

掌装着型全天球カメラを用いた手指の姿勢推定手法

丸山裕介[†] 河野恭之[†]

概要: 本研究では、ユーザが掌に全天球カメラを装着することで手指の自然な姿勢を推定するシステムを開発する。全天球カメラには画角の大きい魚眼レンズが使用されているため手指の検知が容易である。まずカメラで取得した画像から手指を検知して指先を検出することでカメラと指先の位置関係を求め、また各関節間の距離、カメラと手の位置関係を設定し、手指が自然な姿勢をとる際に生じやすい各関節間における依存関係を利用することで逆運動学的に手指の自然な姿勢を推定する。手指姿勢を推定するためのデバイスが全天球カメラのみであり、このデバイスをユーザの掌に取り付けるためユーザの運動範囲を制限せずに手指姿勢を推定することができる。またカメラを用いていることから他方の手やマーカなどを検出対象にできることが考えられる。本研究の応用に手指の姿勢を扱う AR や VR が考えられる。ユーザの運動範囲を制限することなく手指姿勢を推定できることから、本システムを応用した AR や VR を体験する際により直感的な操作が可能となり、臨場感を味わうことができると思われる。現在、本研究の応用にろくろ回しを検討している。ユーザの手指の姿勢推定に加えてろくろを出現させるためのマーカの検出と両手との相対位置姿勢の推定を行うことで仮想的なろくろ回しを実現する。

キーワード: 全天球カメラ, 手指の検知, 手指の姿勢推定

1. はじめに

本研究では、ユーザが掌に全天球カメラを装着することで手指を検知し、手指の自然な姿勢を推定するシステムを開発する。全天球カメラとは複数のレンズによって撮影された画像を合成し上下左右全方位の 360 度パノラマ画像を生成する事ができる装置である。全天球カメラには魚眼レンズが使用されており、一般的なカメラに使用されているレンズと比べて画角が大きい。そのため掌に装着するカメラに全天球カメラを使用することで手指を容易に検知することが可能となる。カメラ画像において手指を検知し指先を検出することでカメラレンズの中心からの指先への方向を求めることができ、カメラと手の位置関係や各関節間の距離を設定した上で各関節角の間に生じやすい依存関係を利用することで逆運動学的に手指姿勢を推定できる。近年、AR や VR などのユーザインタフェースが注目されており、それに伴いヒューマンコンピュータインタラクションなどの分野において手指姿勢を 3D 復元する研究が盛んに行われている。例えば、赤外線カメラを用いて手や指の位置姿勢を計測する Leap Motion[1]や、筋電センサや加速度センサが内蔵されたハードウェアを腕に装着して手指の動きを推定する Myo[2]などのデバイスが開発された。柏木ら[3]は把持される筒状の物体の蓋の上面や透明な物体の内部に魚眼レンズカメラを設置して手指を認識する手法を提案したが、この手法においてユーザは物体を把持していなければならない、また把持対象の物体に魚眼レンズカメラが装着されていることが必要であるため、手指姿勢を推定できる条件が厳しい。上田ら[4]は複数台のカメラでユーザの手を多視点から撮影することで手の形状を推定しているが、カメラが複数台必要であり、そのすべてのカメラの画角に収まる範囲でしかユーザは手を動かすことができない。Gentile ら[5]は Kinect を用いて距離画像から手を検知して



図 1. 全天球カメラの装着例

距離画像の特徴を学習することで手の開き具合を推定しているが、各手指の姿勢を個別に推定することができない。本手法ではカメラを掌に装着するのみで手指姿勢を推定することが可能であるためユーザは運動範囲を制限されない。本システムを AR や VR に応用した場合、ユーザの運動範囲が制限されないことから直感的な操作が可能となり臨場感を味わう体験ができると考えられる。掌に装着したカメラで他方の手やマーカを検出することで両手とマーカとの相対位置を推定する手法を検討しており、また AR 技術を利用したろくろ回しの実現を目標としている。本研究では全天球カメラに RICOH THETA S を用いている。ユーザの手指の動きをできるだけ妨げないよう考慮した装着例を図 1 に示す。

