遠隔操作システム「マイクロドーム」の 視覚系の構築

Construction on Visual System of Teleoperating System "Micro Dome"

正 藤本 英雄(名工大) 正 佐野 明人(名工大) ○竹内 博良(名工大) 梶野 英紀(名工大)

Hideo FUJIMOTO, Akihito SANO, Hiroyoshi TAKEUCHI and Hidenori KAJINO

Nagoya Institute of Technology, Nagoya, Aichi

Key Word: Micro Operation, Stereoscopic Dome Display, Visual Registration, Mixed Reality, Kalman Filter

1 はじめに

本研究では、医療応用を念頭に置いた微細作業を支援する遠隔システムの開発を目的とし、人間の自然な作業動作に則した直感的操作が可能な新しい操作環境を提案する。本稿では、背面投射型ドームディスプレイにおいて、頭部の位置・姿勢を考慮した立体映像を提示する視覚系(映像系)について述べる。

2 マイクロドーム

図1にマイクロドームの概略図を示す。マスタコンソールは、半球ドーム状の球面ディスプレイ内に操作環境を立体提示する。提示可能な映像は、CGで構築した仮想世界(シミュレーション・訓練用)および実世界のライブ映像(遠隔操作用)である。操作系として、マスタデバイス(SensAble Technologies, PHANTOM Desktop)をドーム脇に配置した。

本システムでは、2台の液晶プロジェクタ(EPSON ELP-7100)とCFRP製のメカニカルシャッタ、液晶シャッタ式メガネ(Stereo Graphics, CrystalEyes2)を用い時分割方式の立体視を行う。液晶シャッタ式メガネとメカニカルシャッタは、赤外線を使って60Hzで同期をとっている。

Fig.1 Micro Dome System

システム全体の制御、CG環境およびツールの反力生成などには、Windows NTが稼動するPC(Pentium III, 933[MHz])を用いる。一方、立体視用画像モデルの計算を含む実際の画像処理をグラフィックワークステーションO2(MIPS R10000, 195[MHz])を使用する。PC上で生成した仮想世界のCGは、RGB/ビデオ信号変換器(I・O DATA, TVC-XGA)を介して、O2のビデオ入力端子から取り込む。また、頭部の位置・姿勢の計測にはポヒマス社のFastrakを用いる。Fastrakのトランスミッタをドーム脇に固定し、レシーバを頭部(メガネ)に装着する。

本研究では、道具(器具)を使った操作を想定している。操作者が握っているPHANTOMのグリップの先に仮想のツール先端部を取り付けた形になっている。力覚提示に関しては、ライブラリGHOSTを用いる。通常、GHOSTに用意された関数に基づく接触判定ならびに力覚の提示ポイントは、PHANTOMのジンバル部分である。そこで、本研究では、ツール先端での接触判定ならびに力覚提示に変更した。

3 ルックアラウンドとコラボレーション

球面ディスプレイは連続的な表示面を確保できるので、以下のような機能を付加し、そのメリットを最大限活用することを考えている。

通常、人が作業を行う際には、図2(a)に示すように、良好な視覚情報を得るために頭部を動かして、任意の方向より眺めながら作業が行える必要がある(ルックアラウンド機能)。操作者が周囲から眺める際、頭部の動きによって静止しているはずの物体が動いたり、立体視ができなかったりしないように、頭部の位置・姿勢を計測し、それに応じて映像系を制御する必要がある。そこで、本研究では、頭部運動を考慮した立体映像提示を行う。

