

シースルー型デバイスを用いた過去の閲覧

A browsing method using a video see-through device

坂本 竜基*¹
Ryuuki Sakamoto

北原 格*^{1,2}
Itaru Kitahara

キム ハンソン*¹
Hansung Kim

小暮 潔*¹
Kiyoshi Kogure

*¹ATR メディア情報科学研究所
ATR Media Information Science Laboratories

*²筑波大学 大学院システム情報工学研究科
Graduate School of Systems and Information Engineering, Univ. of Tsukuba

The aiming of Groupware systems, which are asynchronous and storage type, is to support sharing character-based information among the users. As external storage gets higher-capacity and less costly, systems archiving the real space using video and audio are proposed. The video of present systems, however, are captured by environmental or wearable cameras, and thus users cannot browse from their desired viewpoint and angle in the captured space. In this paper, we propose the basic technique of asynchronous type groupware to archive meetings held on the real world using free view point video technique with environmental sound. We also propose an interface to browse the captured meeting as free view point video utilizing a video see-through device.

1. はじめに

計算機を用いた組織活動における協調作業の支援環境の構築を目標にした CSCW (Computer-Supported Cooperative Work) やグループウェアに関する研究は 80 年代より盛んにおこなわれている。これらの研究は、オフィスにおける会議や打ち合わせ (以下、ミーティングとする) が支援対象に設定されることが多く、文字ベースの電子掲示板、スケジュール管理、電子メールといった間接的な支援から、映像、音声、アウェアネスといった現実の会議を直接的に扱うものまで幅広く研究されている。グループウェアは、時間的空間的特性により分類することができるが [石井 94]、リアルタイム型と呼ばれる遠隔地にある会議場や、別々の場所に存在する人同士を映像音声で繋ぎ、遠隔地であることを意識させないミーティング環境を提供するタイプのものが数多く提案されてきた。

一方で、蓄積・非同期型は文字ベースの情報共有支援が多かったが、近年では外部記憶の大容量化、低コスト化に伴って、映像と音声によって授業やプレゼンテーションをそのままアーカイブ化することを目標とした提案もなされている [西口 05]。これをミーティングに応用すれば、参加していない過去のミーティングも、その様子や結果の把握が時間・空間に囚われず可能なため、構成員全体の情報格差を抑える効果が期待できる。しかし、既存の蓄積・非同期型のシステムは固定カメラやウェアラブルカメラによって撮影されており、後に閲覧するユーザが希望する視点で自由に閲覧することはできない。例えば、実際の会議ではある人が発言している時、その発言内容と共に他の人の反応も重要な判断材料となる。このような場合、既存のシステムではその人の顔を写す映像に切り替わるため、会議場全体の雰囲気把握することは難しい。

本稿では、ミーティングを音声付きの自由視点映像としてアーカイブすることによって任意の視点から過去のミーティングを閲覧する非文字ベースの蓄積・非同期型グループウェアシステムを提案する。このシステムでは、自由視点映像を作成する際、ミーティングの閲覧という作業に適合するよう、3D モデルの整合性よりも見た目の美しさを優先する手法が用いられる。また、アーカイブされたコンテンツの閲覧にはビデオカメラと小型モニタを組み合わせたビデオシースルー型デバイス

を用いた直感的なユーザインタフェースが提供され、ユーザの視座と自由視点映像上での視点位置が一致した操作を実現される。

2. システム構成

提案システムは、過去のシーンを撮影して 3D モデルを作成するキャプチャ部とその 3D モデルを自由視点映像として閲覧する閲覧部に別れる。キャプチャ部は複数のビデオカメラ、マイクと、各カメラ、マイクの映像音声を保存する計算機、及び 3D モデルを作成するサーバから構成される。ビデオカメラとマイクは、会議がおこなわれる場所を取り囲むように配置し、ビデオカメラについては事前に色合わせと強校正を済ませておく必要がある。

閲覧部は、ユーザが実際に操作するビデオシースルー型デバイスと、このデバイスの位置姿勢を認識するためのタグ、及びタグの認識と認識されたデバイスの位置に沿った映像を作成する計算サーバから構成されるビデオシースルー型デバイスは XGA が表示可能な小型モニターの背面に USB カメラを取り付けた。タグの認識は ARToolKit を利用する (図 1)。



図 1: 位置認識用タグとビデオシースルー型デバイス

3. 処理概要

キャプチャ部において会議の映像と音声を取得し、映像の各フレームにおいて視体積交差法を用いて会議のメンバーの 3D モデルを構築する。視体積交差法は予め取得しておいた背景

連絡先: 坂本 竜基, 国際電気通信基礎技術研究所, 住所: 〒619-0288 京都府相楽郡精華町光台 2-2-2 ATR メディア情報科学研究所, 電話: 0774-95-2553, Fax: 0774-95-1408

