

# 多視点ビデオデータの時空間コラージュによる追体験空間の構築

Building 3D Virtual Spaces for Re-Experiencing by Spatio-Temporal Collage of Multiple Viewpoint Videos

大高 雄介\*1\*2  
Yusuke OTAKA\*1\*2

角 康之\*1\*2  
Yasuyuki SUMI\*1\*2

岩澤 昭一郎\*2  
Shoichiro IWASAWA\*2

伊藤 禎宣\*2  
Sadanori ITO\*2

間瀬 健二\*2\*3  
Kenji MASE\*2\*3

\*1 京都大学情報学研究科  
Graduate school of informatics, Kyoto University

\*2 ATR メディア情報学研究所  
Media Information Science Laboratories, ATR

\*3 名古屋大学  
Nagoya University

With the recent spread of home videos, video data recorded in our everyday life and social events increase. In order to utilize such video data, we aim to build 3-D virtual spaces, using the video images as textures, where we can re-experience and share our daily experiences and important activities with others. The 3-D spaces will enable us to share personal experiences and knowledge at museums, heritage, educational settings. Most existing works for building 3-D virtual spaces have focused on reproduction of exact 3-D modeled objects and space. Built spaces tend to be homogenized so that it is difficult for virtual visitors of the spaces to find "scent" for re-experiencing. This paper proposes a new method, spatio-temporal collage, to automatically generate 3-D virtual spaces. This method is to provide 2-D perspective images from arbitrary viewpoints by spatio-temporally aligning multiple viewpoints video data.

## 1. はじめに

近年デジタルビデオカメラなどの撮影機器が普及し、日常的に家庭でビデオを撮ることも珍しくない。そして撮影された映像はDVDなどのメディアに記録されている。編集に手間がかかるなどの理由から、現状では映像は蓄積されているだけで、有効利用されていない。この映像を活用する方法が人々のニーズとして存在する。

一方CG技術が成長し、映画、テレビ、ゲームなどの各メディアで3次元仮想空間を利用したコンテンツがつけられている。デジタルシティプロジェクト [3] では現実の都市を電子的に再現しようという試みである。ユーザは実際の場所に行かなくても、仮想的にその都市を歩き回ることが出来る。また STAMP (Spatio-Temporal Association with Multiple Photographs) [5] では複数視点の静止画を用いた擬似3次元空間を提案している。ここでは部分的につながりを持った写真群に対して、表示する画像を連続的に透過処理を施しながら変化させることで、擬似的に時間、空間移動をユーザが体験できる仕組みである。デジタルシティでは現実世界をより正確に再現しようとする。そのためできあがった仮想空間は情報量が多くなり、ユーザは見るものが多すぎて、興味のある対象を自分で探す必要がある。また物体の3次元形状を人間が与えているので、このアプローチは制作に時間がかかり、大量のデータから仮想空間を構築するときこの方法は現実的でない。STAMPでは静止画をベースに構成されているため、当然3次元形状の情報欠落する。しかしカメラで撮影された画像は、撮影者の興味や関心といった情報を含んでおり、時間、空間的にハイライトシーンを切り取る効果がある。写真の間につながりが無い場合でも、擬似3次元空間を構築できるが連続性がなくなるので違和感がある。

そこで本研究では時空間コラージュを使って仮想空間を構築する手法を提案する。時空間コラージュとは複数の視点映像を時空間的に整合させて任意点からの透視映像を作るコラージュ手法のことである。この方法では現実世界が不均一に描画される可能性がある。そのため仮想空間には現れない対象が出てく

る、しかし逆に描画されている対象が強調されることになるので、撮影者の興味を反映していき生きとした映像が生まれる。

この手法ではカメラと対象の位置と向きさえわかれば自動的に仮想空間が構築出来る。これにより作業が自動化され、編集作業などに費やされるユーザの手間を軽減してくれる。応用分野としては、展示会などのイベントなどの参加者間での知識共有や追体験のためのメディア、有形無形の文化財のアーカイブ化、教育支援などに活用できると考える。この手法の実現のためには位置情報と向きの情報が必要である。そのため Local Positioning System を実装した。

