

コンピュータビジョンの最新動向と新たな応用

ロベルト チポラ

カルロス ヘルナンデス

ジョージ ボジャティス

ビヨン ステンガー

東芝欧州研究所 ケンブリッジ研究所 コンピュータビジョングループ
ケンブリッジ大学 工学部

1 コンピュータビジョンの三つのR

コンピュータビジョン技術は携帯電話カメラ、ゲームの入力インタフェース、自動車の駐車や運転の支援、及びコンピュータやインターネット、インターネットショッピングでの画像やビデオ検索など、一般製品に多く使われるようになってきている。

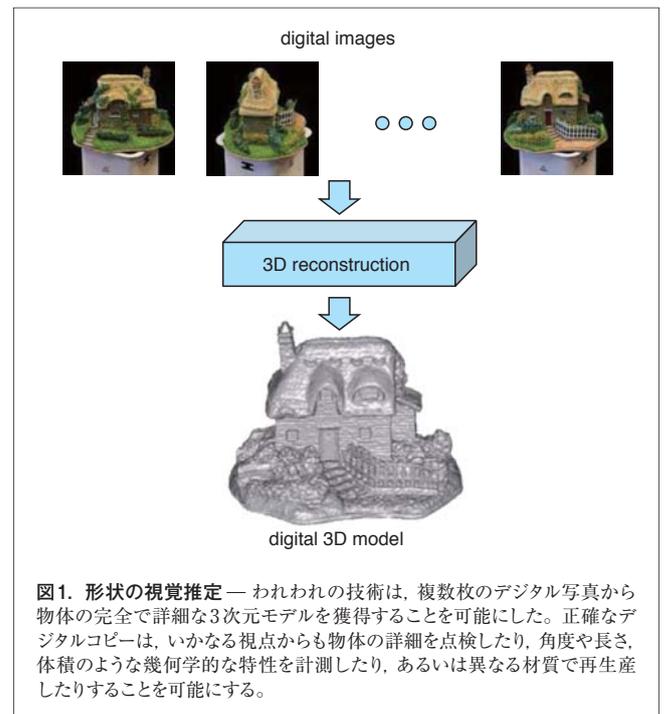
東芝欧州研究所 ケンブリッジ研究所では、多くの実用可能なアプリケーションを見だし、また3次元視覚の分野での新しいコア技術を開拓している。具体的には、再構築 (Reconstruction: キャリブレーションなしの画像からの3次元形状復元) と、登録 (Registration: 今までにないインタフェースによる人間の身体を検出と追跡)、認識 (Recognition: 動画からの物体検出、切出し、及び認識) というビジョンの三つの分野 (三つのR) である。われわれのアプローチの大部分は幾何学であるが、機械学習の最新の実績も含んでいる。

ここでは、①3次元形状の再構築と②手の検出と追跡による簡易で頑健なコンピュータインタフェースの二つの分野での研究を簡単にレビューする。

2 形状の再構築

形状の再構築に関するわれわれの研究は、対象とする物体の画像シーケンスから、現実世界の3次元物体の完全で詳細なモデルを獲得するための課題を調べることから始めた。この課題は計算アルゴリズムの構築を通じて、人間の視覚システムの理解を目指す研究者たちによって、マシンビジョン (例: 参考文献(7)) の初期から大規模に研究されてきている。最近では、コンピュータ処理能力の劇的な改善とデジタル画像技術の利用の進展により、画像からの形状再構築は実用アプリケーションとして興味を持たれるようになってきた。

物体の現実的で今までにないシーン (図1) を合成するための正確な幾何学モデルは非常に価値がある。このようなモデルを獲得するもっとも一般的な方法は、CADプログラムを使って手動で構築するか、レーザレンジスキャン技術 (例: 参考文献(5)) を使うかである。手動による方法は大規模でかつ複雑なモデルに対しては、極めて現実的でなく、またエラーが発生しやすい。一方、レーザスキャンやほかの類似技術は潜在的な応用に幅広く使うには極めて高価である。



このように、シーンのデジタル画像から写実的な3次元モデルの自動的な獲得は、考古学⁽⁹⁾や、建築のモデル⁽¹⁾、彫刻のデジタル化⁽⁴⁾などの応用に対して、安価でかつ手軽で非侵襲な代替案となるものである。

