

手指リハビリテーションのための RGB-D センサーを用いた手の 3D 姿勢推定

RGB-D Sensor 3-D Measurement of Hand Posture for Rehabilitation

○正 松尾 直志 (立命館大) 正 島田 伸敬 (立命館大)

Tadashi MATSUO, Ritsumeikan University, matsuo@i.ci.ritsume.ac.jp
Nobutaka SHIMADA, Ritsumeikan University

The three dimensional extent where a thumb tip can reach is closely related to functions that can be performed by the thumb. However, on the hand and finger rehabilitation, range of motion for each finger joint is used for evaluating the current state of the hand. It requires a therapist to measure the ranges. In addition, it is difficult to evaluate the three dimensional extent of motion from the ranges. We propose a method to automatically estimate three dimensional position of the thumb tip with a contactless depth sensor. To evaluate the relative position to the palm, we also propose a method to estimate three dimensional configuration of the palm. With experiments, we show the effectiveness of the proposed method.

Key Words: 非接触センシング, 医療診断, 3次元計測

1 はじめに

手指を自在に動かせることは物体の把持や操作などにおいて非常に重要である。これらの機能は腱や神経、骨、関節などの複雑な組み合わせによって実現されている。

手指が損傷を受けると、この複雑な構造を回復させるための高度な外科的治療に加えて、日常生活に必要な機能を回復させるためのリハビリテーションが必要となる。リハビリテーションは手術方法や損傷及び回復の状態に応じて行う必要があり、現在の手指の機能状態を把握することが重要である。手指の機能状態を測る指標としては筋力、知覚、可動範囲などがある [1]。この内の可動範囲については現状、各関節について屈曲、伸展及び外転、内転の可能な範囲をセラピストが角度を測る器具をあてがうことで計測している [2]。

親指は他の指に對面して広い範囲を動くことができ、この広がり物体の把持・操作に重要であるが、従来の各関節に器具をあてがう方法では可動範囲の 3 次元的位置を測定するのが難しい。また、患者一人一人に対してセラピストによる測定を行う必要があり多数の患者を相手にする現場では大きな負担となっている。

可動範囲の 3 次元的位置を自動計測が行えればセラピストの負担を軽減できる上、リハビリテーションによって機能状態がどのように改善したかを数値的に評価することで、リハビリテーション手法自体の改善につながると期待できる。

指の可動範囲を測定する方法としては、関節角度を計測するセンサを搭載したデータグローブなどの器具を装着する方法がある。しかし健康者がコンピュータ操作等に用いる場合と異なり、手指を損傷して治療中の患者がこのようなデバイスを装着するのは負担が大きい。また機能状態によっては装着するために人手を借りる必要があり自動測定としては望ましくない。デバイスの装着を伴わない非接触型の計測方法としては X 線を用いる方法もあるが、被爆を伴うため望ましくない。放射線を用いない非接触計測としては Leap Motion Controller [3] といったデバイスが市販されており、指先の 3 次元位置が計測できるとされている。しかし、デバイスから見て掌と指先が重なっている場合には正しく指先が検出できない。これはデバイスから見た手の輪郭を手掛りにして指先を検出しているためと考えられる。手指の機能状態を測るためには特に掌と對面して動く親指の可動範囲が重要であるため、掌と指先が重なっていても指先位置を推定できるものが望ましい。

そこで本発表では、親指先付近の形状を学習して深度センサの出力から親指先の 3 次元位置を推定する方法と、掌の基準点 (手

首、小指付け根、人差し指付け根) を推定する方法を提案する。提案法によって求めた親指先の 3 次元位置と掌の基準点から、親指先の掌に対する相対的な位置が求まる。入力深度センサで得られる情報だけなので人手によらず 3 次元位置の可動領域が計測でき、より正確な回復状況の評価へとつながることが期待される。

