

ぱらぱらアニメによる 体験データの要約・編集支援システム

小関 悠 角 康之 西田 豊明 間瀬 健二

本稿では研究発表会や博物館見学といったイベント空間において取得可能な体験データを、ユーザーが閲覧・編集するためのシステムを提案・実装する。種々のカメラやセンサー機器の発達により大量の取得が可能となった体験データを、ユーザーの扱いやすい形にすることで、その編集や共有を促すことが狙いである。システムは大きく二つの部分に分けられる。一つは体験データを自動的に要約してユーザーに提示するシステムであり、特にセンサー情報を用いることで映像データを「ぱらぱらアニメ」、すなわち、シーンを表現する複数枚の特徴的なスナップショットのセットへと変換する手法について述べる。もう一つは要約された体験データの鑑賞・編集システムであり、こちらでは「ぱらぱらアニメ」の特性を生かし漫画的なレイアウトを組むことで体験データを好みの観点で観賞・編集が出来ることを中心に述べる。本システムは体験データの閲覧や編集へのアクセシビリティを高めるため、Web アプリケーションとして実装した。

This paper proposes a supporting system for browsing, editing and sharing experiences captured in the events such as research exposition and museum touring. The aim of the system is to facilitate users to edit and share their experience data—sets of video data captured in multiple viewpoints—by enabling the users to manipulate the data easily. Our system consists of two parts. The first part is to show the users automatically summarize the experience data by converting segmented video scenes into so-called animated cartoons, i.e., sets of snapshots representing the scenes. The second part is to facilitate the users to edit their experience data by showing the whole experience data with layout of scenes like comics. The system is a Web-based system for the users to easily browse and edit their experiences.

1 序論

1.1 研究の背景

ビデオカメラや各種のセンサー機器が小型化し、また安価になったことから、これらを組み合わせることで体験データを取得する、体験記録システムが複数提案されている。ここでいう体験データとはある空間・時間における詳細な体験の記録のことであり、映像・写真・音声・センサー情報・文章など、様々なメデ

アによる複合的な構成が考えられる。体験データには、例えば日記のように自身が鑑賞することにより体験時を振り返ったり、あるいは卒業写真のように第三者と観賞することにより体験時には気付かなかったような新しい発見を得ることが出来るといった機能がある。また、結婚式のビデオを知人に見せる時のように、第三者へ体験を紹介する際のメディアとしても利用出来るほか、旅行先で撮影した写真を後日に持ち寄って体験者同士で振り返りを行うなど、体験データには様々な形での観賞・共有が考えられる。

体験を簡単に記録する環境が整いつつあるということは、これまでになく大量の体験データを取得する機会が生まれつつあるということでもある。例えば携帯電話のカメラは数百枚の写真を簡単に撮影することが可能であり、blogのように体験を文章として記録し公開する文化も広まってきている。しかしながら、得られた大量のデータをどのように鑑賞する

System of Summarization and Edit Experiences by Using Animated Cartoon.

Yu Koseki, Yasuyuki Sumi, Toyooki Nishida, Kenji Mase, 京都大学情報学研究科/ATR メディア情報科学研究所/名古屋大学情報連携基盤センター, Graduate School of Infomatics, Kyoto University/ATR Media Information Science Laboratories/Infomation Technology Center, Nagoya University.

コンピュータソフトウェア, Vol.24, No.3 (2007), pp.41-50.
[論文] 2006年2月18日受付.

かという問題については、依然として旧来的なスタイルに留まっている。すなわち画像は時系列順に一覧で表示する、映像は早送りや巻戻しで目視確認する、といったスタイルである。これでは膨大な体験データを前にした場合、せっかくそれらを容易に得ることが出来たにも関わらず、その確認だけで非常に時間を費やすことになってしまう。

特に大規模な体験記録システムにおいては、このことは顕著な問題となる。多くの体験記録システムでは、ウェアラブルカメラなど複数のビデオカメラで常時撮影を行うので、その帰結として大量の映像データを取得する。こうした量を対象とする場合、目的の場面を探したり、目視で映像データ全体を確認したりすることが非常に難しいことは明らかである。

さらに問題を難しくさせている点として、これまでに提案されてきた体験記録システムでは、観賞に専用のハードウェアやソフトウェアを必要とするものが多いということが挙げられる。例えば特別に設けられた条件の整った部屋でのみ、あるいは特定のヘッドセットを装着した時のみ、体験の振り返りが行えるというようなものである。結果、興味深い体験データを得ることが出来たとしても、そもそも鑑賞出来るのがごく少数に留まってしまい、本来的な用途の一つであるはずの体験共有が非常に困難となる。