2.1 挑戦

初期のコンピュータビジョン研究者の楽観主義にもかかわらず、完全に自動的な視覚による再構築システムはとらえどころがないままである⁽³⁾。要となる困難さは、参考文献(11)から引用したものであるが、以下のとおりである。

- (1) 高次元性 一般的なシーンの幾何表現は、飛躍的に多くの自由度を要求する。幾何と反射率の双方に強い前提を設けないかぎり、この未知要素を推定することは、通常は実行不可能である。
- (2) 写実的なあいまいさ 画素ごとの観測された強度は、対応するシーンの点の表面幾何や、照明と同様に局所的な反射に依存している。これらの特性は画素ごとの強度から拘束されるが、直接推定することはできない。
- (3) 奥行き喪失 シーンのカメラ画像は、3次元空間を2次元平面に投影することで形成される、この過程では、

シーンとカメラの間で光が移動する距離、すなわち奥行きが失われる。濃淡からの形状のように、単眼のケースでは、このあいまいさは再計算可能という状況ではあるが、人間や人工的な視覚システムは、多様な照明の下であっても多様な視点からの複数のシーン画像を採用している。

われわれは、固体物体の異なるクラスに適する視覚的な形状再構築問題への二つの異なるアプローチを開発した。第1の方法は、リッチなテクスチャ（模様）の材質からなる物体を扱うものである。第2の方法は、白磁のようなテクスチャがまったくない物体を扱うものである。

2.2 テクスチャのある材質

画像対応からの形状推定は、時には単に“高密度ステレオ（dense stereo）”として参照されるが、“異なる”視点からM個の画像が与えられているシーンを再構築するには、たいへん有効な技術である。それは以下の非常に単純な観察に基づいている。

“シーン表面上にある”3次元の点は、ふさがれていないすべての画像での“同様の”見え方がする画像領域に投影される。同等に、この原則は視覚光線の状態で記述されている。上述のように、各画像位置は3次元線分に対応する。同じシーン位置を描写するたくさんの画像位置があれば、視覚光線の交点^(注1)はシーン位置となる。これを描いたのが図2である。

高密度ステレオ問題での多くの研究は、シーンシーケンスを通じて一定した照明と同様に、表面については均等拡散反射モデルを仮定している。画像位置が対応するかどうかの判断

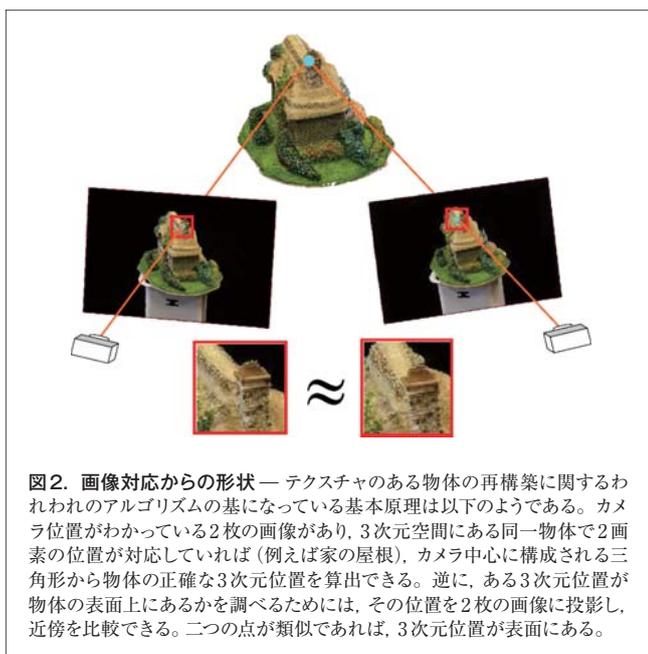


図2. 画像対応からの形状 — テクスチャのある物体の再構築に関するわれわれのアルゴリズムの基になっている基本原理は以下のようである。カメラ位置がわかっている2枚の画像があり、3次元空間にある同一物体で2画素の位置が対応していれば（例えば家の屋根）、カメラ中心に構成される三角形から物体の正確な3次元位置を算出できる。逆に、ある3次元位置が物体の表面上にあるかを調べるためには、その位置を2枚の画像に投影し、近傍を比較できる。二つの点が類似であれば、3次元位置が表面にある。