2. 指先座標の検出

手指の姿勢を推定するにあたり画像上での指先座標を求める。本章では、掌に装着された全天球カメラの画像から手指領域を抽出し、得られた手指領域において指先となる座標を検出するまでの過程を述べる。本章におけるシステムの流れを以下に示す。

[†] 関西学院大学
Kwansei Gakuin University

- i. **肌色領域の検出**：手指は肌色であるためカメラ画像において肌色の色相範囲を基に肌色領域を検出する。
- ii. **手指領域候補の抽出**：手指はカメラレンズとの物理的な距離が小さいため手指領域は比較的面積の大きい領域となると考えられる。この特性を利用して手指領域候補を絞り込む。
- iii. **手指の検知**：手指が画角内に存在する場合にはまず指先が映ることから、各手指の可動範囲を考慮して設定した手指の探索範囲内で各手指の指先画像のテンプレートを用いたテンプレートマッチングを行うことで手指を検知する。
- iv. **手指領域の決定**：手指領域は比較的面積が大きい肌色領域となることから、手指が検知された場合には各手指の探索範囲内で最大の肌色領域を手指領域に決定する。
- v. **指先座標の決定**：カメラを掌に装着すると手指の付け根はカメラレンズの近くに位置し、カメラ画像において外側に位置すると考えられるため、画像上に手指の付け根の座標を設定し、手指領域の輪郭においてこの座標から最も距離が大きい座標を指先座標に決定する。

まず全天球カメラで得た画像から肌色領域を抽出する。カメラ画像の例を図2に示す。Sherrahら[6]が提案した肌色の色相範囲を基にカメラ画像を2値化することで肌色領域と非肌色領域に分離する(図3)。検出された肌色領域にはカメラが装着されている手の指以外も含まれている可能性があるため、肌色領域から手指とされる領域を抽出する。カメラを装着した手の指はカメラレンズとの距離が小さいためカメラの画角内に手指が存在する場合は他の肌色領域と比べ面積の大きい領域となると考えられる。そこで肌色領域検出画像に距離変換処理を施す。距離変換とは、二値画像においてある画素から最近傍の0画素までの距離をすべての非0画素に対して求め各画素の画素値とする処理である。抽出された肌色領域に対してこの距離変換処理を施し、閾値を設定して再度2値化することで比較的面積の大きい肌色領域を抽出する(図4,5)。またカメラ画像に手指領域が存在する場合、画像に指先が映る。指先の形状は丸みを帯びており、図2から2値化画像においてその特徴が反映されていることがわかる。そこで2値化された各手指の指先画像のテンプレートを用意し、テンプレートマッチングを行う(図6,7)。各手指が内転(外転)するときの可動角から各手指の探索範囲を設定してテンプレートマッチングを行い、閾値以上の類似度が表れた探索範囲には手指領域が存在すると仮定する。内転(外転)とは掌に対して平行な手指の運動を指す。各手指の内転(外転)の可動角を表1に示す。手指が存在するとされた探索範囲において面積が最大の領域を手指領域に決定する。