Fig.2 Advantage of Spherical Display

また、共同作業(コラボレーション機能)のスタイルをとる場合は、図2(b)に示すように、サイズアップしたマイクロドームに複数のウィンドウを開くことが考えられる。

4 頭部運動を考慮した立体映像提示

本研究では、図3に示すような、ドーム系、トランスミッタ系およびフェイス系の3つのブロックから成る座標系を設定した。トランスミッタ系には、ドーム脇にトランスミッタ座標として $\Sigma_{transmitter}$ を設定する。フェイス系には、 Σ_{eye-r} (右眼)と Σ_{eye-l} (左眼)を設定し、その中央にフェイス座標として Σ_{face} を設定する。ここで、フェイス座標原点にレシーバを固定する。なお、視線に関しては、ドーム中心に向かうと仮定する。ただし、視線をドーム中心よりずらした場合でも、ウィンドウの範囲内であれば、ほぼ違和感のない立体視ができると考えている。

Fig.3 Coordinate Frames on Dome Screen and Face

各座標系の回転角には、ロール ϕ 、ピッチ θ 、ヨー ψ を用いる。眼球座標はフェイス座標系に固定されており、回転しないものとする。また、視線がドーム中心に向かうとすると一意に決定でき、眼球座標で表現したベクトルとして取り扱う。

頭部運動に応じた立体映像提示は、左右の眼に対応した提示領域を求め、画像を投影することで実現できる。したがって、各座標間の関係を明確化することにより、眼球座標をドーム座標で記述し、提示領域を導出する必要がある。そこで、ドーム座標系 Σ_{dome} からフェイス座標系 Σ_{face} への座標変換行列は、次のように表わされる。

$${}^{dome}T_{face} = {}^{dome}T_{transmitter} \cdot {}^{transmitter}T_{face} \quad (1)$$

ここで、

$${}^{dome}T_{transmitter} = \text{Trans}(-l_{mtx} - r, 0, l_{mtz} + r) \quad (2)$$

$$\begin{aligned} {}^{transmitter}T_{face} &= \text{Trans}({}^{t-m}x_e, 0, 0) \text{Trans}(0, {}^{t-m}y_e, 0) \\ &\quad \text{Trans}(0, 0, {}^{t-m}z_e) \text{Rot}(z, {}^{t-m}\phi_e) \\ &\quad \text{Rot}(y, {}^{t-m}\theta_e) \text{Rot}(x, {}^{t-m}\psi_e) \end{aligned} \quad (3)$$

$$\begin{aligned} {}^{t-m}E_f &= [{}^{t-m}X_f \quad {}^{t-m}\eta_f]^T \\ &= [{}^{t-m}x_f \quad {}^{t-m}y_f \quad {}^{t-m}z_f \\ &\quad {}^{t-m}\phi_f \quad {}^{t-m}\theta_f \quad {}^{t-m}\psi_f]^T \end{aligned} \quad (4)$$

ただし、 ${}^{t-m}E_f$ はカメラ座標において、フェイス座標 Σ_{face} の位置・姿勢を、3次元位置と疑似ベクトルを用いて表わしたものである。

次に、 Σ_{face} で表わされた右眼の位置 ${}^fE_r = [l_{hx}, 0, 0]$ を、式(1)を用いてドーム座標系 Σ_{dome} に変換すると、

$${}^{dome}E_r = {}^{dome}T_{face} {}^fE_r = [{}^d x_{er} \quad {}^d y_{er} \quad {}^d z_{er}] \quad (5)$$

を得る。さらに、 Σ_{dome} では極座標を採用しているので、 ${}^{dome}E_r = [{}^d \alpha_{er} \quad {}^d \beta_{er} \quad {}^d l_{er} + r]$ を用いて表わすと、

$${}^d l_{er} + r = \sqrt{({}^d x_{er})^2 + ({}^d y_{er})^2 + ({}^d z_{er})^2} \quad (6)$$

$${}^d \alpha_{er} = \sin^{-1} \left(\frac{{}^d y_{er}}{{}^d l_{er} + r} \right) \quad (7)$$

$${}^d \beta_{er} = \sin^{-1} \left(\frac{{}^d x_{er}}{({}^d l_{er} + r) \cos {}^d \alpha_{er}} \right) \quad (8)$$