(注1) わずか2枚の画像でカメラ中心が動けば、視覚光線はすべて一致しなくなり、交点が決められる。

を簡単にするこの条件は、シーンの点が見えている画像で同一強度の画素に投影することを意味している。加えて、物体はテクスチャを持っているので、物体表面の一部は複数画像で一意に同定できる。

われわれが開発した3次元再構築問題向けの体積型定式化は、グラフの最小切断を使って、計算的に制御しやすく大局的な最適化をしやすくなっている。われわれのアプローチは、以下に示す2条件で構成されるコスト計算のもとに、3次元空間を“物体あり”と“物体なし”にラベル分けする最適分割を探索するものである。



図3. おもちゃの家 — この図はわれわれの技術で獲得した現実物体の3次元モデルの例である。上段には、おもちゃの家の4枚の画像が並んでいる。下段には、類似の視点から見た3次元モデルの描画結果が並んでいる。最初の3枚のイメージは入力シーケンスとして使われたもので、最後の4枚目は入力には使われていない。この小さなおもちゃの家（直径約10 cm）のモデルは、フェンスや屋根のレリーフのような1 mm以下の詳細まで正確に再構築している。

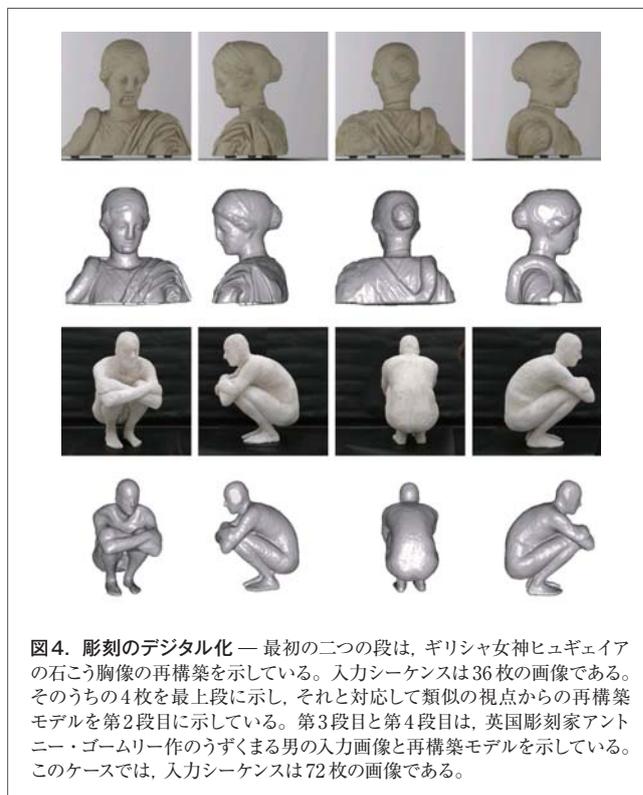


図4. 彫刻のデジタル化 — 最初の二つの段は、ギリシャ女神ヒュギエイアの石こう胸像の再構築を示している。入力シーケンスは36枚の画像である。そのうちの4枚を最上段に示し、それと対応して類似の視点からの再構築モデルを第2段目に示している。第3段目と第4段目は、英国彫刻家アントニー・ゴームリー作のうずくまる男の入力画像と再構築モデルを示している。このケースでは、入力シーケンスは72枚の画像である。

- (1) “物体あり”と“物体なし”の二つの領域の境界が、写真として一貫性がある位置を通過させる条件
 - (2) “物体あり”領域を膨らませるための膨張条件
- (1)の条件で遮蔽(しゃへい)の影響を考慮して、再構築物体について幾何的な知識を仮定しない正規化相互相関(NCC: Normalised Cross Correlation)をもとに、遮蔽に頑健で写真として一貫性がある尺度を用いる。大局的に最適な3次元分割は、重み付けグラフの最小枝刈り解法により可能である。この方法を用いてデジタル化した物体のいくつかを図3と図4に示す。

