

RGB-D カメラと高速カメラの組み合わせによる 手の動きの3次元形状計測

○郭 雨キン, 諸岡 健一, 辻 徳生, 能登 裕子, 原田 博子, 倉爪 亮 (九州大学)

A Method for Measuring 3D Hand Motion by Combining RGB-D Cameras and Multi High Speed Cameras

○ Yuxin Guo, Ken'ichi MOROOKA, Tokuo TSUJI, Yuko NOTO
Hiroko HARADA, and Ryo KURAZUME (Kyushu University)

Abstract: This paper presents a new method for measuring 3D hand motions by using multiple RGB-D cameras and multiple high speed cameras. The 3D hand shapes are recovered from the depth images obtained from RGB-D cameras. Moreover, the high-speed cameras can capture the hand motion which the RGB-D cameras can not provide. Therefore, we reconstruct subtle hand motions by the high speed cameras.

1. 緒言

近年, 日本 および中国では少子・高齢化社会となりつつあり, 高齢者を介護・支援する看護師へのニーズが高まっている。一方, 医療現場では, 看護師不足問題が深刻になっていると指摘されている。超高齢社会の到来を踏まえると, 熟練看護師と同等の技術を有する若い看護師を効率的に育てることは重要である。

ベットから車いすへの患者の移乗, 包帯交換や皮下注射など, 看護学生は様々な看護動作を習得する。現在の看護教育では, 指導する熟練看護師が手本を示した後に, 看護学生がその動作を模倣しながら手技を修得する。しかし, 熟練看護師が動作を伝達する手段は, 口頭によるものであり, また, 伝える情報は, 看護師の経験や知識など, その多くが定性的である。特に, 医療道具を使用する場合, 指先や腕の動きを口頭で表現することは極めて難しい。また, 定量的な評価ができないため, 常に同じ挙動ができるとは限らない。このような理由から, 看護学生は繰り返し訓練しながら看護動作を習得する必要がある, 看護教育に多くの時間を要す。

この問題を解決するために, 我々は, 看護師の3次元動作を計測し, そのデータを使って手技の熟練度を定量的に解析するシステムを開発している。これにより, 看護手技を定量的に表すことができ, 看護行為の手技伝承を容易かつ安定的に行うことが期待される。また, 熟練看護師と看護学生の動作の比較も可能となり, より効率的かつ効果的な教育が可能となる。本発表では, 看護動作として皮下注射を対象とし, 3次元形状とカラー情報が同時に取得可能なRGB-Dカメラと, 高速カメラを組み合わせ, 手の動きを3次元計測するシステムを提案する。

1.1 関連研究

皮下注射では, 看護師は, 目標となる血管や組織へ正確に注射針を穿刺するために, 繊細な手の動きを要求される。近年, 手の動きを計測するシステムが開発されており, これらはマーカ使用の有無によって大別できる。マーカを使用するモーションキャプチャは, 手や道具にマーカを直接付着する, あるいは光学・磁気式マーカが付いた特殊な手袋を用いることで, 手の3次元を精度良く計測することができる。しかし, マーカや手袋の使用によって, その重みなどにより, 実際とは異なる不自然な動きになる可能性があり, マーカを用いないマーカレス計測が望ましい。また, 体や腕だけ

でなく, 指や注射器などの細かな動きも計測すべきである。

これらの要求を満たす計測法として, 1台のRGB-Dカメラを用いるものがある^{1,2)}。RGB-Dカメラは, 対象物の表面形状とカラー情報を同時に取得可能であり, 体全体や腕の動きの3次元データを得ることができる。一方, 市販のRGB-Dカメラは, 対象物から40[cm]以上離れて撮影しなければならない。また, フレームレート(30fps)や空間解像度は高くなく, したがって, 手の細かな動きを対象とした場合, その計測精度は粗い。別の方法として, 高速カメラなどの単眼カメラのみを用い, 取得したカラー画像と手の3次元モデルを組み合わせることで, 対象手の形状を計測方法がある³⁾。この方法では, 3次元形状を得るために, 事前に様々な形の手のモデルを多数準備する必要がある。

本システムでは, 皮下注射動作における細かな手の動きを計測するために, RGB-Dカメラと高速カメラを同時に使用する。RGB-Dカメラと同一フレームレートのCCDカメラを組み合わせる手法⁴⁾と比較すると, 高速カメラは, ズームによって撮影範囲を変えながら, RGB-Dカメラより高解像度のカラー画像を, 1秒間に1,000枚取得可能である。これにより, 本システムは細かな手の動きを計測可能である。