1.2 研究の目的と概要

このような背景から、膨大な体験データを対象とした際も、ユーザーが容易に取り扱うことが出来るような観賞・編集システムが求められている。また、データを第三者とも共有・鑑賞することを考えると、システムは特別なハードウェアやソフトウェアを必要としないことが望ましい。

本稿では以上のような要求を満たすシステムを提案し、その実装を説明する。本システムは体験キャプチャシステム [1][2][3][4][7] を体験記録システムとして背景に用い、得られた情報を利用して映像データの要約を自動的に行う。本システムにより、体験データを分かりやすい形で一覧したり、好みの観点で編集することが出来る。また、データはインターネット上で簡単に配布出来る上、観賞に特別なソフト

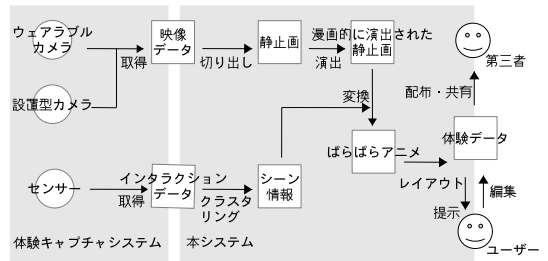


図1 体験キャプチャシステムと本システムの動作

ウェアやハードウェアを用いないため、標準的な環境で誰でも簡単に共有することが可能である。

体験キャプチャシステムと本システムを用いた体験データの利用イメージは以下ようになる。

1. 体験データを記録する

記録端末としてウェアラブル型と設置型の二種類を用いる。展示会や説明会など、参加者と撮影範囲があらかじめ定まっているような、閉鎖的な環境での利用に向いている。

2. 記録者がデータを編集する

どのような体験データにおいても、最初に振り返りを行うのは体験者本人であろう。本システムではまず体験者が容易に体験を見直すことが出来るよう、コンピュータ・アプリケーションを用いて要約したデータを提示する。アプリケーションは様々な観点で体験データを鑑賞・編集が行えるような機能を充実させる必要がある。

3. 編集したデータを配布・共有・交換する

配布・共有する相手には二種類が考えられる。同じ体験の参加者など記録時点の環境情報を知る者と、そうした情報を全く知らない者である。前者には単体で観賞・把握可能なメディアとして、後者には記録者本人が体験を紹介する際のメディアとして、利用出来ることを目標とする。

体験キャプチャシステムと本システムによる動作概要は図1の通りである。体験キャプチャシステムの各種カメラ、センサーなどから映像データやインタラクション・データを取得し、前者は静止画に切り出して漫画技法を用いた演出を施す一方、後者はクラスタリング手法を用いてシーン情報に変換する。その後、これらを併わせてばらばらアニメとし、レイアウトを

行ったものを体験データとしてユーザーに提示する。ばらばらアニメとは複数の静止画を順に切り替えていくことでシーンの様子を簡易に見せるための手段で、ここでは場の臨場感といった情報を損わずに映像データよりも短い時間で観賞出来るという利点がある。

1.3 本稿の構成

本稿の構成は次の通り。まず、関連研究を簡単にまとめ(第2章)、背景技術である体験キャプチャシステムの概要について簡単に説明する(第3章)。以降はシステムの動作順序に従い、初めに得られたデータをシーンという概念を用いながら要約する手法について述べる(第4章)。次にセンサー情報を用いて映像データから特徴的な静止画を切り出し、漫画技法を用いて背景情報を可視化させる手法、得られた静止画をばらばらアニメへと変換する方法について説明する(第5章)。そして要約したデータを、ユーザーに提示する方法について述べ(第6章)、その後、得られた体験データがユーザーが好きな形に編集出来ること、これにより体験が様々な観点で見直せることについてまとめる(第7章)。最後に、この過程で編集された体験データが第三者へ容易に配布可能であることと、その利点について述べる(第8章)。

2 関連研究

体験データを扱う研究としては、本システム同様にウェアラブル機器を用いるもの[9]のほか、データをRFIDタグを装着したオブジェクトに結び付けることで直接的に整理を行うもの[10]、シンプルなカード型の機器を提案するもの[11][12]などが挙げられる。後者は特に、本システムで唱えるところの「シーン」を物理的に扱うものとして、共通点は多い。また、コミックダイアリ[6]も、非常に近い観点を持った研究である。ただしこの場合、提供されるコミックは既存のテンプレートに従ったものであり、完全に自動的に生成されるため、本システムのようにその内容に対してユーザーが編集を加えたりすることは出来ない。また生成される要約にはユーザー自身が記録した映像や音声が存在しないため、「自分自身の体験」と捉え、振り返るにはリアリティに欠ける。