3 照度差ステレオ法 (Photometric Stereo) とテクスチャなしの材質

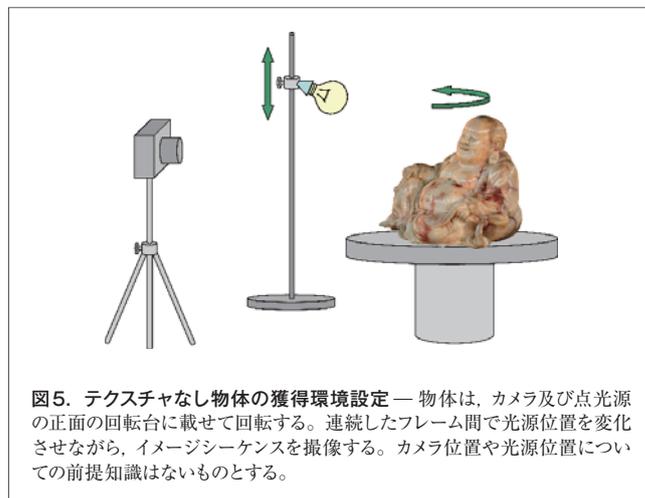
高密度ステレオ技術が詳細で完全な3次元再構築を可能とする。一方、この方法は、形状獲得用の三角形化された複数画像上の位置間の対応を獲得するために、リッチなテクスチャをもつ物体に依存している。その結果、テクスチャがない物体では表面特徴が欠如するために、この高密度ステレオ法を直接応用できない。

一方、照度差ステレオ法は、照明の変化として物体表面の点の画像強度での変化を観察することで働くものである。これらの変化は、3次元形状を規定する点での局所的な表面の方位を明らかにする。

照度差ステレオは奥行きを再生する積分を行うので、まして正規化は必要ではなく、結果は一般的により詳細化される。更に、照度差ステレオは表面のテクスチャや反射についてほとんど前提を設けないので、参考文献(2)でデモされているようにほぼ完璧に任意である。しかしながら、複数画像で表面の強度が同一の点を集めるもっとも簡単な方法は、カメラ視点を一定にすることである、換言すれば、各画素はいつも表面の同じ点に対応している。これが照度差ステレオ法の制限事項である。したがって、照度差ステレオ法では、彫刻のような複雑で多側面物体の完全な3次元幾何の再生はできない。この制限により、参考文献(12)、(6)のような注目すべき例外はあるが、現存の照度差ステレオ技術は、奥行き情報を調べる(例: 参考文献(10))程度にしか使われてきていない。また、著者らが、複数視点から2.5次元の再構築を行う技術を示している。しかしながら、これらの照度差ステレオ法では、多側面物体の完全な再構築はいまだに不可能である。照度差ステレオで複数視点から得た物体の複数の奥行き情報を一つの3次元表現に合わせ込むことも行われているが、実際には、処理が複雑でエラーが発生しやすい。

われわれは強力なシルエット手がかりを有効に活用して、この問題への異なる解法を見いだした。古典的な照度差ステレオ法を修正し、カメラが物体をぐるりと回り、照明が変化する

複数視点のフレームワークで動くようにした。この技術の環境設定を図5に示した。まず、参考文献(8)に類似した技術を使ってカメラモーションを再生するのに物体のシルエットはよく使われる。次に、今までにない頑健な推定により、すべての画像での光方向と強度を正確に推定する。



第2に、物体の表面はメッシュ切りされ、視覚構造から初期化され、予測された見え方が撮像された画像に対応するまで展開される。メッシュの各面は、可視で強度が正しい画像に投影される。事前に計算された強度と照度から、最小二乗法を解いて各面に法線方向をアサインする。メッシュの現実表面の法線をカバーするまで、この計算は繰り返す。メッシュが真の表面をカバーするまで、この2段階処理は繰り返される。われわれのアプローチの利点は以下のとおりである。





図7. 色つき大理石の再構成 — 大理石の仏像 (中国, 清王朝) は, 36枚の画像シーケンスから再構成されている。

- ・ キャリブレーション不要である。光やカメラ姿勢のキャリブレーションなしで, シーンに物体があればよい。
- ・ 複雑で, 光っている, テクスチャのない物体の完全な3次元幾何が正確に再構築される。ほかの方法では, 不可能であった。
- ・ 簡単な獲得環境設定で示されているように, 実用的かつ効果的である。