2. システム概要

本システムでは, 複数台の高速カメラ(カシオ社製・EX100PRO)と, RGB-Dカメラ(Asus社製・Xtion pro live)を用いる。ここで, RGB-Dカメラと高速カメラ, および高速カメラ間は, キャリブレーションにより同期して撮影可能である。Fig. 1は, 時刻 t_1 および t_2 で, RGB-Dカメラと高速カメラの両方で手を計測した場合の, 本システムによる手の3次元形状復元の流れを示す。 t_1 と t_2 で, RGB-Dカメラから, 距離画像とカラー画像を取得する(Fig. 1: Step1)。同時に, N 台の高速カメラによってカラー画像が取得される。

時刻 t_1 と t_2 では, RGB-Dカメラより得られる距離画像から, 手の3次元形状を取得可能である。一方, 時刻 t ($t_1 < t < t_2$)では, 高速カメラのみ手の動きを計測可能である。したがって, 時刻 t では, 高速カメラのカラー画像列を使って手の形状を復元する(Fig. 1: Step2,3)。

一般に, ステレオ視による対象物の形状復元では, ステレオ視を構成する複数枚の画像間の対応関係を求めることが

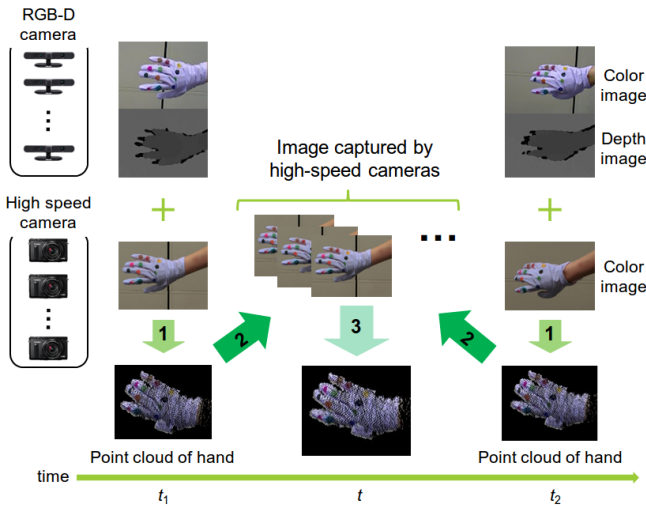


Fig.1 Overview of our proposed measurement system.



Fig.2 Medical gloves with colored points.

重要である。しかし、カラー画像間の対応点探索問題は、オクルージョンや画像に含まれる雑音や歪みなど様々な要因のため、複雑な問題である。この対応点問題を簡単化するために、Fig. 2 に示すように、医療用手袋に点マーカを記入し、この手袋を装着した手を計測する。通常、医療用手袋を着用して皮下注射を行っている。また、1.1 節で挙げた光学・磁気式マーカと異なり、単に手袋に点マーカを記述しているため、この手袋を使用しても、通常の医療行為と同じ条件で看護動作を計測できる。本研究では、点マーカは指の関節 (Fig. 3) の位置にあると仮定し、次節で述べる方法により手の形状を復元する。

3. 高速カメラによる手の形状復元

簡単化のため、1 台の RGB-D カメラと、2 台の高速カメラを使って手を計測し、その計測データから手形状を復元する方法を述べる。ここで、時刻 t_1, t_2 において、RGB-D カメラから、距離画像 S_{t1}, S_{t2} およびカラー画像 $I_{t1}^{(r)}, I_{t2}^{(r)}$ を取得する (Fig. 1)。また、時刻 t ($t_1 < t < t_2$) では、高速カメラのみ手の撮影が可能である。時刻 t で 2 台の高速カメラから取得したカラー画像を、それぞれ $I_t^{(h1)}, I_t^{(h2)}$ と表す。また、1) 3 つのカラー画像 $I_t^{(r)}, I_t^{(h1)}, I_t^{(h2)}$ 、および 2) RGB-D カメラの距離画像 S とカラー画像 $I_t^{(r)}$ の各幾何学的関係は、キャリブレーションによって既知であるとする。

3.1 前処理

時刻 t_1 で取得した、3 つのカラー画像 $I_{t1}^{(r)}, I_{t1}^{(h1)}, I_{t1}^{(h2)}$ から、画像処理により手袋の点マーカを抽出する。また、

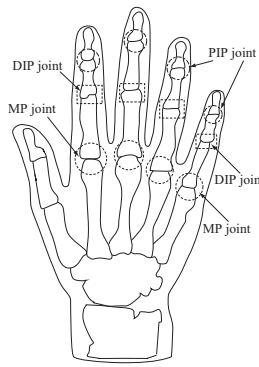


Fig.3 Hand skeleton.