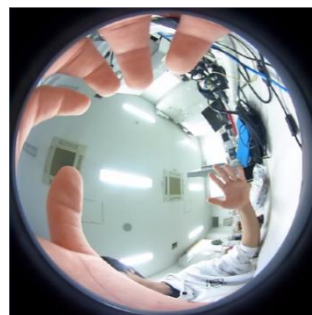


図2. カメラ画像



図3. 肌色領域検出



図4. 距離変換処理



図5. 図4の2値化

	母指	示指	中指	環指	小指
内転(外転)可動角	80	40	40	40	40

表1. 内転(外転)可動角

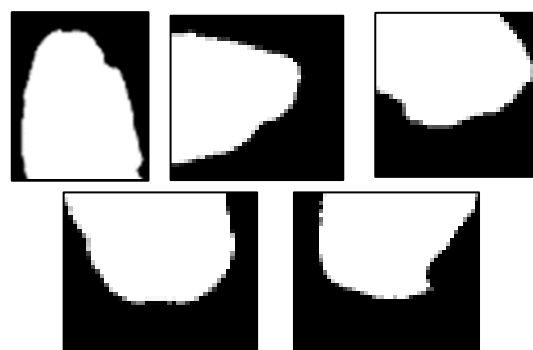


図6. 各指先のテンプレート画像

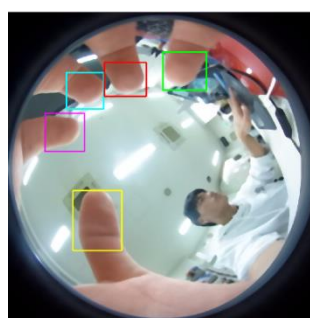


図7. テンプレートマッチング

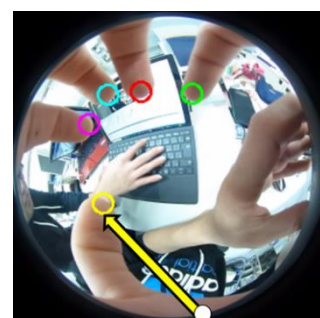


図8. 指先検出

次に手指領域から指先座標を決定する。実際のカメラと手指との位置関係より各手指の付け根はカメラの筐体に最も近く、カメラ画像において外側に位置する。そこでカメラ画像に手指の付け根の座標を仮定したとき、指先は手指領域の輪郭においてこの座標から最も距離が大きい座標である。仮定した手指の付け根の座標と手指領域の輪郭上のそれぞれの座標との距離を求め最も距離が大きい輪郭上の座標を指先の座標に決定する(図 8)。図 8 では親指を例に示しており、仮定した付け根の座標である白い点から手指領域の輪郭において最も距離が大きい点を指先座標に決定し、黄色い円で囲んでいる。

3. 手指姿勢の推定

本研究における手指の姿勢は各関節間の距離と各関節の角度によって成り立つものとする。本章では、検出した指先座標から手指の姿勢を求める手法について述べる。本章におけるシステムの流れを以下に示す。

- i. 指先とカメラとの位置関係を推定：魚眼レンズの画像への射影方法から、カメラ画像の中心と指先座標との相対位置を計算することでカメラレンズの中心からの指先への方向を求める。
- ii. 手指の各関節の角度を推定：各関節間の距離、カメラと手との位置関係を設定し、手指が自然な姿勢をとる際に生じやすい各関節角間の依存関係を利用することでカメラレンズの中心からの指先への方向から逆運動学的に手指姿勢を推定する。

人が日常生活の中で自然と手指を動かすとき、手指の各関節が屈曲するときの角度には依存関係が生じやすい。屈曲とは手を握る方向への手指の運動である。手指の各関節の屈曲角、各関節間の距離の値、カメラと手の位置関係を設定すると、この性質を利用することで指先座標から逆運動学的に各関節の角度を求められる。

3.1 θ_{tip} の計算

本研究ではカメラレンズの中心と指先を結ぶ直線とカメラレンズの底面との成す角度を θ_{tip} と定義する(図 9)。まず手指が動く軌道を、その付け根とカメラレンズの中心を通る直線の上であると仮定する。本研究ではこの直線を軌道直線と呼ぶ。画像において検出した指先座標から仮定した手指の付け根座標と画像の中心を結ぶ軌道直線に対して垂線を降ろしたときの交点を仮の指先座標に定め(図 10)、このときの θ_{tip} を求めることで軌道直線上における手指の姿勢を推定する。全天球カメラの魚眼レンズによる射影方法から θ_{tip} を求める。RICOH THETA に装着されている魚眼レンズでは立体射影により画像が生成されている。立体射影の性質と θ_{tip} の求め方を図 11 に示す。魚眼レンズによる仮想の半球上の点は、仮定した手指の付け根の座標と射影された指先の座標との距離 d 、半球の半径を r としたと