図2 ウェアラブル型・設置型端末

一方で取得可能なメディア・データのうち、一部を特徴的に利用する研究もある。具体的にはセンサーデータを映像データの要約のためだけに用いることで、体験のサマリービデオを提供するもの[8]、音声のみを用いて空間の雰囲気や伝えたり、行われたプレゼンテーションを自動的に再利用するもの[5]などである。本システムではばらばらアニメという特徴的なレイアウトを採用することで、こうした従来のシステムと同等以上の一覧性の高さや扱いやすさを目指した。

3 体験キャプチャシステム

体験キャプチャシステムは、体験記録システムとして提案・実装されているもので、展示会や説明会など、主に閉じた空間におけるイベントでのユーザーの体験を各々の視点で記録するものである。このシステムには、イベントに参加するユーザーが装着するウェアラブル型と、ポスターや空間の各所に備える設置型の二種類の端末が存在する(図2)。それぞれの端末はカメラ・赤外線IDタグ・赤外線センサー・マイクといった機器で構成されている。

システムは映像や音声、センサー情報の常時記録が行えることに加え、個々のセンサーとタグの認識情報から「見る」「話す」といった基礎的なインタラクションについて「誰が」「いつ」「何に対して」行ったかを自動的に類推し、データベースに蓄積することが可能である。このようにして得たデータのことを、以下ではインタラクション・データと呼ぶ。

4 データの自動要約

4.1 概要

本システムでは、体験とは複数のシーンの列なりであると仮定している。ここでいうシーンとは、体験を

時間的に、その意味に応じて適切に分断したものと定義出来る。我々は映画や小説などにおいて、時間の経過や主体や客体の変更、位置の移動に応じて、シーンが変わったと認識出来る。そして「あのシーンは重要だった」「最初のシーンは最後のシーンの伏線だった」というように振り返ることが出来る。普段の体験にも複数のシーンが内在されていると考えれば、それらを捉えることで振り返りが容易になるはずである。

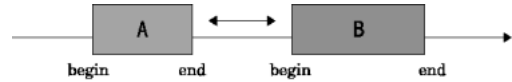
このシーンという概念を用いると「体験データを適切に要約する」とは、適切にシーン分けし、それぞれのシーンの重要度を適切に定め、重要なシーンは大きく、そうでないシーンは小さく扱うことと言える。

一方でシーンは「誰かと話す」「何かを見る」といったインタラクションの組み合わせとも言える。意味のあるシーン切り替えには、前述の通り客体の変更や位置の移動などを伴うことがほとんどだからである。

以上より、インタラクションを適切に組み合わせしていくことで、体験データからシーンを推定していくことが可能と考えられる。本システムでは体験キャプチャシステムを背景に用いるため、体験中のインタラクションについては、自動的に得ることが出来る。ただし、得られた個々のインタラクション・データには重み付けがされておらず、また前後のインタラクションとの関係性が明らかにされていない。そのため、それぞれのインタラクションがどのようなシーンを構成するものであるのか、構成されたシーンはそれぞれどれくらいの重要度を持つのか、推測する必要がある。

4.2 距離を用いたクラスタリング

シーンを推定するために距離という概念を導入する。距離とはインタラクション間の関連性であり、短いほど同じシーンに入れるべきであることを意味する。具体的には、インタラクション間の時間差を用いる。すなわち、短い間に起きたインタラクションは同じシーンに入れるという考え方である。ただしインタラクションの対象が異なる場合は時間差に1以上の定数 p をかけたものとする(図3)。同じ相手とのインタラクションは、違う相手とのインタラクションより p 倍結びつきが強いということである。 p を大きく取ると、それだけ同じ相手とのインタラクションを重要



$$f(x) = \begin{cases} p \times (B.begin - A.end) & (A \neq B) \\ B.begin - A.end & (A = B) \end{cases}$$

図3 距離

表1 インタラクションの点数化

点数	種類
0点	いる / 見られる
1点	見る
2点	見つめ合う / 話す / 話される
4点	話し合う

視する。

次に距離を変数に階層的クラスタリング・アルゴリズムである群平均法を導入し、全てのクラスター間の距離があらかじめ定めた閾値を越えるまでクラスタリングを行う。閾値は求める要約によって変化するが、大きいほど、シーンの数が少なくなる。