2 システム

画像上の各ピクセルまでの距離を計測できる深度センサとしては Microsoft 社の Kinect [4] を用いた。図 1 のようにセンサの前に手を差し出して撮影すると各ピクセルにセンサからの物理的な距離情報が格納された深度画像 (図 2) が得られる。図 2 ではセンサからの距離 (深度) と各ピクセルの輝度を対応させ、センサから離れるほど暗くなるよう表示している。ここではセンサから手までの距離を 700 [mm] ~ 800 [mm] としており、この距離範囲には手のみしか入っていないものとする。

手全体の形状は非常に多岐に渡るが、親指先付近の形状はある程度限定されていると考えられる。そこで提案法ではまず親指の爪が入る程度のウィンドウを設定 (図 2) し、その内側での深度分布を手掛りにして親指先位置を推定する。ウィンドウサイズは 32×32 [pixel] としており、これは手までの距離を 700 [mm] ~ 800 [mm] とした場合 4 から 4.5 [cm] 四方に相当するサイズである。親指位置推定は以下のような多段階の推定を用いて行う。

1. 深度画像内のウィンドウごとに、ウィンドウ内深度分布からそのウィンドウが親指先を含むかどうかを判別する (図 3(a))。
2. 親指先を含むと判別されたウィンドウについて、ウィンドウ内の親指先と思われる位置を推定し、親指先位置候補を集める (図 3(b))。
3. 親指先位置候補の内、最も多くのウィンドウから指示された点を推定結果とする (図 3(c))。

求めた親指先位置はセンサを基準とした絶対座標であるが、指の可動範囲を知るためには掌に対する相対的な座標が必要となる。そこで手全体に渡る深度分布から手首、小指付け根、人差し指付け根の 3 つの基準点の 3 次元位置を推定し、掌を基準とした親指先の座標を求める。

3 深度分布特徴ベクトル

ウィンドウ内の深度の分布は、各ピクセルのセンサからの絶対的な深度よりもウィンドウ内でセンサに比較的近い領域がどのように分布しているか、といった相対的な深度情報も重要となる。そこで提案法では、ウィンドウ内の相対的な深度分布をより強く反映するよう、以下のような方法で特徴ベクトルを構成する。

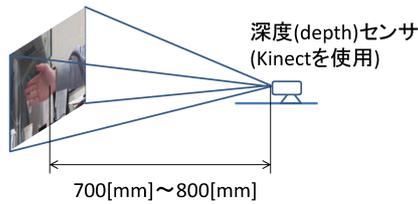


Fig.1 センサとの位置関係

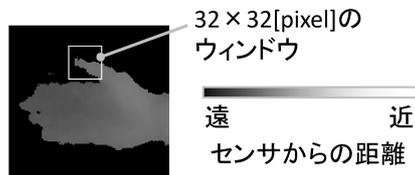


Fig.2 深度画像



Fig.4 教師ラベル

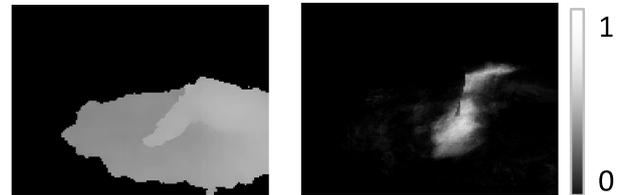


Fig.5 親指先を含む確率

成分。

相対関係成分 32×32[pixel]の深度分布をそれぞれが8×8[pixel]サイズのブロックに分割し、ブロックの全てのペア (${}_{16}C_2 = 120$ 組) についてのブロック内深度平均値の差として得られる120次元の成分。

4 ウィンドウが親指先を含むかどうかの判別

ウィンドウが親指先を含むかどうかの判別には Random Forest Classifier [5] を用いる。Random Forest 内の分類木は100本とし、弱識別器に採用する特徴の評価には Gini impurity [6] を用いる。