き図 11 中の式に基づいて画像へ射影されることから、指先座標を用いて θ_{tip} を求められる。

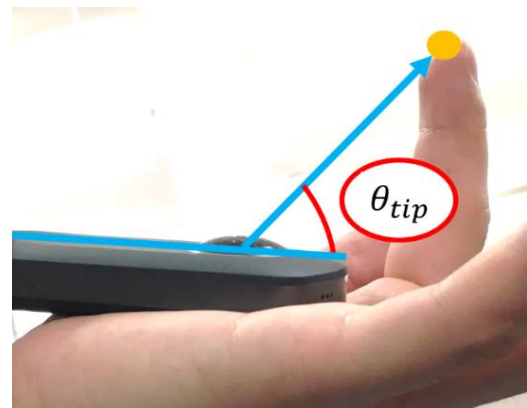


図 9. θ_{tip} の定義

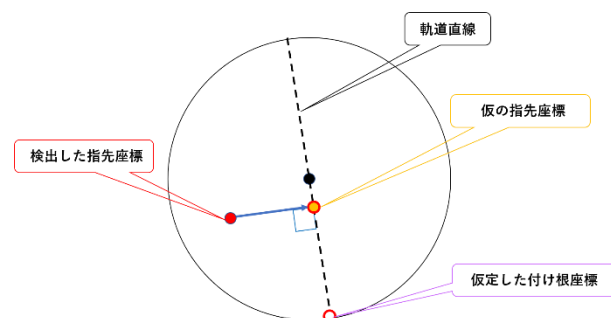
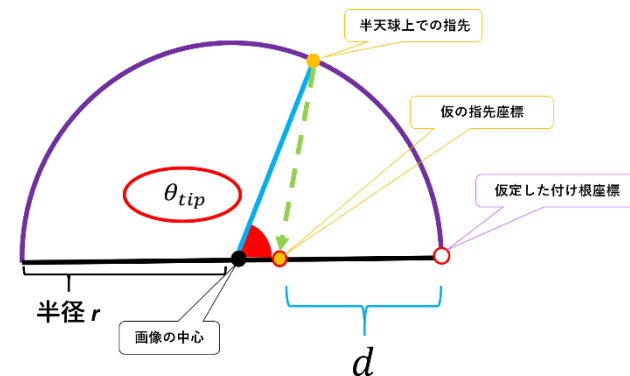


図 10. 仮の指先座標



$$\theta_{tip} = \begin{cases} 2 \tan^{-1} \frac{d}{r} & (0 \leq \theta_{tip} \leq \frac{\pi}{2}) \\ \pi - 2 \tan^{-1} \frac{d-2r}{r} & (\frac{\pi}{2} < \theta_{tip} \leq \pi) \end{cases}$$

図 11. θ_{tip} の計算

3.2 各関節の屈曲角の推定

各関節の屈曲角を求めることで軌道直線上における手指の姿勢を推定する。仮の指先と各関節、カメラと手の位置関係を図 12 に、各関節の名称を図 13 に示す。各関節間の距離 a , b , c 、掌の厚みの中心面から全天球カメラのレンズまでの高さ h 、カメラレンズの中心から MP 関節までの距離 l は予め与えるものとする。この図から各屈曲角

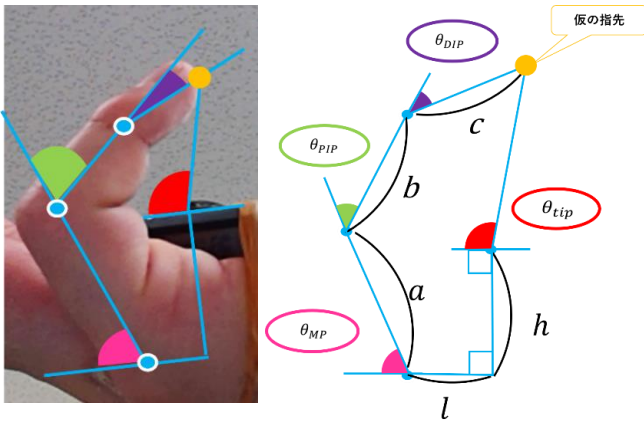
θ_{MP} , θ_{PIP} , θ_{DIP} と θ_{tip} を用いた方程式が導かれる. しかしこの方程式には未知数が3つ存在するため解くことができない. そこで, 人が日常生活の中で自然と手指を動かすときに生じやすい各関節間における屈曲角の依存関係を利用する. 藤木らが提案した依存関係を図14に示す(母指については括弧内の関節に該当). 図14から手指における3つの屈曲角の間に2つの依存関係を利用することができるため, 各屈曲角は1つの変数で表わすことができる. このことから図12によって導かれた方程式に存在する未知数は1つとなり方程式を解くことができるため, 各手指における3つの関節の屈曲角を求められる. 表2に示した屈曲角の可動範囲内で各屈曲角を求める.

