTTグループ（インフラならびに端末提供、コンテンツ制作）の各社が参加した。

トライアル実施期間中ならびに実施後の調査・分析により、遠隔傾聴ならびにコンテンツ共有の効果に関して下記のような結果を得た。

（遠隔傾聴の効果）トライアルに参加した傾聴ボランティア6名に対して「対面傾聴を100として、対面傾聴に比べ遠隔傾聴はどの程度利便性があると感じるか」という問いに対し、平均90%の利便性と回答。また、1被験者の例であるが、傾聴開始後5分間のビデオ解



■ 図-14 対面傾聴と遠隔傾聴の比較

析により図-14の結果を得た。

(コンテンツ共有の効果1) 共有コンテンツは遠隔傾聴のために利用者の9割以上が必要性を実感

- 話のきっかけづくりに役立つ【傾聴ボランティア全員】。
- コミュニケーションの幅を広げるのに役立つ。
- 共有コンテンツのズーム・パン機能は、介護施設入居者の注意を引く作用が働き刺激となって集中力が増すとともに、見やすくなるため好評。

(コンテンツ共有の効果2) SD法による遠隔傾聴の利用感に関する主観評価の結果、「安らぎ感」、「包容感」、「洗練感」、「信頼感」の順で寄与しており、すべての因子で共有コンテンツ共有ありの方がいい場合より優れる。

## まとめ

本稿では、情報セラピープロジェクトの概要ならびに最近の成果について述べた。自己表現や意思表示の機会が健常者に比較して少ない認知症者を見守る仕組みの構築のためには、視線情報はとりわけ重要であると考えている。視線情報のみで認知症者の内面を押し量することは困難であるが、視線情報とコミュニケーションロボットの機能を利用して、徐々に本人の意思を引き出し、最終的に適切な情報提示を行う、あるいはコミュニケーションのチャンネルを開く手法の開発などを現在進めている。本稿で述べたさまざまな手法を融合したトータルなシステム構築を今後進めてゆく予定である。

### 参考文献

- 1) Kuwahara, N., Kuwabara, K., Abe, S., Yasuda, K. and Tetsutani, N.: Semantic Synchronization: Producing Effective Reminiscence Videos, 4th International Semantic Web Conference (ISWC2005) Demo Papers (2005).
- 2) McCarthy, M.: Fractionating the Neglected Space: The Relevance of Reference Frames for Defining Left and Right in Spatial Neglect, Cortex, Vol.38, pp.465-477 (2002).
- 3) 山鳥 清: 神経心理学入門, 医学書院 (1985).

- 4) 石田 彩, 内海 章, 川戸慎二郎, 桑原和宏, 渋谷 雄: 高次脳機能障害者のビデオ視聴行動の観察と情報セラピーインタフェースのための映像コンテンツ切り替え法, ヒューマンインタフェース学会論文誌, Vol.8, No.1, pp.157-166 (2006).
- 5) Utsumi, A., Kawato, S. and Abe, S.: Attention Monitoring based on Temporal Signal-Behavior Structures, In Proc. of Computer Vision in Human-Computer Interaction (ICCV2005 workshop), LNCS 3766, pp.100-109 (Oct. 2005).
- 6) 桑原教彰, 桑原和宏, 安部伸治, 須佐見憲史, 安田 清: 写真のアンテーションを活用した思い出ビデオ作成支援-認知症者への適用と評価-, 人工知能学会論文誌, Vol.20, No.6 SP-B, pp.396-405 (2005).
- 7) Kuwahara, N., Abe, S., Yasuda, K. and Kuwabara, K.: Networked Reminiscence Therapy for Individuals with Dementia by Using Photo and Video Sharing, Proceedings of The 8th International ACM SIG ACCESS Conference on Computer and Accessibility (ASSETES2006), pp.125-132 (2006).
- 8) 内海 章, 山添大丈, 安部伸治: ユーザ意図検出のための単眼カメラによる視線推定, ヒューマンインタフェースシンポジウム 2006, 3522, pp.1143-1146 (2006).
- 9) Shi, J. and Tomasi, C.: Good Features to Track, Proc. CVPR94, pp.511-518 (2001).

(平成 18 年 11 月 15 日受付)

※本研究は情報通信研究機構の研究委託「軽度脳障害者のための情報セラピーインタフェースの研究開発」により実施した。

### 安部伸治

sabe@atr.jp

1984年北海道大学工学部原子工学科卒業。1986年同大学院工学研究科修士課程修了。同年日本電信電話(株)入社。ヒューマンインタフェース研究所・サイバソリユーション研究所にて、画像・映像ハンドリングなどに関する研究に従事。現在ATR知能ロボティクス研究所にて認知症介護支援システムに関する研究に従事。日本物理学会、ヒューマンインタフェース学会、電子情報通信学会各会員。

### 内海 章(正会員)

utsumi@atr.jp

1991年大阪府立大学工学部金属工学科卒業。1993年大阪大学基礎工学研究科情報工学修士課程修了。同年、ATR通信システム研究所に入社。画像処理、ヒューマンインタフェースの研究に従事。現在、ATR知能ロボティクス研究所ならびにATRメディア情報科学研究所主任研究員、博士(工学)。電子情報通信学会、ヒューマンインタフェース学会、システム制御情報学会各会員。

### 桑原教彰(正会員)

kuwahara@atr.jp

1987年東京大学工学部精密機械工学科卒業。同年、住友電気工業(株)に入社。2002年ATRメディア情報科学研究所に転出。以来、五感メディア、協調メディアの研究に従事。現在、ATR知能ロボティクス研究所主任研究員、博士(工学)。