こうして得られた個々のクラスターはインタラクションのまとめり、すなわちシーンと見なせる。個々のシーンの重要度は、クラスター(=シーン)の大きさと近似的に見なせるだろう。たくさん話したシーンの方が、少しだけ話したシーンより重要であると一般的に考えられるからである。また同じ長さであれば、誰かを見ていたシーンより誰かと話していたシーンの方が、受動的に「見られ」「話される」シーンより能動的に「見て」「話す」シーンの方が、それぞれ重要である可能性が高い。以上よりシーンの重要度は、それが含むインタラクションの種類に応じて点数化したものの合計と考えられる。この点数も要約に求める観点によって調整が必要だろうが、一つの例として表1を挙げる。この場合、「Aさんと話して」「ポスターBを見る」シーンは3点である。

5 映像データのばらばらアニメ化

5.1 概要

映像データは原則的に、撮影時間と同じだけ鑑賞時間が必要となる。もちろんノイズや不要な場面をカッ

トしたり、早送りやメタタグを用いた頭出しといった機能を用いることで、時間をある程度削減することは可能である。しかしいずれにせよ一覧性が低く、検索に不向きなメディアであることに違いはない。

こうした問題を解決するため、映像データから特徴的なカットを静止画として切り出し、まとめて「ばらばらアニメ」に変換する。そして得られたばらばらアニメを映像の代用として用いる。映像データからちょうどハイライトと呼べるようなカットだけを選び抜き、連続的に見せることが理想である。加えて、切り出した静止画にはそれぞれの時点のインタラクション情報を提示するため、漫画技法を用いた演出を行う。

ばらばらアニメを映像の代用として用いることの一の利点は、必要な鑑賞時間を大幅に減らすことにある。この結果、例えば一日の出来事を大雑把に把握したり、目的のシーンをあやふやに探し出すというような、膨大な映像データが対象である場合は非常に難しい作業を簡単にこなすことが可能となる。

5.2 静止画の切り出し

ばらばらアニメへの変換のためには、まず映像データから特徴的なカットを静止画として切り出す必要がある。ユーザーの行動についてはインタラクション/シーンといった概念を用いて要約を行っているの、ここでは個々のインタラクションに応じて静止画を一枚選び出すことにする。すなわち「Aさんと話す」インタラクションがあった場合はそれを示すような静止画を映像データから一枚切り出す、といった格好である。とは言え、無数のフレーム(静止画)の集合である映像データから、誰と何をしているのか一目で理解出来るような「良いカット」を自動的に切り出すのは難しい。良いカットの条件としては、まずインタラクションの対象が静止画にはっきりと捉えられているものが挙げられるが、ウェアラブル・カメラを用いた場合、むしろインタラクションの対象、すなわち見ているポスター、話している相手などを捉え続けていることはまれである。すなわち映像データからランダムに静止画を切り出すと、高い確率で悪いカットを引き当ててしまうことになる。図4は良いカットと悪いカットの例である。



図4 良いカット・悪いカット

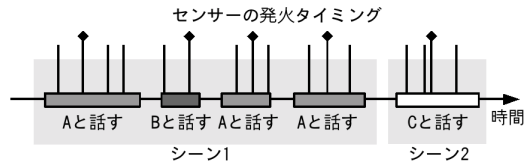


図5 タグ認識と静止画の切り出し

この問題を解決するため、体験キャプチャシステムが事前に得ている赤外線センサーがタグを認識するタイミングを利用して、静止画の切り出しを行う。図5はあるユーザーのインタラクションとタグ認識例である。体験キャプチャシステムでは一つのインタラクションが複数のタグ認識結果から類推されている。例えば「Aさんのセンサーが一定期間内にポスターXのタグをn回捉えた」「AさんがポスターXを見ている」といった具合である。そこで、図中「」のような、なるべく個々のインタラクションの中心に存在するタグ認識時刻に合わせて映像データから静止画を切り出す。これはセンサーがタグを認識した瞬間、同方向に備えられたカメラも当然、インタラクション対象を視野に捉えているからである。実際、これによりほぼ確実にインタラクション対象を捉えた静止画を切り出すことが可能となった。

5.3 漫画技法を用いた演出

加えて人口密度の高いイベント空間などでは、一枚の静止画に複数の人物が写り込むことが多々ある。こうした場合、映像データでは自明であったはずの、インタラクション対象が誰/どれであったのか、といった背景情報が切り出された静止画においては明確ではなくなってしまう。図6左はそのような例である。加えて、静止画では当然、話していたのか見ていたのかといった、インタラクションの種類も把握出来ない。