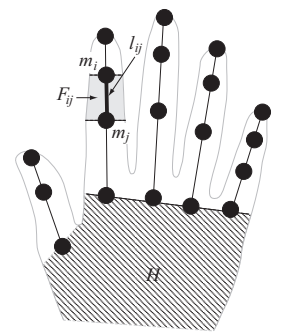


Fig.4 Point markers and links of hand.

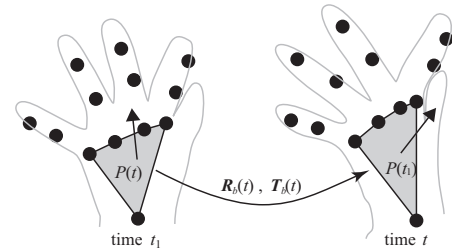


Fig.5 Shape deformation of hand back.

S_{t1} と $I_{t1}^{(r)}$ の対応関係より、 t_1 での各点マーカの 3 次元位置 $m(t_1)$ を得る。更に、 $I_{t1}^{(r)}, I_{t1}^{(h1)}, I_{t1}^{(h2)}$ の幾何学的関係から、 $I_{t1}^{(r)}$ を、2 台の高速カメラのカラー画像 $I_{t1}^{(h1)}, I_{t1}^{(h2)}$ にそれぞれ写像することで、 $I_{t1}^{(h1)}$ と $I_{t1}^{(h2)}$ のマーカの対応関係を求める。時刻 t_2 の画像に対し、上述と同様の処理を施し、 t_2 でのマーカの 3 次元位置 $m(t_2)$ 、および高速カメラのカラー画像 $I_{t2}^{(h1)}, I_{t2}^{(h2)}$ でのマーカの対応関係を求める。

次に、医療用手袋のマーカに基づいて、手を複数の領域に分割し、その領域ごとに S_{t1} を構成する点データ群を分ける。具体的には、Fig. 4 に示すように、隣接する 2 つのマーカ m_i, m_j ($i \neq j$) を線分リンク $l_{i,j}$ で結ぶ。全てのマーカをリンクで連結した後、 S_{t1} の点データごとに、その最近傍にあるリンクを求める。これにより、各リンク $l_{i,j}$ は、それを最近傍とする点データの集合 $F_{i,j}$ を有する。一方、この処理により残った点データは、指以外の手領域、すなわち手の甲部とみなし、その点データの集合を H で表す (Fig. 4)。

3.2 手形状モデルの変形による手形状推定

時刻 t ($t_1 < t < t_2$) では、各高速カメラで取得したカラー画像 $I_t^{(h)}$ ($h = h_1, h_2$) を使って、以下の処理を行う。まず、1 時刻前の画像 $I_{t-1}^{(h)}$ を用いて、2 つの画像 $I_t^{(h1)}$ と $I_t^{(h2)}$ のマーカの対応関係を求め、マーカの 3 次元位置を算出する。

次に、各マーカの移動量に合わせて、時刻 t_1 と t_2 で得られた手形状 S_{t1}, S_{t2} を変形し、時刻 t の手形状を推定する。手の甲部の点マーカは同一平面上にあり、手首を起点として甲部が回転するとみなせる。そこで、Fig. 5 が示すように、時刻 t_1 と t において、甲部の点マーカが乗る平面 $P(t_1), P(t)$ をそれぞれ求める。また、 S_{t1} と S_t の手首の点マーカの位置が一致する並進ベクトル $T_b(t)$ を求める。 $T_b(t)$ を適用し S_t を平行移動した後、 $P(t_1)$ と $P(t)$ の法線ベクトルが一致する回転行列 $R_b(t)$ を求める。最終的に、手の甲部に

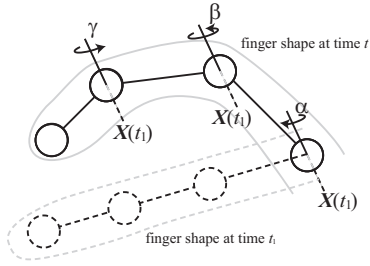


Fig.6 Shape deformation of fingers.