学習の際には図4のように親指に青色の指サックをはめた状態で深度画像とカラー画像を同時に撮影したものを用いた。まずカラー画像から色を用いて指サック領域を抽出してその重心位置を求め、深度画像上の各ウィンドウについて、ウィンドウが指サック領域重心位置を含めば親指先を含むものとしてラベル付ける。このようにして得られるラベルとウィンドウ内の深度分布特徴ベクトルとをペアにして教師データとし、Random Forest Classifier を訓練する。

今回は掌を広げて親指以外の指は自然に伸ばした状態で、指サックをはめた親指を動かしながら85枚のシーンを撮影した。各シーンについて手領域付近の様々な位置にウィンドウを設定して4800個の教師データを収集、合計408000個の教師データを用いて訓練を行った。

学習した Random Forest Classifier を用いて親指先を含む確率を求めた結果を図5に示す。入力した姿勢は親指先が掌に接触しており、輪郭での検出ができない状態であっても親指先付近で確率が高くなっていることが分かる。

5 親指先位置の推定

親指先を含むウィンドウについて、親指先位置の相対位置を出力するモデルとしては Random Forest Regression [5] という回帰モデルを用いる。4と同じく Random Forest 内の分類木は100本とし、学習時の特徴選択における評価には平均二乗誤差を用いる。回帰モデルの入力は親指先を含むウィンドウの深度分布特徴ベクトルであり、出力はそのウィンドウに対する親指先の相対位置である。

学習には4で用いたシーンから、親指先を含むウィンドウの深度分布特徴ベクトルを抽出して用いた。各シーンについて手動で親指先位置を指定し、上記の深度分布特徴ベクトルのそれぞれと、その抽出元ウィンドウから見た相対位置を組み合わせる教師データとした。今回の実験では85枚のシーンから、21760個の教師データを作成し回帰モデルを訓練した。

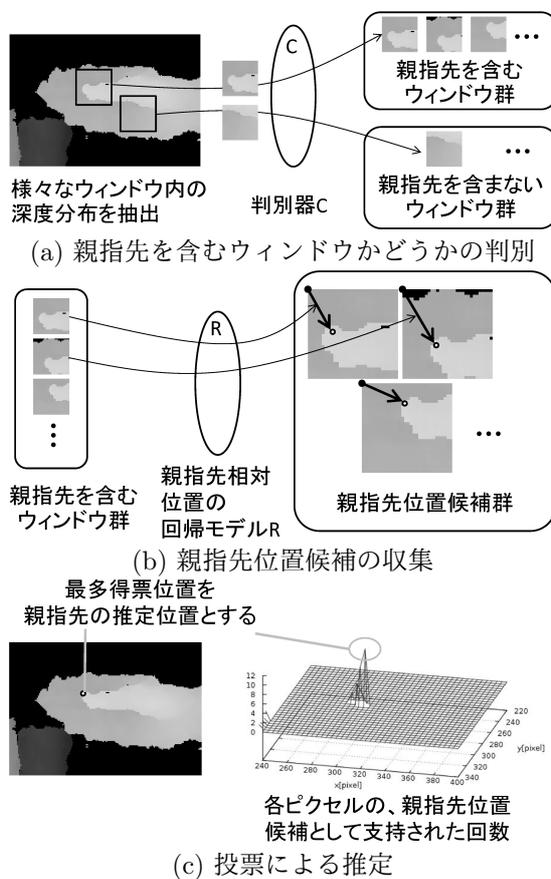


Fig.3 推定の流れ

1. ウィンドウ内深度値を正規化する。センサからの深度基準で手と背景を分離して、手領域内の深度の中央値を求める。ウィンドウ内各点の深度からこの中央値を引き、さらに背景領域には深度値100[mm]を与える。
2. 正規化した深度分布から絶対値成分と相対関係成分を作成し、それらを合わせて $256 + 120 = 376$ 個の成分を持つベクトルを構成する。

低解像度成分 32×32[pixel]の深度分布を4×4[pixel]サイズの定数フィルタで平滑化し、それを2×2[pixel]間隔でサンプリングして得られる $16 \times 16 = 256$ 次元の