	母指	示指	中指	環指	小指
CM関節	0~90	MP関節	0~90	0~90	0~90
MP関節	0~70	PIP関節	0~100	0~100	0~100
IP関節	0~50	DIP関節	0~80	0~80	0~80

表2. 屈曲可動角

3.3 手指の内転(外転)角の推定

手指の内転(外転)角を求めることで軌道直線上と仮定した手指の姿勢を実際の手指の姿勢に近づける. 手指の内転(外転)角は指先座標と軌道直線とのずれから計算することができる. カメラレンズの中心から手指の付け根までの距離 l の半球球を考えると, 半球球上のそれぞれの指先が軌道直線を中心になす角度 ω , h を用いて図15中の式を導くことができる. この式における ϕ が手指の内転(外転)角に対応する.



$$a \cos \theta_{MP} + b \cos(\theta_{MP} + \theta_{PIP}) + c \cos(\theta_{MP} + \theta_{PIP} + \theta_{DIP}) + l = \frac{a \sin \theta_{MP} + b \sin(\theta_{MP} + \theta_{PIP}) + c \sin(\theta_{MP} + \theta_{PIP} + \theta_{DIP}) - h}{\tan \theta_{tip}}$$

図12. 屈曲角の計算

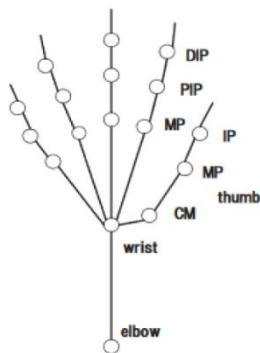
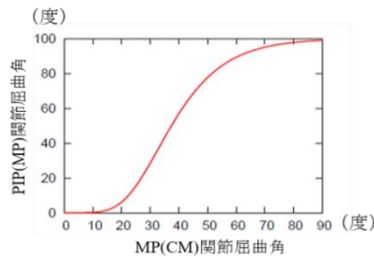
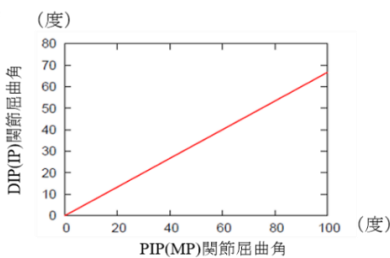