となる。また、左眼に関しても同様の変換を行う。

以上のようにして得られた位置で、ドーム上での映像生成式を用いることによって、両眼それぞれに対応する提示領域を生成することができ、頭部運動を考慮した立体映像が提示できる。ここで、映像生成式の導出は、ドームに接する仮想スクリーンを用いて、ドームを介して行われる(文献(1)参照)。

5 カルマンフィルタによる運動予測

提案する操作系では、一つのツールを実マスタと仮想スレーブから構成する形で複合現実感(MR: Mixed Reality)を生成する。しかし、操作入力 1kHz で更新されるのに対し、映像の更新は 30Hz で行われるので、映像提示に遅れが生じる(end-to-end system delay問題)。よってこの遅れが大きいと、素早い動きに対して実と仮想のツールが分離してしまい、動的レジストレーションエラー⁽²⁾が起こる。そこで、この遅れによる影響を補償するために、カルマンフィルタによる運動予測を考える。

カルマンフィルタは、観測毎にシステムの現在の状態の推定値と次の状態の予測値を計算するアルゴリズムである。離散時間系のマスタツール先端部の状態方程式は、以下のようになる。

$$\mathbf{x}_{k+1} = \mathbf{F}\mathbf{x}_k + \mathbf{D}_k(\mathbf{x}_k)\mathbf{u}_k \quad (9)$$

$$\mathbf{y}_k = \mathbf{H}\mathbf{x}_k \quad (10)$$

ただし、状態変数 $\mathbf{x}_k$ は 12×1 のベクトルでツール先端部の重心の位置、回転角、速度、角速度から成る。また、 $\mathbf{y}_k$ はツールの位置情報、 $\mathbf{u}_k$ はツール先端部にかかる力とモーメントである。

上記のシステムは、姿勢成分に非線形性を有しているため、そのときの状態変数周りで線形化し、以下の拡張カルマンフィルタを得る。

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k[\hat{\mathbf{y}}_k - \mathbf{H}(\hat{\mathbf{x}}_{k|k-1})] \quad (11)$$

$$\hat{\mathbf{x}}_{k+1|k} = \mathbf{F}\hat{\mathbf{x}}_{k|k} + \hat{\mathbf{D}}_k\hat{\mathbf{u}}_k \quad (12)$$

$$\mathbf{K}_k = \hat{\Sigma}_{k|k-1}\mathbf{H}^T(\mathbf{H}\hat{\Sigma}_{k|k-1}\mathbf{H}^T + \Sigma_{vk})^{-1} \quad (13)$$

$$\hat{\Sigma}_{k|k} = \hat{\Sigma}_{k|k-1} - \mathbf{K}_k\mathbf{H}\hat{\Sigma}_{k|k-1} \quad (14)$$

$$\hat{\Sigma}_{k+1|k} = \mathbf{F}\hat{\Sigma}_{k|k}\mathbf{F}^T + \hat{\mathbf{G}}_k\Sigma_{wk}\hat{\mathbf{G}}_k^T \quad (15)$$

上式によって得られた未来の状態量 $\mathbf{x}_{k+1}$ の推定量を用いて、ツール仮想部を描画する。

6 まとめ

遠隔微細作業支援システムとしてマイクロドームを開発した。また、頭部の動きに基づいた立体映像の再現を行った。そしてend-to-end system delay問題に対しては、カルマンフィルタによる補償を提案した。今後は、本視覚システムの有効性を実験的に評価していく予定である。

参考文献

- [1] 佐野明人, 藤本英雄, 梶野英紀 “マイクロドームにおけるウィンドウのビジュアルフィードバック制御”, 日本機械学会東海学生会第30回学生員卒業研究発表会予稿集, pp.146-147, 1999.
- [2] R.T. Azuma: “A Survey of Augmented Reality,” Presence, Vol.6, No.4, pp.355-385, 1997.