Fig.7 Measurement system composed of one RGB-D camera and two high speed cameras.

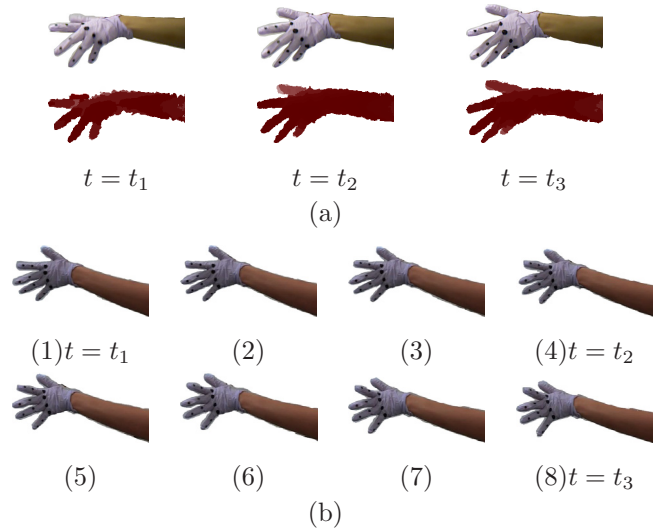


Fig.8 Exmaples of (a) the depth images captured by RGB-D camera and (b) the color images obtained from high speed camera.

ある点データ $p \in H$ を

$$\tilde{p} = R_b(t)p + T_b(t) \quad (1)$$

と移動させることで、時刻 t での甲部の形状データを得る。

指の場合は、Fig. 3 が示すように、指の根元にあり、基節骨と中節骨を接続する中手指節間関節 (以後、MP 関節) を基準に、各点マーカの回転角度を求める。具体的には、Fig. 6 が示すように、時刻 t_1 での、MP 関節上の 4 つの点マーカを結ぶ直線ベクトル $X(t_1)$ を、MP 関節の回転軸とする。先に求めた $R_b(t)$ と $T_b(t)$ を S_{t1} へ適用した後、各指ごとに、MP 関節、近位指節間関節 (PIP 関節)、遠位指節間関節 (DIP 関節) の順で回転角度を求める。この時、各マーカの回転軸は、そのマーカを通り、且つ $X(t_1)$ と平行なベクトルと仮定する (Fig. 6)。まず、時刻 t_1 と t での PIP 関節の位置が一致するように、回転角 $\alpha(t)$ を求める。対象指をなすマーカに、 $\alpha(t)$ による回転移動を適用し、次に、DIP 関節が一致する、PIP 関節での回転角 $\beta(t)$ を算出する。最後に、指先の位置が一致するように、DIP 関節での回転角 $\gamma(t)$ を求める。

以上述べた方法により、 S_{t1} を変形させ、手形状 $\tilde{S}_{t1}$ を求める。同様の方法で、時刻 t_2 での S_{t2} を変形させ、手形状 $\tilde{S}_{t2}$ を得る。最終的に、時刻 t での手形状 S_t は、

$$S_t = w\tilde{S}_{t1} + (1 - w)\tilde{S}_{t2} \quad (2)$$

によって生成される。ここで、重み係数 w は、時間関数として定義し、本研究では、

$$w(t) = \frac{t - t_2}{t_1 - t_2} \quad (3)$$

とした。

4. 実験

提案手法の有効性を検証するために、1 台の RGB-D カメラと、2 台の高速カメラからなる計測システム (Fig. 7) を使って、手の形状を復元する実験を行った。実験で使用

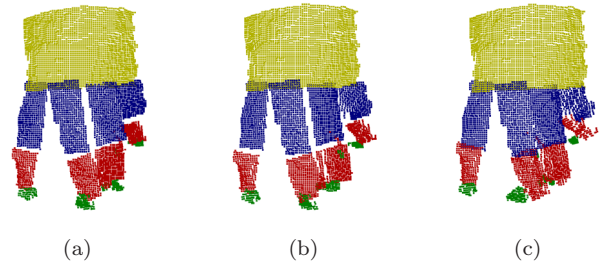


Fig.9 3D hand shapes recovered by our proposed method.

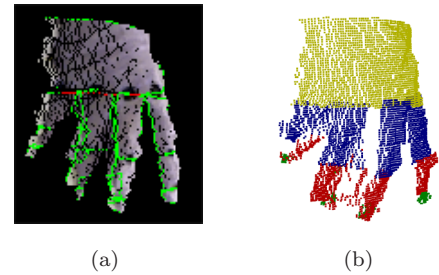


Fig.10 3D hand shapes at time t_2 : (a) S_{t2} measured by RGB-D sensor; (b) $\tilde{S}_{t2}$ obtained from our proposed method.