図13. 各関節の名称

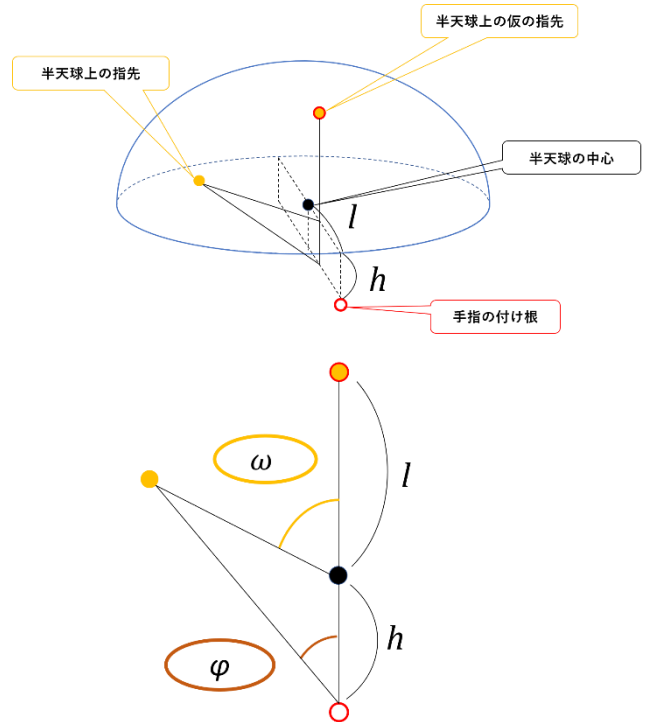


(a) MP-PIP(CM-MP)関節間



(b) PIP-DIP(MP-IP)関節間

図14. 屈曲角の依存関係



$$\phi = \tan^{-1} \frac{l \sin \omega}{l \cos \omega + h}$$

図15. 内転(外転)角 ϕ の計算

4. 評価実験

4.1 実験手法

本システムにおける手指姿勢推定の精度を評価する実験を行った. まず全天球カメラを装着した手の指を, 図10で示した θ_{tip} と各屈曲角を画像上で計測できるように横から撮影して θ_{tip} と各屈曲角を計測し, 同時に全天球カメラで得た画像から本手法を用いて θ_{tip} と各屈曲角を推定する.

手指を横から撮影した画像を手指画像とし、手指画像から計測した θ_{tip} と各屈曲角、全天球カメラで得た画像から本手法で推定した θ_{tip} と各屈曲角のそれぞれの値を比較して誤差を求める。またこのときの手指画像と本手法で再現した3Dハンドモデルを観察する。この実験を複数の手指姿勢において行うことで本手法の姿勢推定の精度と姿勢の再現性を評価する。手指画像と3Dハンドモデルを観察する際にオクルージョンが発生しないよう小指の姿勢を実験対象にする。また本研究では自然な手指の姿勢の推定が目的であり、この条件下での小指の内転(外転)角の変化が著しくないことから屈曲角のみを精度評価の対象に指定している。

4.2 結果

手指画像と本手法で推定した姿勢から再現した3Dハンドモデルを図16に示す。また図16(a)~(e)のそれぞれの場合における小指の手指画像から計測したと各屈曲角、また全天球カメラで得た画像から本手法で求めた θ_{tip} と各屈曲角、また双方におけるそれぞれの値の誤差を表3に示す。小指の手指画像から計測した θ_{tip} と各屈曲角の値を正解とし、本手法で推定された θ_{tip} と各屈曲角の値との誤差から本手法を評価した。表2に示した各関節の屈曲角の可動範囲に対し、図16(a)~(c)の場合においてそれぞれの誤差は5%未満であったため、実際の手指に近い姿勢の推定が可能であることを確認した。正確に指先を検出することができたため高精度に姿勢を推定できたと考えられる。しかし図16(d)、(e)に示した場合では指先が検出されず、また手指の再現性が低いハンドモデルとなってしまった。図16(d)の場合では、実際の手指がカメラレンズに極端に近くなり図17(a)に示すようなカメラ画像になることで手指を検知することができなかった。また図16(e)の場合では、手指が振り返った姿勢をとっているため図17(b)に示すように手指がカメラの画角内に現れず手指を検知することができなかった。これらの実験結果から、本手法において推定できない場合の手指姿勢を把握することができた。

4.3 考察

図16(d)のように手指とカメラレンズとの距離が極端に小さい場合には手指が画角を覆ってしまうため図17(a)のようなカメラ画像に、図16(e)のように手指が振り返っている場合には手指がカメラの画角内に現れず図17(b)のようなカメラ画像になり手指を検知できないため姿勢を推定できないとわかった。そこで本手法の改善に、時系列的に指先の動きを推定することで手指の姿勢を推定する手法が考えられる。手指を検知できなくなった場合に、その直前までに推定された画像上での指先の動きから指先座標を補完し、姿勢を推定することで本手法における問題点を解決する。