図 6 二人を捉えた静止画

これを解決するため、漫画技法を用いる。吹き出しやコマ割り，集中線，あるいは汗や涙の記号に代表される漫画様式は，今日では最も広く受け入れられた情報の提示手法の一つであり，静止画へインタラクション情報を加える本システムの意図と相性が良い。

まず，明示的にインタラクションの対象を示すよう，端末に備えられた赤外線センサーの位置情報を用いてハイライト処理を行う。センサーはタグを認識する都度，その方向を X-Y 座標で体験キャプチャシステムに記録する。静止画は前述の通り，このタグ認識に対応して切り出されるので，ちょうど切り出した静止画のどこにタグが存在するのか知ることが出来る。よってこの値を用いると，インタラクションの対象が静止画のどの位置に表示されているか大まかに捉えることが出来る。図 6 右はハイライト処理を行った結果である。インタラクションの対象でない左側の人物がモノクロになって潰れている一方，右側の人物はカラーのままハイライトされる。これにより，右側の人物がインタラクションの対象であったことが見てとれる。

さらにインタラクションの種類についても静止画に投影する。本システムで採用されるインタラクションは大別して「見る」と「話す」であるので，区別を簡単に示すため，「話す」のカットの場合には吹き出しを描く。インタラクションの対象が「話す」カットの場合は，吹き出しについてもセンサーの位置情報を用い，顔があると推定される方向とは逆になるように表示する。一方，撮影者が「話す」場合は枠外から吹き出しを表示する。図 7 は二つの吹き出しが描かれた「話し合う」インタラクションの表示例である。

これにより，同じような絵面のカットを並べた場合でも，どのようなインタラクションが体験の中心だったのか，感覚的に把握することが出来る。



図 7 吹き出し

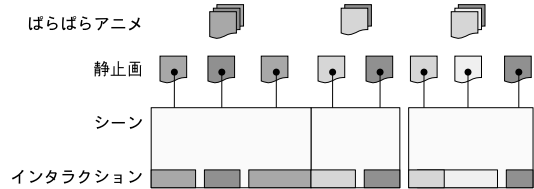


図 8 意味形式と表示形式の対応

吹き出しの中身にはマイクから得た音声を認識し表示することが理想的である。しかし体験キャプチャシステムは様々なイベント空間で用いる前提があるため，得られるであろう種々の自然な会話音声を汎用的に認識することは非常に困難である。そのため，吹き出しの中身には文脈に因らず記号列で代用する。ただし，長い発話であった場合は吹き出しを大きく，中の文字を多くし，短い発話の場合はその逆とする。これにより交互に「話す」「話される」が表示される対話シーンで，どちらが会話の主導権を握っていたか，といった情報を簡単に見て取ることが出来る。

5.4 ばらばらアニメへの変換

こうして得られた静止画，映像データから抜き出した特徴的なカットを，シーンごとにまとめて，ばらばらアニメへと変換する。ばらばらアニメとは個々のカットをゆっくりとクロスフェードさせながら一つずつ表示させていく形式である。これにより映像データの意味的情報を失わず，ある程度の臨場感を保ったまま，時間的には大幅に圧縮することが可能となる。

図 8 に本システムにおける意味的な形式と表示形式の対応をまとめる。一つのばらばらアニメが一つのシーンに対応し，それぞれのシーンを構成するインタラクションごとに，一枚ずつの静止画がばらばらアニメに内包されることが見てとれる。

6 体験データの提示

6.1 概要

これまでの作業により、体験データは自動的に要約され、ぱらぱらアニメという形式へ変換することが出来た。本章からはユーザーが実際に操作する、体験データ観賞・編集アプリケーションへと話を移し、ぱらぱらアニメという素材をどう提示し、どう観賞や編集を促すべきかという点について検討する。

6.2 サイズ

体験データがシーンの数のぱらぱらアニメへ変換されたので、先に演算したシーンの重要度に従って、ぱらぱらアニメを大・中・小の三段階のサイズへと変更する。重要度とサイズの対応は、重要度の点数化同様、求める提示結果によって様々な可能性が存在するだろうが、例えば表2のような値が考えられる。

重要度がユーザーの意図と異なる結果になる可能性は大いに考えられる。しかし重要度を定めずに全てのシーンを画一的に表示するよりも、正しいか不確定であっても様々なサイズを用意して表示する方が、体験データを観賞する際のとっかかりになるはずである。システムが一時的に定めた重要度に対して「なぜこのシーンが大きく表示されているのか」「このシーンはもう少し小さくすべきだ」といった反応をユーザーから引き出すのが、ここでの一番の目的である。