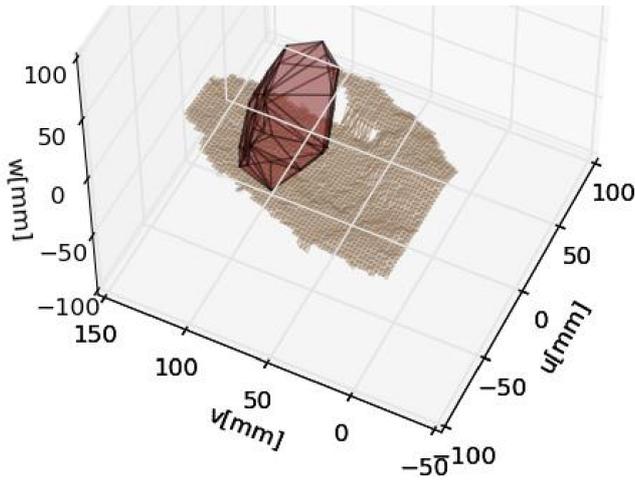


Fig.6 親指先のみを可能な限り動かした場合の指先位置の凸包

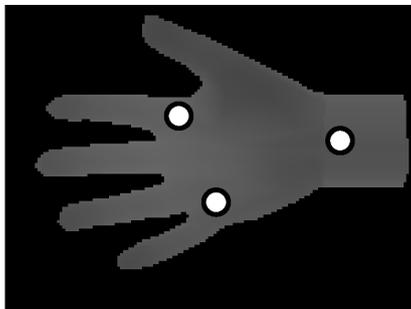


Fig.8 3つの基準点

親指先を含むウィンドウについて、その深度分布特徴ベクトルを回帰モデルに入力すれば親指先の推定位置が求まる。これを親指先候補位置とする。親指先を含むと判別された各ウィンドウについて、そこから求まる親指先候補位置の画素に1票入れる投票を行い、得票数最多の画素に対応する3次元位置を親指先位置とする。

図3(c)は、投票を行い親指先位置を求めた様子を図示している。図6は掌を広げた状態で手はほとんど動かさず、親指先位置のみを可能な限り動かした場合の3次元推定位置から、凸包として可動範囲を図示したものである。実際の指先の動きに対応した形状が得られている。

6 掌基準点の推定

指先の相対的な位置を求めるための基準として手首、小指付け根、人差し指付け根の3つの点を用いる(図8)。これらの基準点は指先のような検出しやすい特徴的な形状ではないため、掌を含めた手全体の深度分布から推定する。深度画像上の手領域の面積は手がセンサに近いほど大きくなる。手の位置に依らず常に手全体の形状を取り出せるよう、基準点の推定には下記のように正規化した深度分布を用いる。

1. センサ平面の中心軸上に手が写っていると仮定し、深度画像上でセンサの正面に相当する小さな領域(センサからの距離が800[mm]の位置で2[cm]四方に相当)について深度の中央値を求め、代表深度とする。
2. センサ平面の中心軸上、代表深度の位置にセンサ平面に平行な200[mm]四方の矩形の代表領域を設定する。深度画像上の、基準領域に相当する領域の深度分布を抽出する。

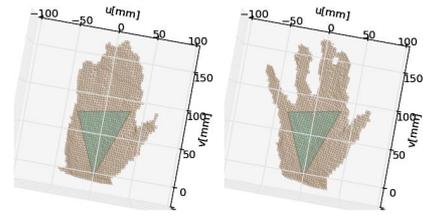


Fig.9 基準点の推定結果

3. 抽出した深度分布を32×32[pixel]に縮小した上、各画素の深度から代表深度を引いて正規化深度分布とする。

この正規化深度分布と、代表領域の中心に対する基準点の相対的な3次元位置との関係を図7のような構造を持つConvolutional Neural Network(CNN)[7]で学習し、正規化深度分布から基準点の位置を推定する。

