

Scopehandを用いた走査形空間撮像方式

鈴木 元・岡田 敦嗣・河野 隆志 (NTTヒューマンインタフェース研究所)

Abstract: This paper proposes a new video input method for a wide and flat space. A movable video camera and a position and angle sensor is used to detect a position of view point. Captured video images are mapped to a video image memory space. This method creates a useful shared workspace for visual collaborative communication.

Keywords: virtual space, visual communication, shared workspace
video scanning, video mapping

1. まえがき

離れた場所の人間同士が遠隔で協同作業を行う際、映像通信が大きな役割を果たすものと考えられる。映像通信を用いる事により、音声だけでは伝えられない、見せたいもの、見たいものを伝える事ができ、遠隔地の協同作業の効率向上に対して大きな役割を果たすものと期待される。

ここで、映像通信の対象となる物は、従来のテレビ電話が対象としていた話相手の人物の顔画像が主ではなく、協同作業や対話の対象となっている資料や機器、素材、現場状況等の話題材料が主と考えられる。対話者同士が話題対象物をあちこち色々な方向から自由に眺めながら相談できる機能が必要と考えられる。このような話題材料を効果的に撮像、表示するためには、従来のテレビ電話のような視点が

固定された形態のカメラやディスプレイでは不自由であり、見たい物、見せたい物を自在に入力撮像でき、見ている位置を自由に移動できる、視点の自由度を備えたヒューマンインタフェースの実現が重要課題と考えられる。

ビデオカメラ自身の撮像範囲はサイズ、解像度が固定されているが、カメラを人間が持つて動かすことにより、自由に見せたい位置を撮像できる。この際、カメラと一体化した位置センサにより、カメラの撮像位置を検出し、可動カメラを動かして種々の位置から撮像した画像を画像メモリ空間の対応する位置に空間的に張り付けて記憶し、大きな空間全体の状況を表示して伝える方法が考えられる。このようないわば空間全体を可動形ビデオカメラで走査して伝える方法は、大きな解像度の必要な資料や図面

Scanning Video Input Method for a Wide Space
SUZUKI Gen, OKADA Atsushi, KOUNO Takashi
(NTT Human Interface Laboratories)

等をビデオカメラで撮像して説明する場合には非常に効果的と考えられる。

筆者らは上記の自由な視点を持つヒューマンインタフェース実現の観点から、ビデオカメラ、表示器、位置センサ、を一体的に組み込んだ、単眼被覆形ハンドセット (Scopelland) を用いた、対話空間創出方法の検討を進めている [1]。本稿ではScopellandを用い、上述したようなカメラの撮像位置を検出し、それに合わせて画像を空間的に写像して張り付けて、大きなサイズの空間を伝える方式について、コンセプトと原理実験の結果を報告する。

2. 仮想協同作業空間の課題

従来人間同士が協同作業を行う場合は、会議室、作業室、等の協同作業空間に書類、品物、機器等の話題材料を持ち寄って、同席して面談しながら相互のインタラクションが行われてきた。遠隔の人間同士で協同作業を行う場合も、仮想的な協同作業空間を用意し、そこに自分の机上の資料や品物を提示しそれらの話題材料を囲んで対話を進める事が有効と考えられる。

この際、話題材料はいずれかの参加者の近傍の場所に実在し、参加者が自身がビデオカメラで撮像し、他の参加者へ映像通信により提示する形態が想定される。このような話題材料を実際に撮像して協同作業空間へ提示する場合、撮像機能、表示機能として以下の課題がある。

(1) 高精細大空間

文書等の資料や、細かい図面等の提示にはFAX以上の解像度が必要であり、さらに種々の機器を比較したり、並べてみるには大きな作業机に相当するような、高精細で大きな仮想空間が必要である。通常のNTSCテレビ画像の解像度では文書や大空間の撮像・表示には到底足りない。

(2) 見せる視点の自由度

見せたい話題材料を見せたい方向から自在に撮像する機能が必要。

(3) 見る視点の自由度

対象物に近付いて詳しく見たり、離れて全体を見たりできる機能が必要。さらに、共通の作業空間の中で、複数人が各々自由にその中を見回す等、視野

と視点を個人が自由に選択できる必要がある。

(4) 視点の把握

共通の作業空間の中で、相手が何処を見ているのか把握できる機能が必要である。すなわち、相手の視線の方向から何処を見ているか推測できる機能や、あるいは相手が見ているものを直接把握できる機能が必要。

このような機能を実現した仮想協同作業空間のイメージを図1に示す。遠隔に離れた人間同士でひとつの大きな画像空間を共有し、各人はその中で自由に視野を動かせるとともに、特定部分の注視や、空間全体の概観ができる。また共有空間全体の中で相手がどこを見ているかの把握をする事ができる。