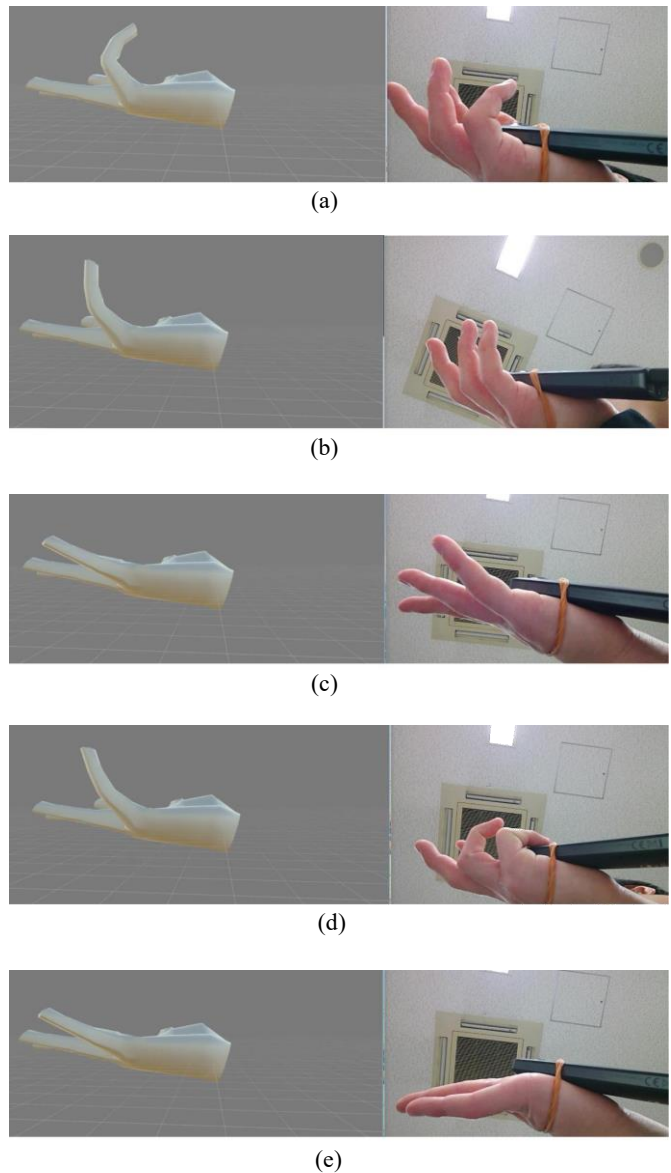


図16. 各手指姿勢の再現

5. まとめと今後の展望

本研究では、掌に装着した全天球カメラで得た画像から自然な手指の姿勢を推定する手法を提案しシステムを作成した。まず、カメラ画像から手指を検知し、指先を検出することで指先座標を得た。検出した指先座標から軌道直線上に仮の指先座標を定め、魚眼レンズによる射影方法から仮の指先座標を用いて θ_{tip} を求め、各関節間の距離などの値を予め与えた上で各屈曲角における依存関係を利用することで一意に各屈曲角を求めた。また仮の指先座標と実際の指先座標との位置関係から内転(外転)角を求め、姿勢を推定した。

提案手法による姿勢推定の精度評価を小指の姿勢から行った。複数の手指の姿勢において、実際の手指の姿勢と本手法での θ_{tip} と各屈曲角を求めてそれぞれを比較することで本手法の姿勢推定精度を評価し、また本手法で再現した3Dハンドモデルと実際の手指を画像上で目視して姿勢の

再現性を評価した。実際の小指の姿勢における θ_{tip} と各屈曲角の値を正解とし、正解の値に対する本手法で得た各値との誤差から、正確に指先を検出できる場合には高精度な姿勢推定が可能であることが確認できた。しかし実際の手指の姿勢が、カメラレンズとの距離が極端に近い場合や振り返っている場合には手指を検知できず姿勢を推定できな

	本手法	写真上	誤差
θ_{tip}	76.7	80.0	3.3
θ_{MP}	51.1	53.3	2.2
θ_{PIP}	77.3	79.7	2.4
θ_{DIP}	45.6	47.0	1.4

(a)

	本手法	写真上	誤差
θ_{tip}	61.5	61.0	0.5
θ_{MP}	41.1	40.7	0.4
θ_{PIP}	60.5	59.4	0.6
θ_{DIP}	35.7	35.0	0.7

(b)

	本手法	写真上	誤差
θ_{tip}	17.9	17.0	0.9
θ_{MP}	11.9	11.3	0.6
θ_{PIP}	3.0	2.6	0.4
θ_{DIP}	1.8	1.5	0.3

(c)