CNNの学習には教師信号として基準点の座標が必要となるが、実際の手について基準点の3次元位置を求めるのは難しいので3次元CGモデルで作成した手を模したモデルを用いた。3次元CGモデルの位置を±2[cm]の範囲で、手首角度や指の各関節角度を0~5°の範囲でそれぞれ独立に、ランダムに変更して正規化深度分布を1000パターン生成し、それを用いてCNNの学習を行った。

図9に実際に人間の手を入力して基準点の位置を推定した結果を示す。手首、小指付け根、人差し指付け根の3つの基準点を正しく推定できていることが分かる。

図10は、指を伸ばした状態で指の関節角度は変えずに手首の角度のみを変えて撮影を行い、各フレームについて親指先位置を推定してその結果を凸包として表示したものである。図中の手は、動作開始時の位置と姿勢を描画したものである。図10(a)はセンサ基準の座標系を用いて親指先位置の移動範囲を表示しており、手首の角度変化による手全体の移動も含むために移動範囲が広がっている。図10(b)は各フレームについて推定した3つの基準点によって定めた手首を原点とする相対座標系を用いており、親指先が掌に対してあまり動いていないことを反映して移動範囲は狭くなっている。

図11は、親指を、その指先が輪を描くように動かしたときの推定指先位置の凸包を表示したものである。図中の手は、動作開始時の位置と姿勢を描画したものである。図11(a)では手全体の動きによる移動も含めて凸包が作られており、掌に対して実際に動かした範囲よりも細長い形状となっている。図11(b)では相対的な位置を用いているので、手全体の動きが含まれず実際の動きに近い形状となっている。

7 おわりに

深度センサの出力から親指先の3次元位置を推定する方法と、掌の基準点(手首、小指付け根、人差し指付け根)を推定する方法を提案した。推定した基準点により相対座標系を定めれば、掌に対する相対的な位置が求められ、掌に対する指先の相対的な移動範囲が得られることを示した。しかし実際の推定精度は評価できておらず、今後高精度な3次元センサ等を用いて提案法の精度を評価する必要がある。

References

- [1] 独立行政法人:“職業性外傷:手のリハビリテーション”, <http://www.research12.jp/gaisho/08.html>.
- [2] 米本, 石神, 近藤:“関節可動域表示ならびに測定法”, リハビリテーション医学, **32**, 4, pp. 207-217 (1995).
- [3] I. Leap Motion: “Leap motion controller”, <https://www.leapmotion.com/product>.
- [4] Microsoft: “Kinect for windows”, <http://www.microsoft.com/en-us/kinectforwindows/>.