## 2. 時空間コラージュとは

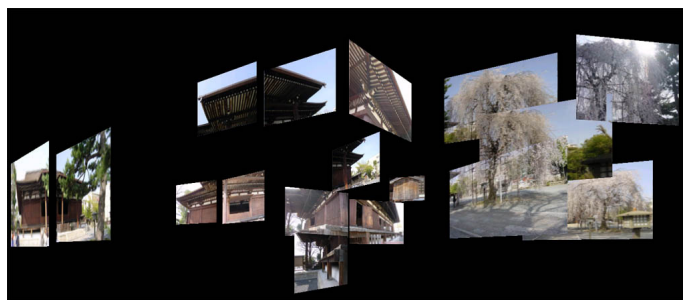


図 1: 時空間コラージュの例

本研究では3次元仮想空間を構築する新たな手法として時空間コラージュを提案する。コラージュとは様々なものを画面に貼りつけて、特殊な効果を出す技法である。これを写真に応用したものがフォトコラージュである。時空間コラージュは、例えば、撮影された時間が異なる複数の画像を同時に表示したり、撮影地点を変えて撮られた画像を繋ぎ合わせたりして元の時間と空間をわざと崩すことによって特殊な効果を狙うものだ。コラージュの例??で挙げたのは複数視点の画像から立体的な映像を作るためのコラージュの一例である。この図では撮影者の位置に画像を配置する、この際画像の向きはカメラの向きに一致させる。仮想空間内では、撮影地点に行くと撮影者の

連絡先: 大高雄介, 京都大学情報学研究科知能情報学専攻, 京都府京都市左京区吉田本町, otaka@lab1.kuis.kyoto-u.ac.jp

視点を追体験できることになる。現実世界を均一に再現した場合、仮想空間内でユーザは何を見ればよいかわからない。これに対して、撮影者が興味を持った対象だけ描かれるのでユーザはハイライトシーンを効率よく見ることが出来る。

時間、空間に対してそれぞれ2種類に分けることによってあわせて4通りの異なる意図を持ったコラージュを検討することが出来る。その4つの分類とは以下のとおりである。

- 時間も空間も現実世界に忠実にコラージュするもの。これは従来のアプローチに近く、現実世界に近づけるようにコラージュする。
- 同一時間内で空間的広がりを持たせる。普通の視点では見えないような部分を空間をずらして表示する。また注目された対象ならば強調して表示する。方法としては、対象の拡大、明るさの増加などがある。
- 同一空間で時間的広がりを持たせる。これはある時刻に描画すべき対象をそれと前後した映像、例えば対象が歩いている場合はその軌跡を描くとか、おなじ空間で起こったイベント、かかわる人間などを描画することが考えられる。
- 最後に時間、空間ともにはばらばらというものでこれはアーティストが担うべき領域であり本研究ではこれに関しては触れない。

本研究ではこれらのコラージュ手法の内、一つ目の手法を実装した。これについては次節で述べる。

### 2.1 画像交差によるコラージュ

これは画像を撮影者の視点に合わせて回転させて、対象を元々あった場所に交差させて配置することによって撮影者から対象までの距離に応じて配置する画像のサイズを変える。これは距離に対して線形にして、遠くの対象ほど画像サイズを大きくする。またウォークスルーのユーザの視線ベクトルと画像の法線ベクトルの角度のずれから画像の透明度を変えて配置した。間の角が90度以下のものは向きが逆になるので描画されない。このとき視線ベクトルの自由度は水平方向360度だけ考えて上下方向は考えない。

### 2.2 対象の軌跡の描画

仮想空間に時間的広がりを持たせる方法として対称の軌跡を描く方法が考えられる。これは描画すべき時刻の映像だけでなくそれと時間的に前後した映像を使うことによって物体の移動、表情の変化を効果的に描写することが出来る。実際の方法としては前後の映像データを保持しておき対象が連続的に変化しているときにそれらをその時々位置にあわせて配置する。そのとき現在時刻に近いほど透明度を低く設定し、離れれば離れるほど薄くして消えていくような表現をする。

## 3. システム概要

本研究では体験キャブチャシステム [4] で用いられたデバイスを使用してシステムを構成している。体験キャブチャシステムは図のようなウェアラブルセットや環境側にもセンサを用いることによってユビキタス環境を実現し、人々の協調的インタラクションを記録する試みである。体験キャブチャルームでは、人と人、人と物、人と環境の間のインタラクションを検出するために、各種センサを利用している。まず、対象物の認識・位置測定を行うために、赤外線 ID タグと、それを認識する赤外