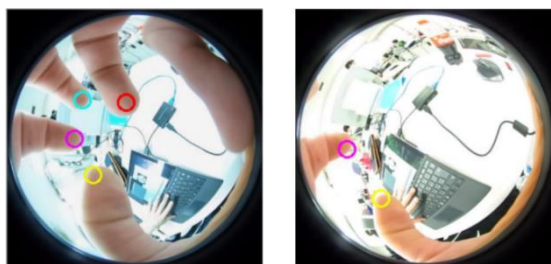
	本手法	写真上	誤差
θ_{tip}	-	180.0	-
θ_{MP}	-	90.0	-
θ_{PIP}	-	100.0	-
θ_{DIP}	-	60.0	-

(d)

	本手法	写真上	誤差
θ_{tip}	-	-23.0	-
θ_{MP}	-	-23.0	-
θ_{PIP}	-	0.0	-
θ_{DIP}	-	0.0	-

(e)

表 3. 実験結果



(a)

(b)

図 17. 姿勢推定精度が低かった場合のカメラ画像

かった。そこで手指の姿勢推定精度を改善するために、時系列的に指先の動きを推定することで手指の姿勢を推定する手法を検討している。手指が検知されなくなる直前までに推定された画像上での指先の動きから指先座標を補充することで、手指を検知できなくなった場合でも姿勢を推定できる。

提案手法は手指の姿勢を扱う AR や VR などのシステムへの応用が可能であり、ユーザの運動範囲を制限しないためユーザはより直感的な操作が可能となり、臨場感を味わうことができると考えられる。また画角の大きい全天球カメラを用いているため広範囲で他方の手やマーカなどの検出が可能であると考えられる。その他に、カメラに加えて加速度センサやジャイロセンサなどを扱うことで手の動きや姿勢をより高精度に推定できると考えられる。現在、本研究の応用に AR ろくろ回しを検討している。ろくろを出現させるためのマーカと左右それぞれの手の相対位置姿勢をリアルタイムに推定することでビデオシースルー型 HMD を用いた AR ろくろ回し体験を実現する。本研究における手指姿勢推定の手法では手指の自然な姿勢を対象にしており、例えば机を持ち上げるときなどの MP 関節のみが曲がるような依存関係を無視する姿勢を対象にすることはできない。しかし依存関係に従わない手指姿勢では手指に負担がかかりやすく、手指は依存関係に従う自然な姿勢をとる傾向にあるため、物体を把持するシステムに利用しない限り本研究には有用性があると考えられる。本研究では手指姿勢推定のためのカメラに筐体が薄く画角の大きい魚眼レンズが装着された RICOH の THETA S を用いているが、手指の動きを妨げないよう提案した装着方法ではユーザは掌からカメラの筐体がはみだしていることに注意して行動しなければならない。そのため本手法の優位性が認められてより小型のカメラが開発されることを期待する。

参考文献

- [1] Weichart, F., Bachmann, D., Rudak, B., and Fisseler, D.. Analysis of the Accuracy and Robustness of the Leap Motion Controller. Sensors, 2013, Vol.13, pp. 6380-6393.
- [2] "Thalmic Labs". <https://www.myo.com/>, (参照 2017-10-17).
- [3] 柏木直諒, 杉浦裕太, 宮田なつき, 多田充徳, 杉本麻樹, 齋藤英雄. 組込み型カメラによる把持動作時の人体手形状 3D 復元. 映像情報メディア学会誌, 2017, Vol.71, No.11, pp. J283-J286.
- [4] Ueda, E., Matsumoto, Y., Imai, M., and Oasawara, T.. A Hand-Pose Estimation for Vision-Based Human Interfaces. IEEE Transactions on Industrial Electronics, 2003, Vol.50, No.4, pp. 676-684.
- [5] Gentile, V., Sorce, S., and Gentile, A.. Continuous Hand Openness Detection Using a Kinect-Like Device. CISIS, 2014, pp. 553-557.
- [6] Sherrah, J. and Gong, S.. Skin Colour Analysis. http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/GO_NG1/cvOnline-skinColourAnalysis.html, (参照 2017-10-17).
- [7] 藤木隆司, 陣維英, 吉本廣雅, 有田大作, 谷口倫一郎. 手の制約知識を用いた逆運動学解法による実時間三次元手形状推定. 画像の認識・理解シンポジウム, 2005, pp. 1380-1387.