6.3 体験データの例

こうして得られたぱらぱらアニメを時系列順に並べた例が図9である。複数のシーンが重要度に応じたサイズで、漫画のように並べられていることが分かる。

6.4 音声の同時再生

体験データの観賞に臨場感を付与するため、ぱらぱらアニメ再生時にはそのシーンの音声の再生も行う。映像データから静止画を切り出したのと同じように、音声は事前にそれぞれのインタラクションについて数秒間ずつ切り出しておく。音声の切れ目がどこにあるかといった知的な切り出しは全く行っていないため、観賞時にはぱらぱらアニメのコマが変化していく

表2 コマのサイズとシーンの重要度の対応例

サイズ	横 × 縦	シーンの重要度
大	320 × 240	全体の点数の 20%以上
中	208 × 156	全体の点数の 10%以上
小	96 × 72	それ未満



図9 生成された体験データ

のに伴い、ぶつ切りの音声が流れてくることになる。しかし誰と誰が話しているのかといった情報を映像以上に即座に伝えることが可能であり、どのような雰囲気や会話が交わされているのかといった、文字や静止画では伝えにくい部分についても分かりやすく提示することが出来るため、その効果は大きいと思われる。

7 レイアウトの編集

7.1 概要

得られた体験データは、あくまで体験キャプチャシステムのデータから自動的に導かれたものであり、ユーザーの意図は全く反映されていない。そのため、自分にとって重要と思われたシーンが小さく表示されていたり、あるいは一続きであると考えていたシーンが複数に分割されて表示されている場合がある。また通常、シーンは時間順に並べられているが、人によってはこれはインタラクションの対象ごとに並べた方が都合が良いと思うかもしれないし、あるいは面白かつ

たポスター順に並べたり、関係のある人物を近付けるといった、全く異なるレイアウトをとることで意外な発見を得られるかもしれない。

よってこれまでに述べた手法で得た、シーン分けとそれぞれのシーンの重要度、配置は一つの叩き台として捉え、最終的には体験記録システムのユーザー本人が、それぞれの意図をもって体験データのレイアウトを自由にデザイン出来るようにすべきである。

本システムはこうしたニーズに応えるため、体験データの観賞アプリケーションでそのまま、レイアウトを簡単に編集することが出来る。これによりユーザーは、与えられたレイアウトにこだわらず、自由な観点で体験データを見直すことが可能である。

なお、本アプリケーションは Macromedia Flash によって書かれているため、Windows や Mac, Linux など多くのパソコンで標準的に動作し、ほとんど全ての操作がマウスだけで行える。

7.2 シーンの組み換え

シーンをマウスで選択すると縁が赤くなり、図 10 のように縦に三つのアイコンが表示される。上からシーンの「早送り / 巻き戻し」(- のアイコン)「配置変更」(= のアイコン)「表示中の静止画切り出し」(= のアイコン) という機能を示している。

このうち「配置変更」はドラッグによって対象のシーンを好きな場所へ移動させ、レイアウトを変更するもので、「表示中の静止画切り出し」は図 11 のように、表示中の静止画 / インタクションを切り出して新しく一つのシーンとするものである。「配置変更」や「表示中の静止画切り出し」により、配置済のシーンと移動中のシーンが重なった場合、二つのシーンは自動的に結合する。シーン内の静止画 / インタクションは自動で時間順に並べ変えられる。これにより複数のシーンから好きな静止画だけを切り出し、それらを一つのシーンにまとめるといった作業がマウス操作だけで容易に行える。

7.3 その他の機能

レイアウトに関する機能としては他に、個々のシーンのサイズを変更するもの、注釈をつけるもの、特定



図 10 選択されたシーン

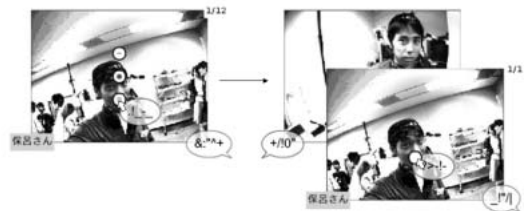


図 11 ドラッグによる切り出し

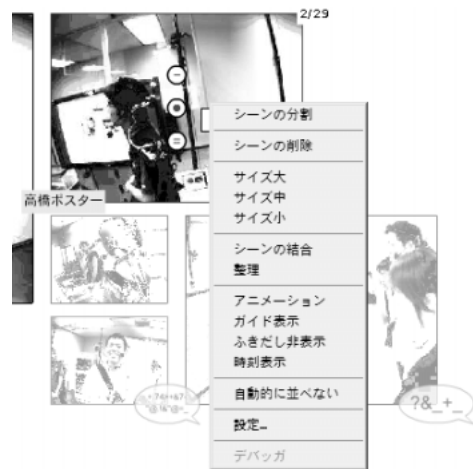


図 12 メニュー

の条件に該当するものだけを表示するよう絞り込みを行うもの、削除するもの、などが実装されている。また、レイアウトを根本的に変更したい時のために、シーンを時間やインタクションの対象に応じて整理、あるいは分割する機能も存在する。これらは図 12 のようなメニューなどから簡単に実行出来る。