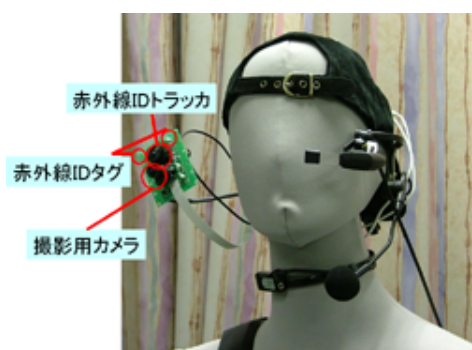


図 2: 装着型センサ



図 3: LPS 用赤外線 ID トラッカ

線 ID タグを利用している。使用機器は体験キャブチャシステム用に作られたもの [6] で以下の通りである。

- 赤外線 ID タグ 図 2 3 赤外線に反応して LED タグから発光される ID 情報を受け取りその ID を持つタグが画面のどの位置にあるか認識するもので、その座標と ID 番号、それらを取得した時刻をセットにしてデータベースに書き込む。このデータはセンサごとに管理される。
- 撮影用カメラ 図 2 Point Gray Research 社製 Dragon Fry を使用した。このカメラは複数のカメラ間で同期を取って撮影することが出来る。
- 赤外線タグ 各タグごとに固有の ID 情報を発光している。このシステムでは LPS 用のタグと物体を認識して仮想空間を構築するためのタグの2種類がある。

## 4. LPS システム

装着型センサと設置型センサで取得されたタグデータと映像はデータベースに書き込まれる。LPS システムで必要となるのは図 3 の赤外線センサによって得られた天井の赤外線 ID タグの位置である。天井のタグは図 4. のように配置されている。タグの世界座標は人の手によって予め実測している。データベースには各センサごとにタグ情報が書き込まれている。まず頭部赤外線センサのタグデータを取得する。取得したデータは歪んでいるのでひずみ除去を行う。ひずみ除去を行ったタグ位置と、対応する3次元座標を位置姿勢計算プログラムに入力する。計算された位置情報がセンサごとにファイルに書き込まれる。位置情報は、時間と座標をセットにして保存される。

LPS システムでは赤外線タグの座標を赤外線センサによって取得し、予め測定しておいたタグの世界座標の幾何学的関係





図 4: LPS システム

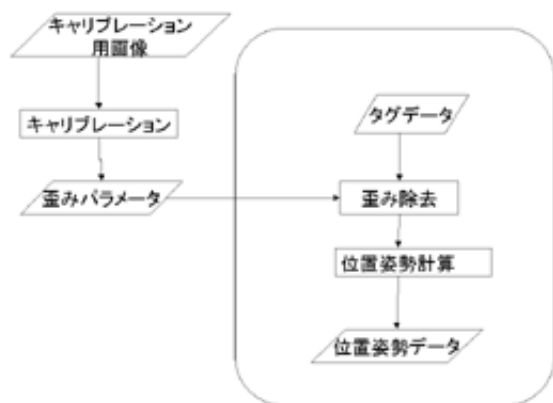


図 5: LPS システムのフローチャート

からセンサの光学中心の位置を計算する。このシステムでは位置計算のために最低 4 点の異なるタグの座標とそれに対応する世界座標を必要とする。このときセンサ側の問題で同一時刻に 4 点を取るといことがありえないためある程度の時間幅を同一時刻とみなして計算している。このため本来得られるはずの幾何学情報以外で計算していることがあり、これが位置推定において大きな問題となる。位置推定のアルゴリズムには Dementhon [2][1] による POSIT アルゴリズムを採用した。

## 5. 追体験空間

5.01 実験環境LPS 用の赤外線タグを 60 個天井に設置して行った。間隔は縦横それぞれ 20cm 間隔で全体で 380cm \* 40cm の長方形領域になる。装着型センサセットをつけた人間が 3 人と赤外線タグだけをつけた人間が一人でインタラクションを撮影した。場所は天井にタグが設置してあるところで撮影された。およそ 10 分ほど撮影し LPS システムによって 3 人の位置と姿勢が計算された。赤外線タグだけをつけた人間は位置と向きを固定として扱った。頭部搭載カメラで撮影された映像を使って 3 次元仮想空間を自動的に構築した。