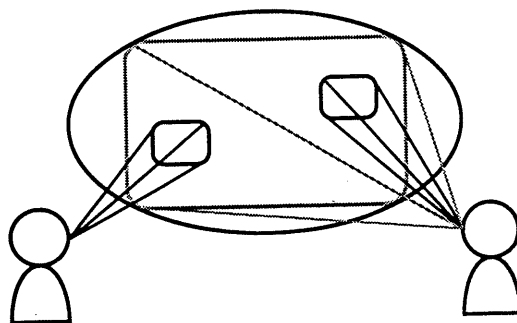


図1 共有空間のモデル

3. 空間走査によるビデオ撮像方式

上記の課題を実現する方法として、本稿では位置センサ付可動カメラを用いた空間走査によるビデオ撮像方式 (略称: VideoScan) について検討を行う。

一般に対象物の形状、位置と、視点条件 (位置、画角) を把握して、ビデオ画像を空間構造モデルへマッピングする方法はテクスチャマッピングの写像手法として、CAD等の分野で使われている。この方法を可動形ビデオカメラで撮影したビデオ画像に対して適用し、撮像対象の構造、撮像条件に合わせて画像メモリ上の仮想空間に写像して、張り付けて記憶・表示させる方法が考えられる。

このような写像手法を用いるためには、以下の点を実現する必要がある。

- (1) 対象物（空間）の形状、位置の把握
- (2) 撮像条件（位置、画角）の把握
- (3) 写像位置の算出

対象物（空間）が複雑な形状をしている場合は、その形状、位置の把握は難しい。このため、対象とする空間を比較的簡単な形状にモデル化する方法が考えられ、空間を半球のドーム形にモデル化して撮像する方式の研究が行われている [2]。

本稿では図2に示すような、対象物の空間を机や壁等のような板状の平面形状にモデル化する方法について検討する。資料や図面、素材等を机の上に置いて眺める場合には、このような平面状の空間モデルが適用できる。平面状のモデルを拡張し、複数の平面で囲まれた空間をモデル化すれば、壁や部屋の形状に対しても適用できる。

撮像条件（位置、画角）の把握に関しては、カメラをロボットアームや機械により駆動する場合はエンコーダ等を用いる事により、比較的精度よく検出できる。

本稿では図2に示すように、ビデオカメラと位置センサが内蔵されたハンドセット (Scopellhand) [1] を手で持って、空間的に走査して撮影する方式を検討した。

撮影画像の写像位置の計算方法に関する、撮影対象物とカメラの位置関係のモデルを図3、図4に示す。撮影対象物が x, y 平面上にあり、 x, y 平面上から z 軸方向に距離 l だけ離れた点 V からカメラで撮像しているとする。レンズの歪を無視すれば、カメラを x, y 平面に垂直にして撮像した場合は撮像画像の中の相対的位置 $s(x, y)$ は、 x, y 平面上の実際の位置 $S(x, y)$ と一致している。 x, y 平面上の点 $S(x, y)$ を撮影すると、画像中の相対的位置 $s(x, y)$ に撮像される。一方、カメラが x, y, z 軸に対して、各々 θ, ϕ, ω だけ回転した状態で撮影した事がわかっている時、得られた画像の中の点 $s(x, y)$ が実際に x, y 平面上のどこにある点か写像計算により求める。