7.4 レイアウトの保存

レイアウトの可能性は体験データごとについて無数に考えられ、その中に何か一つ決定的な答えがあるわけではない。そのため本システムではレイアウトを複数、好きな時に保存が可能である。保存されたレイアウトはまた、好きな時に呼び出して再編集することが出来る。同一の体験を様々な観点で見比べるといった利用も考えられる。

一つの体験データに対して受ける印象が、時々によって変化していくことは多い。写真では、撮影時とその後で受ける印象が異なるということが多々ある。本システムについても同様に、同じ体験データに対してレイアウトを複数保存しておくことで、後々にそれらを見直すことを可能にし、結果として新しい気付きを生むきっかけになると考えられる。

8 体験データの共有

得られた体験データとそのレイアウトは、前章で説明した体験データの観賞・編集アプリケーションによりひとまとめのファイルとして生成される。これらのファイルは、同じアプリケーションさえあれば誰でもすぐに観賞・編集を行うことが出来る。すなわち多くの体験データ観賞システムと異なり、本システムは一つのアプリケーションと体験データだけで、他人の体験データを簡単に扱うことが可能である。

体験データを電子メールに送付したり、ウェブページにアップロードすることで、体験の共有も簡単に行うことが出来る。データを得た第三者は、体験者の用意したレイアウトを楽しめる一方で、編集を独自に加えることで全く新しいレイアウトを作成することも可能である。例えば、体験者自身があまり重要視しなかったシーンから面白い会話を発見し、そのカットを大きく配置したレイアウトを作成して体験者本人に返すことで、新たな気付きをフィードバックさせる……というような利用方法が考えられる。インターネットには Flickr や orkut のように、自分が公開したデータが他人によって評価され、意味付けされていくサービスが生まれつつある。本システムはそれらと同様に、データが半ば体験者本人の手から離れて第三者に渡ること、今まで本人には気付かなかった評価や意

義が生まれてくることを目指している。

もっとも、会話内容や視線情報などを含んだ体験データは多くの場合、無制限に配布・共有すべきものではない。本システムによる体験データの取り扱いには、具体的には以下の二つのどちらか、あるいは両方の条件を満たすべきである。

1. 限定的な利用者

一番簡単なのは、体験データの利用者に制限を設けることである。例えばイベント空間で記録を行った場合は、そのイベントに参加した者のみが他人の分も含めて体験データを見ることが出来る、といった具合である。学会、研究会といった場での活用が考えられる。

2. 公的な空間

記録された体験データは自由に配布・共有される、といった前提条件のある空間を設け、その下でのみ記録を行う考え方である。美術館や博物館など、もともと行動に規範を求められるような場、すなわち公的な空間での活用が考えられる。

言い換えれば以上の条件のどちらかを満たせば、体験データ取得に対する障壁はそれほどない。すぐに考えられる本システムの用途としては、学会発表者の体験データをまとめて学会のメンバーサイトに載せる、美術館での体験データをそのまま blog に掲載する、といったものが挙げられるだろう。

9 評価と課題

体験要約という、明確な正解の無い問題に取り組んだため、本システムの評価を行うのは難しい。それでも映像データから静止画の切り出しを行う部分、インタラクション群をシーンとしてクラスタリングする部分については、それぞれ実際にシステムを利用していく中で、ある程度の有用性を確認出来た。例えば、研究発表会であるブースに陣取ったプレゼンターの体験をシーンに分割した場合、次々に訪れるビジターごとのシーンが生まれた。逆にビジターの体験を分割した場合は、訪れたブースごとのシーンとなった。多量のインタラクション・データを用いることで、比較的素朴なアルゴリズムを用いても、意味の通ったシーンを類推することは十分可能であると考えられる。