した高速カメラは、1,200[fps] までフレームレートを上げることができるが、それに対し取得する画像の解像度は低下する。このような解像度とフレームレートのトレードオフの関係を考慮して、今回の実験では、フレームレートを 120[fps] とし、RGB-D カメラと比べ、4 倍の撮影スピードで 640×480 [pixel] の画像を取得した。

RGB-D カメラで取得可能な 3 フレーム分の時間 t_1 , t_2 , t_3 ($t_1 < t_2 < t_3$) において手の形状変化を計測した。Fig. 8 は、RGB-D カメラで取得した 3 フレーム分のカラー・距離画像 (それぞれ、Fig. 8(a) 上・下) と、同時時間帯 t_1 から t_3 で得られた高速カメラのカラー画像例 (Fig. 8(b)) を示す。

このうち、Fig. 8(b-1), (b-4), (b-8) が、それぞれ t_1 , t_2 , t_3 で取得した画像である。全カメラから得られる距離・カラー画像から、手領域を抽出し、その領域内のデータに対し 3. 節で述べた手法で手形状を復元した。提案手法により復元した、時刻 t ($t_1 < t < t_2$) での手モデルを、Fig. 9 に示す。

今回の実験では、画像を取得した後、オフラインで形状を復元した。形状復元に要した時間は、13.4[sec] であり、その多くは画像間の対応点探索に要した。本研究では、オンライン処理は要求されていないが、例えば、対応点探索に並列化を導入することで高速化でき、オンライン形状計測の可能性がある。

提案手法による手形状の計測精度を検証するために、時刻 t_1 と t_3 の高速カメラの画像から、時刻 t_2 の手モデル $\hat{S}_{t_2}$ を推定し、 t_2 で RGB-D カメラから得られる S_{t_2} の形状データと比較した。Fig. 10 は、 S_{t_2} と、提案手法で推定した手形状 $\hat{S}_{t_2}$ を示す。ここで、 $\hat{S}_{t_2}$ と S_{t_2} の差 E を、対応する点マーカー間の平均距離として定量化した。

Fig. 10 に示す手モデルの誤差は、1.4[mm] であったが、同図に見られるように、リンク部において推定誤差が含まれている。この原因として、カメラ間のキャリブレーション誤差や、指の関節の回転軸の算出法の精度、などが挙げられ、これらの問題を解決することが今後の課題である。また、本手法は、高速カメラからの手形状復元では、RGB-D カメラで取得した形状データのみを使っている。そのため、Fig. 9 で見られるように、RGB-D カメラの撮影時では隠れ部にあたり、且つ高速カメラのみの撮影時では観測可能な手表面部は、その形状データが欠損している。今後は、高速カメラのカラー画像を利用した、手形状データの補完を行う必要がある。

5. 結言

本稿では、RGB-D カメラと、高速カメラを組み合わせ、手の動きを 3 次元計測するシステムを提案した。実験結果から、RGB-D カメラの形状データと、高速カメラのステレオ視を融合することで、細かな手の動きを計測することが可能である。皮下注射では、注射器や患者の手の動きを計測する必要がある。この場合、看護学生と患者、注射器の間で隠れが生じやすくなり、データが得られない可能性がある。今後は、隠れに対しても安定して各動作データを取得できる方法を構築する。

参考文献

- [1] I. Oikonomidis, N. Kyriazis and A.A. Argyro, "Efficient Model-based 3D Tracking of Hand Articulations using Kinect," Proceedings of the British Machine Vision Conference (BMVC), pp. 101.1-101.11, 2011.
- [2] C. Qian, X. Sun, Y. Wei, X. Tang, J. Sun, "Realtime and Robust Hand Tracking from Depth," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1106-1113, 2014.
- [3] 星野聖, "単眼高速カメラによる三次元手指形状推定をベースにした新しいインターフェース," 日本機械学会誌, vol.111, no.1080, pp.920-923, 2008.
- [4] S. Sridhar, A. Oulasvirta, C. Theobalt, "Interactive Markerless Articulated Hand Motion Tracking using RGB

and Depth Data," Proceedings of International Conference on Computer Vision (ICCV), pp. 2456-2463, 2013.