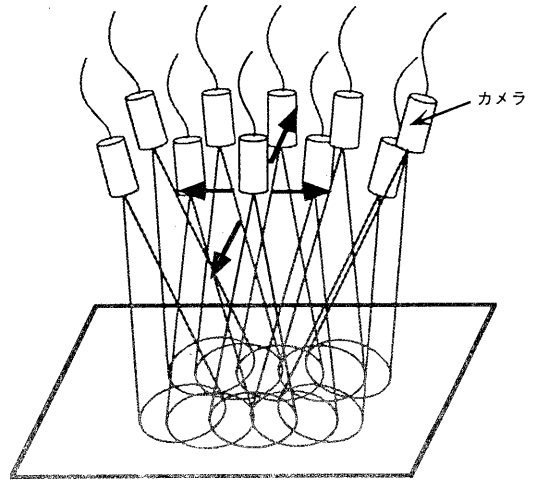


図2 平面上の対象物をカメラで走査するモデル

(1) y 軸方向の回転 θ

実際の位置を $S1(X1, Y1)$ とすると、

$$x = l \tan \alpha \quad (1)$$

$$y = l \tan \beta \quad (2)$$

$$X1 = l \tan (\theta + \alpha) \quad (3)$$

$$Y1 = l \cos \alpha \cdot \tan \beta / \cos (\theta + \alpha) \quad (4)$$

(2) x 軸方向の回転 ϕ

実際の位置を $S2(X2, Y2)$ とすると、

$$x = l \tan \alpha \quad (1)$$

$$y = l \tan \beta \quad (2)$$

$$X2 = l \cos \beta \cdot \tan \alpha / \cos (\phi + \beta) \quad (5)$$

$$Y2 = l \tan (\phi + \beta) \quad (6)$$

(3) z 軸方向の回転 ω

実際の位置を $S3(X3, Y3)$ とすると、

$$r^2 = x^2 + y^2 \quad (7)$$

$$\tan \gamma = y / x \quad (8)$$

$$X3 = r \cos (\omega + \gamma) \quad (9)$$

$$Y3 = r \sin (\omega + \gamma) \quad (10)$$

上記の写像の計算を組み合わせる事により、 x, y 平面上の対象物を任意の方向から撮像した場合の、画像の位置を算出する。

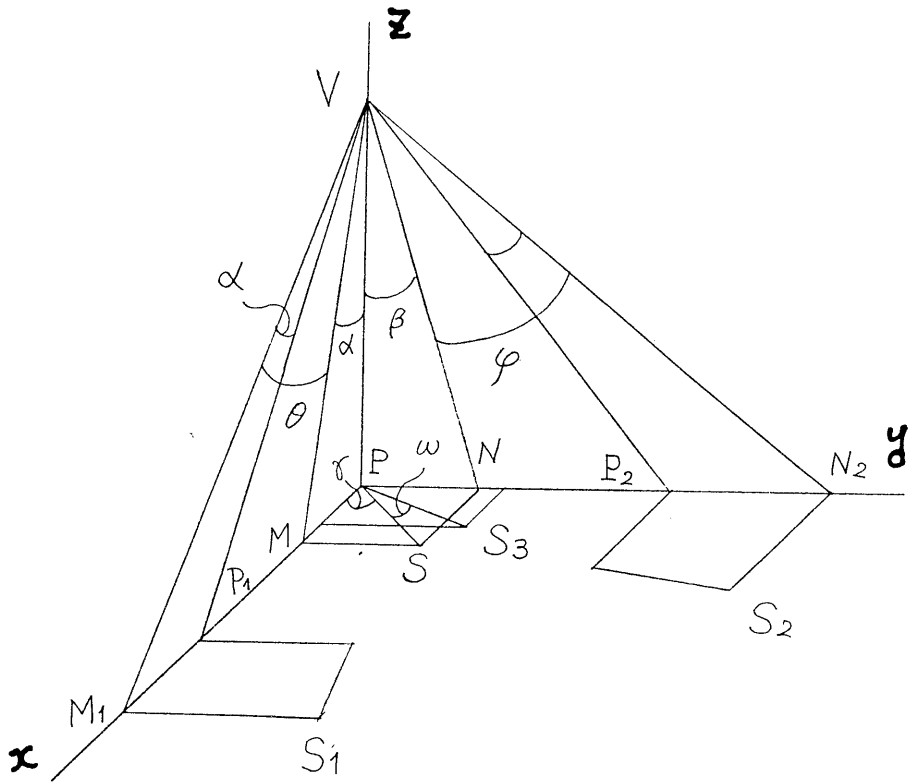
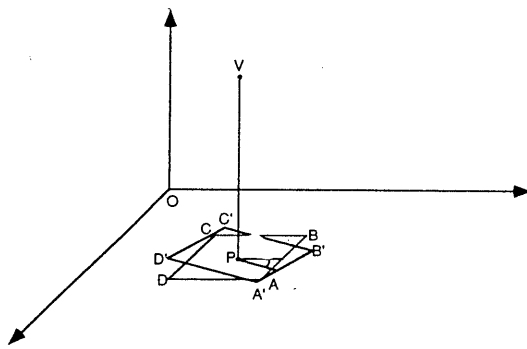
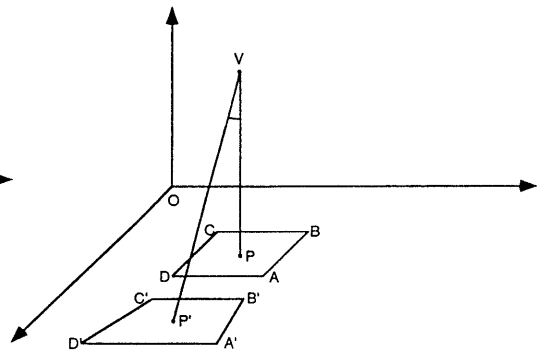