他方、要約部全体を評価する観点としては、想定されるユーザーの要求にどれだけ従っているものを最初に提示出来るか、また観賞・編集部分においてはどれだけ自由な利用が可能で、インターフェイスがどれだけ優れているか、といったものが考えられる。ただ、どのような要約が正しいのかというのは一義的に答えられない問いである。そこで、先に述べた通り、要約部はユーザーの反応を引き出すためのものと捉え、最終的にはユーザー自身が手を加えることにより、満足のいく体験データを生み出すことが出来ればよいと考えた。実際の利用の中では、ユーザーが得た要約は必ずしも想定のものとは一致していなかったが、逆に「そういえばこんな体験をしたのだった」「意外とこの人とは長く話していたんだった」というような気付きが得られていた。

今後の課題としては大きく二点が考えられる。一つには、対象とするイベント空間の性質によって、要約がどのように変化し、表1のようなパラメータをどのように調整していくか、現在掴んでいる感覚的な理解から、定式的な考えかたへ昇華させること。すなわち、センサーに基づいた要約アルゴリズムとして、本システムに限らない何か基礎的な提案が出来ないかを模索する方向である。もう一つは、現時点では体験キャプチャシステムに依存した格好の本システムを、より広範囲に利用出来ないかを考える方向である。特に「ばらばらアニメ」を用いた編集部分は、膨大な画像データを扱う手法として、もう少し普遍的な利用方法があるはずである。モデルケースとしてデジタルカメラで撮影した写真を、撮影時の背景情報を用いながら「ばらばらアニメ」へと変換し、本システム同様の形にユーザーへ提示するシステムを実験的に作成してみたが、観賞システムとしては十分に利用に堪えるものであった。

その一方、豊富なセンサー情報と大量のビデオ映像に支えられた本システムには独自の特色があることも確かである。一番重要なのは、体験キャプチャシステムのような常時撮影では、本人の記憶していないものが体験として記録され、その中には有意義なものが

潜んでいるかもしれないということである。本稿では「気付き」という言葉を何度か使ったが、実際に自分が撮影したはずのデータから、新しい「気付き」を得た時の面白みは、写真などの他メディアに無いものである。この特性を生かし、ユーザーが抱く意図とは一歩離れることで「気付き」をなるべく提供するシステムを今後も考えたい。

謝辞 本研究は情報通信研究機構の委託研究「超高速知能ネットワーク社会に向けた新しいインタラクティブ・メディアの研究開発」により実施した。

参考文献

- [1] 角康之, 伊藤禎宣, 松口哲也, Sidney Fels, 間瀬健二: 協調的なインタラクションの記録と解釈, 情報処理学会論文誌, Vol. 44, No. 11 (2003), pp. 2628-2637.
- [2] 小関悠, 角康之, 西田豊明, 間瀬健二: ばらばらマトリクス: 漫画技法を用いた映像要約システム, インタラクション 2005, 2005.
- [3] 小関悠, 角康之, 西田豊明, 間瀬健二: シーン推定と漫画技法を用いた体験要約システム, 人工知能学会全国大会, 2005.
- [4] 小関悠, 角康之, 西田豊明, 間瀬健二: ばらばらアニメによる体験データの要約・編集支援システム, WISS2005, 2005.
- [5] 角康之: 映像と音声のコラージュによる体験のウェアネス支援, 日本バーチャルリアリティ学会誌, Vol. 10, No. 2 (2005), pp. 83-87.
- [6] 坂本竜基, 角康之, 中尾恵子, 間瀬健二, 國藤進: コミックダイアリ: 漫画表現を利用した経験や興味の伝達支援, 情報処理学会論文誌, Vol. 43, No. 12 (2002), pp. 3582-3595.
- [7] 高橋昌史, 伊藤禎宣, 土川仁, 角康之, 間瀬健二, 小暮潔: インタラクション解釈における階層構造の検討, 人工知能学会全国大会, 2004.
- [8] 熊谷賢, 中原淳, 角康之, 間瀬健二: 体験要約のためのビデオ自動編集手法, 人工知能学会全国大会, 2004.
- [9] 相澤清晴: ライフログの取得と処理—ウェアラブル, コピキタス, 車一, 人工知能学会全国大会, 2005.
- [10] 河村竜幸, 福原知宏, 武田英明, 河野恭之, 木戸出正継: 実世界対象物とのインタラクションによる体験映像の整理, 日本認知科学会第20回大会, 2003, pp. 240-241.
- [11] 徳村篤志, 野島久雄: 思い出コミュニケーションの活性化のためのカード型写真ビューアーのデザイン, 日本認知科学会第19回大会, 2003, pp. 192-193.
- [12] 徳村篤志, 野島久雄: 思い出コミュニケーションの活性化のためのカード型写真ビューアーのデザイン (2), 日本認知科学会第20回大会, 2003, pp. 256-257.