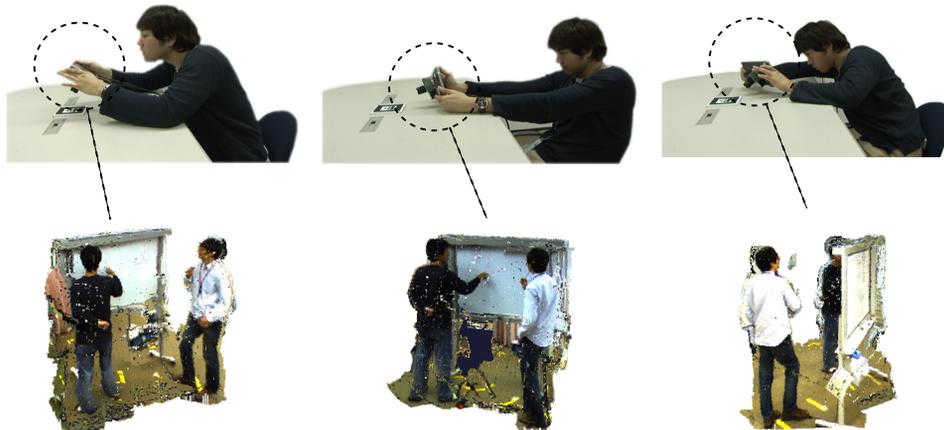


図 2: 上段: 会議を閲覧している様子. 下段: ビデオシースルー型デバイスに写し出されている映像

画像との差分を前景としてボクセルベースの 3D モデルを構築する.

閲覧部は, この 3D モデルを読み込み, 閲覧デバイスの位置に基づいて映像を逐次レンダリングする. レンダリングはボクセルのままおこなうのではなく, 後に述べる Microfacet Billboarding 法をもちいてテクスチャの不連続性を抑制する. 閲覧デバイスの姿勢位置は, 閲覧デバイスが取得する画像からおこなわれるタグの姿勢位置の認識結果をそのまま利用する. また, モデルが作成される源泉となった画像が撮影されたタイムスタンプに基づいて音声の簡易な同期再生をおこなう.

4. テクスチャの連続任意視点映像の作成

複数の映像から 3D モデルを作成する方法としては, 視体積交差法 (Shape-from-silhouette 法 [Laurentini 94]) を用いて 3 次元ボクセルモデルを推定する手法が有名である. この手法では, ボクセルと呼ばれる微小な立方体の集合により量子化した撮影対象空間 (以下, ボクセル空間と呼ぶ) に対し, 撮影物体の形状を推定することから始まる. まず, 撮影した多視点画像中から 3 次元形状を推定したい物体のシルエットを検出し, 画素毎にシルエット領域であるか非シルエット領域であるかを基準に二値化する. 次に, その二値化情報を仮想空間中に投影し, すべての多視点画像からシルエット領域が投影されたボクセルだけが残るようにボクセル空間を削ることにより, 撮影対象の 3 次元形状を凸多面体形状で近似推定する. 最後に, 仮想カメラのパラメータに基づき適切な多視点画像を 3 次元形状モデル上にテクスチャ・マッピングし, 3 次元モデルを生成する. しかし, 上記の方法で作成される自由視点映像は, 仮にモデルの形状誤差が少なかったとしても必ずしも十分な精度で映像を出力できるわけではない. これは, 通常の視体積交差法では, 作成された形状モデルの表面に実写画像をテクスチャとして貼りつけるが, 3 次元空間をボクセルで量子化した誤差と, 形状を凸多面体で近似した誤差の影響により, レンダリング画像中にテクスチャの不連続が発生し, 画質を著しく低下させてしまうことに起因する.

一方で, Microfacet Billboarding 法と呼ばれるアプローチが提案されている [Yamazaki 02]. 微小ピルボードを細かい間隔で 3 次元形状上に配置していくため, 回り込みのような大きな視点変化による見え方の変化を正しく再現することができる上, サイズが小さいため, そのような場合でも個々の見え方の歪みは小さく抑えられる. また, 平面間のマッピングによりテクスチャを貼り付けるので, テクスチャの不連続の問題は回避される. 本研究ではモデル形状の正しさよりも映像の

品質のほうがより優先されるべきである. よって, Microfacet Billboarding 法を用いてレンダリングしたものをを用いる.

5. 実験

提案システムを用いて実際にミーティングを記録し, それをビデオシースルー型デバイスを用いて閲覧する実験をおこなった. まず, 二人の人物がホワイトボードを用いて議論している様子を 7 台の USB カメラを用いてキャプチャリングした. この USB カメラから約 15fps で撮影された映像を自由視点映像化して 3D モデルを作成した. 前景として背景差分された内容は, 人物とホワイトボードであり, ボクセルは 1cm 四方で物体内部のボクセルを削除して作成したところ約 11 万個作成された. これを机の上に置いたタグを用いて再生した様子を図 2 に示す. 閲覧している角度に従って任意の角度の視点が生産されていることが判る.

6. まとめ

本稿では音声付きの自由視点映像によるミーティングのアーカイブ化と, ビデオシースルー型デバイスを用いた閲覧インタフェースによる蓄積・非同期型グループウェアの基盤技術を提案した.

謝辞

本研究は情報通信研究機構の委託研究により実施したものである.

参考文献

- [石井 94] 石井 裕: CSCW とグループウェア, オーム社 (1994)
- [Laurentini 94] Laurentini, A.: The Visual Hull Concept for Silhouette-Based Image Understanding, *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, Vol. 16, No. 2, pp. 150-162 (1994)
- [Yamazaki 02] Yamazaki, S., Sagawa, R., Kawasaki, H., Ikeuchi, K., and Sakauchi, M.: Microfacet Billboarding, in *In Proc. the 13th Eurographics Workshop on Rendering*, pp. 169-179 (2002)
- [西口 05] 西口 敏司, 村上 正行, 亀田 能成, 角所 考, 美濃 導彦: 受講者撮影機能を持つ双方向コミュニケーション記録型講義自動アーカイブシステム, 知能と情報 (日本知能情報ファジィ学会誌), 実践研究論文, Vol. 17, No. 5, pp. 587-598 (2005)