図3 撮影画像の写像位置関係のモデル



z 軸方向の回転



y 軸方向の回転

4. 実験

4.1 実験方法

実験に用いたシステムの構成を図5に示す。また実験条件を表1に示す。

実験の手順を以下に示す。

(1) 机においた撮影対象物(写真)をビデオカメラで撮影する。カメラを手で持ち、机を上下左右に走査しながら撮影する。

(2) 撮影と同時に、カメラと一体にした磁気コイル形的位置センサにより、机面に対するカメラの位置・角度を測定する。

(3) 測定したカメラの位置・角度に基づいて、撮像した画像の位置を算出する。

(4) 机面に相当する画像メモリ空間に、撮像した画像を写像して、連続的に張り付ける。

(5) 画像メモリ空間の全体または部分をビデオ信号に変換して出力し、ディスプレイで見る。顔の向きに応じて表示する位置を切り換える。

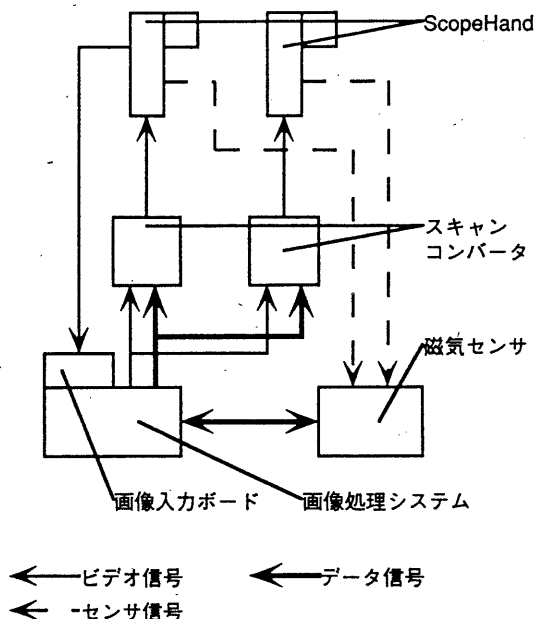


図5 実験システムの構成

ビデオ画像の取り込み、位置の算出、画像メモリ空間への写像はグラフィックWS (IRIS) を用いた。

4.2 実験結果と考察

ビデオ画像を写像して連続的に張り付けた結果の一例を図6、図7に示す。画像の写像と張り付けはビデオ画像一枚当たり約1秒で連続的に処理できた。

図からわかるように、写像した画像の張り付けはかなりずれが目立っている。このようなずれが生じる原因として以下の点が考えられる。

(1) カメラのレンズの歪による誤差

撮影された画像の周辺部ではレンズの歪によるずれが大きくなっている。このようなレンズの歪による誤差を抑えるためには、①レンズの歪を含めた写像位置の算出を行う、②画像の中心部分のみを使用する、等の対策が考えられる。

(2) 位置センサの検出誤差

本実験で位置センサとして用いた磁気コイル形センサは磁場の乱れ等の影響を受け易く、特に周りに金属があると検出誤差が大きくなっている。

(3) 対象物の形状の凸凹

対象物が紙や写真の場合は机と同一平面と見なす

表1 実験条件

画像入力部	カラーCCDカメラ 画素数：41万画素 焦点距離：7.5mm 画角：45° (H)×35° (V)
処理システム部	IRIS 310/VGX 画像入力ボード： Video Lab
位置センサ部	検出方式 磁気センサ 精度 (RMS) 位置 2.5mm 角度 0.5°



図6 実験結果の例（50枚のビデオ画像で走査した場合）



図7 実験結果の例（100枚のビデオ画像で走査した場合）

ことができるが、凸凹のある対象物の場合は平面モデルとのずれを生じる。

本方式を実際に協同作業で使う場合には、画像の張り合わせ部分でのずれがあつて見にくい場合には、撮像した相手に頼んで、その部分を再度撮像してもらう事により詳しく見直すことができる。このため、張り合わせのずれの問題点があつても、空間全体の理解を自由な視点からできる利点の実用上の効用は大きいと考えられる。画像メモリの範囲内で、画面サイズに制約されず、大きな空間を自由に使って撮像でき、仮想空間の効用を活用できる。

5. まとめ

カメラ、ディスプレイ、位置センサを一体化したハンドセット(ScopeHand)を用いて、平面状の話題材料を空間走査的に撮像して入力し、表示する方式について検討を行った。この方式は、画像の張り合わせ位置のずれという問題点はあるものの、自由な視点で空間全体を眺め回す事ができ、協同作業を支援するためのヒューマンインタフェースとしては有効と考えられる。

今後の課題としては、写像による張り合わせのずれを少なくするための補正処理方法の検討、より複雑な形状の対象物に対する撮像方法の検討を進めていく。

参考文献

- [1]鈴木、河野、山森：“ハンドセット形対面对話通信”、信学技報、IE90-91(1991)
- [2]廣瀬：“バーチャル・リアリティの実装”、日本音響学会平成4年度春季研究発表会講演論文集、2-5-7、pp.465-468(1992)