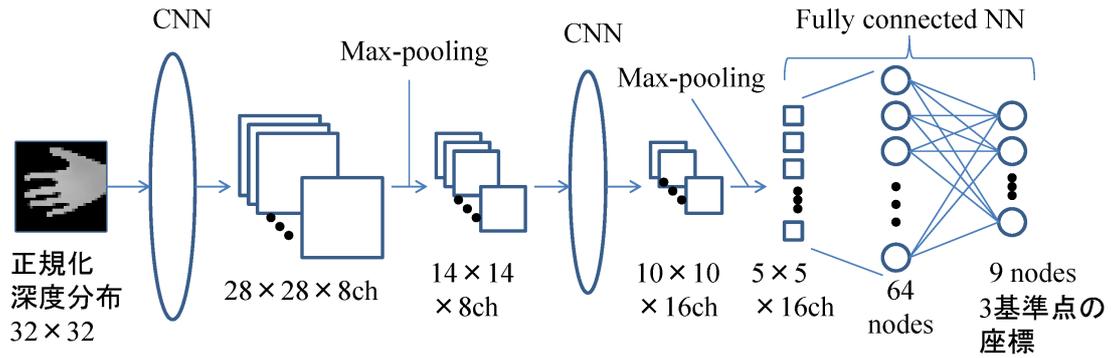
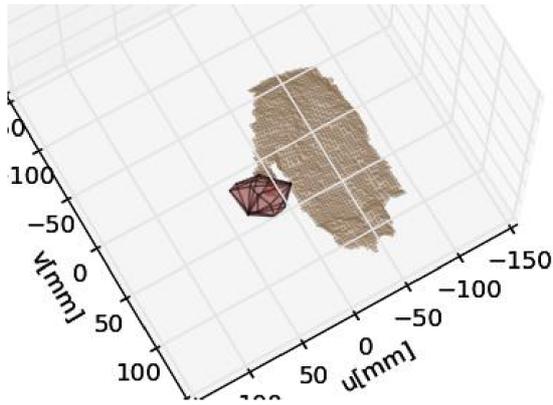
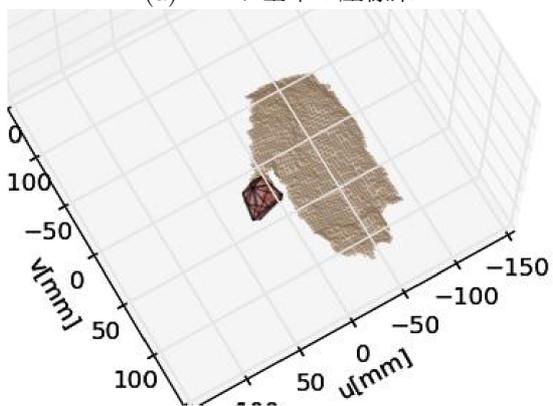


Fig.7 Convolutional Neural Network



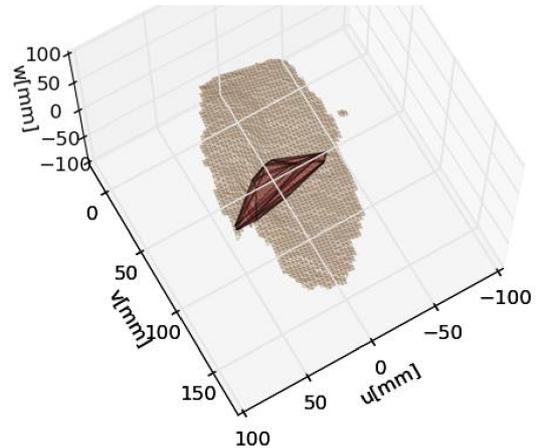
(a) センサ基準の座標系



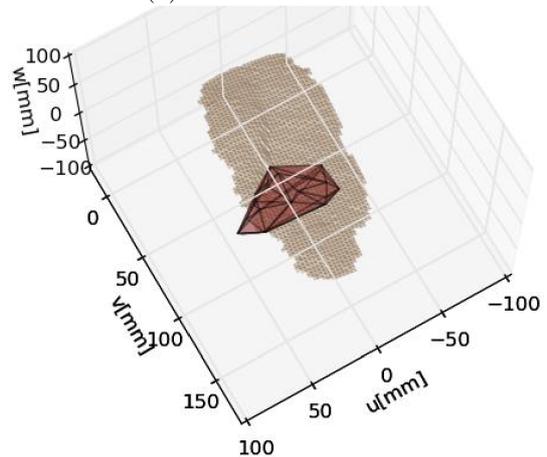
(b) 各フレームの基準点によって定めた相対座標系

Fig.10 手首のみを動かした際の親指先の移動範囲

- [5] L. Breiman: "Random forests", Machine Learning, **45**, 1, pp. 5-32 (2001).
- [6] L. Breiman: "Technical note: Some properties of splitting criteria", Machine Learning, **24**, 1, pp. 41-47 (1996).
- [7] Y. Lecun, L. Bottou, Y. Bengio and P. Haffner: "Gradient-based learning applied to document recognition", Proceedings of the IEEE, **86**, 11, pp. 2278-2324 (1998).



(a) センサ基準の座標系



(b) 各フレームの基準点によって定めた相対座標系

Fig.11 親指先の移動範囲