### 5.1 実験結果

人間 4 人のインタラクションを撮影しその結果をウォークスルー空間で表示させた。できあがったウォークスルー画面は図 6 の通りである。図では 4 人の人間の位置が頭上にあるポリゴンの輪で表現されている。その下に画像が貼り付けてある。この画像は 3 人の頭部搭載カメラで撮影された映像からフレームを切り出したものである。4 人のうち赤い丸で囲まれた



図 6: ウォークスルー画面

人物が注目されているオブジェクトである。2つのカメラからの映像が使われている。それとは対照的にまったく描画されていないオブジェクトが存在する。これはそのときこのオブジェクトを捕らえているカメラが存在しないからである。この対照的な描画はその場の人々の興味を反映しているといえる。

### 5.2 考察

今回は追体験空間を構築する新しい手法として時空間コラージュを提案した。コラージュに必要な位置姿勢情報の取得のために LPS システムを実装した。赤外線センサと赤外線 LED タグを用いて実装された LPS システムは安価で特殊な装置を必要としない。しかしコラージュするためには位置だけではなく向きに関する情報も必要だ。使用した LPS は、向き情報も使うには精度が悪く改善が望まれる。改善のためにはセンサの精度を上げるか、計算に使う 4 点が同一時刻に撮られたことが保証されるような点の選び方をする必要がある。

試作したウォークスルー空間は、画像を交差させて立体的な効果と撮影者の興味を反映させる効果を狙ったが立体的に見えるには位置の調整が難しく、撮影者の向きまで反映させることが出来なかった。注目された対象は複数のカメラの映像を使って表現されるため追体験空間内でも強調されて表示された。これは撮影者の興味を反映しているのでこの手法が有効であることは確かである。しかしウォークスルー空間を魅力あるコンテンツにするためにはコラージュ手法の工夫が必要だ。提案した時空間コラージュはほとんど実装できなかったので今後の課題としてこれらの実装が挙げられる。

このアプリケーションをより一般的するには位置情報を取得するシステムを考えなければならない。屋外ならば GPS、GIS を利用したアプローチが考えられる。将来ビデオカメラに GPS 機能がついて映像データと同時に位置情報が取れるようになると作業の自動化が進むだろう。

## 6. おわりに

本研究では時空間コラージュという新たな仮想空間の構築手法を提案し、そのために必要な LPS システムを実装し、ウォークスルー空間を試作した。映像データをそのまま用いることによって撮影者の興味を反映させた仮想空間を作ることを目指した。

## 7. 謝辞

研究するにあたり、適切な助言とご指導を頂いた角康之助教授、実験環境を提供していただき、数々の有益な助言を下された ATR メディア情報科学研究所の皆様にご感謝致します。なお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ション・メディアの研究開発により実施した。

## 参考文献

- [1] D. DeMenthon D. Oberkampf and L.S. Davis. Iterative pose estimation using coplanar feature points. *CVGIP: Image Understanding*, Vol. 63, No. 3, pp. 495–511, 1995.
- [2] D. DeMenthon and L.S. Davis. Model-based object pose in 25 lines of code. *International Journal of Computer Vision*, Vol. 15, pp. 123–141, 1995.
- [3] Hideyuki Nakanishi, Chikara Yoshida, Toshikazu Nishimura, and Toru Ishida. Freewalk: Supporting casual meetings in a network. In *Proc, CSCW'96*, pp. 308–314, 1996.
- [4] 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, pp. 2628–2637, 2003.
- [5] 田中浩也, 有川正俊, 柴崎亮介. 写真画像群の重なりを用いた広域的な擬似 3 次元空間. 暦元純一 (編), *インタラクティブシステムとソフトウェア IX (WISS 2001)*, pp. 75–84. 日本ソフトウェア科学会, 近代科学社, December 2001.
- [6] 伊藤禎宣, 角康之, 間瀬健二. 赤外線 id センサを用いたインタラクション記録装置, 2003. HI104-4.