図6と図7は, ケンブリッジのフィッツウイリアム美術館の収集品からデジタル化した美術品を示している。

4 変形する形状の再構成

前章で剛体物体 (変形しない物体) の3次元モデリングの最先端技術がどれだけ進歩しているかを紹介したが, 今までのところでは, 非剛体物体 (変形する物体) シーンについては, ほんとうにわずかの研究しか達成されていない。変形物体のモデリングは生地のようなタスクだけでなく, 身体や顔のような一般的なダイナミックな表面についても非常に重要である。この章では, 変形表面の詳細な3次元形状, 曲げやしわを獲得するために, 連想キャプチャ方法論を説明する。既知の表面特性や, 構造化された光, シルエットに依存する方法では, 移動する3次元データを得ることは困難であった。

マルチスペクトルの照度差ステレオは, テクスチャのない表面から高密度法線フィールドを再生できるので, 魅力的な代替案である。空間的に分離された赤, 緑, 青の光源の下で撮像されたテクスチャのない生地の動画シーケンスを用いて実験を行っている。

現実の動いている生地から複雑なモーションデータを獲得するための提案技術は, 通常のビデオカメラと3色の光源からなる実用的な環境設定 (図8) を使っている。要となるのは, 赤, 緑, 青の光が異なる方向から照射されている環境で, 均等拡散表面が, どの周波数もミックスすることなく, 同時に3色を反射することである。赤, 緑, 青の反射量は表面法線方

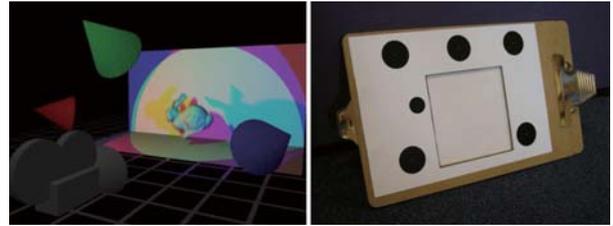


図8. 環境設定とキャリブレーションボード — 左: マルチスペクトルの環境設定の概要, 右: 2枚のクリップボードが繋がっていて, そこに印刷されたキャリブレーションパターンがはり付けられている。パターン平面の方向を計算して平面の対象を追跡できる。生地は2枚のボードの間の四角の穴に挟み込まれ, 色と方向の連想から計算ができるようになる。

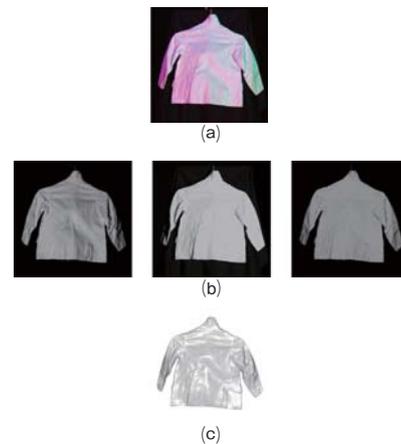


図9. 色付き照明での照度差ステレオ — (a)三つの異なる色付き照明で同時に照らされたジャケットのシーケンスからの1枚, (b)それぞれが異なる色に対応する三つのグレースケール画像, (c)古典的な照度差ステレオ法での再構成¹⁷⁾での入力例。

向の線形関数である。カラーカメラがこれらの量を計測し, その量から表面法線方向を推定できる。変形物体の動画シーケンスにこの技術を応用することで, 物体の法線マップのシーケンスを得ることができ, それらを統合して奥行きマップのシーケンスを作成できる (図9)。

4.1 奥行き情報ビデオ

この節では, Kontsevicらの研究¹⁵⁾の説明を行う。簡単のために, はじめに法線方向 \mathbf{n} の表面で位置 P の均等拡散表面を方向 \mathbf{l} から照らす一つの離れた光源があるケースに焦点を合わせる。光源のエネルギー分配を波長 λ の関数として $S(\lambda)$ とする。表面点での反射特性を表現するスペクトル反射関数を $\rho(\lambda)$ とする。カメラ画素はスペクトルの異なる部分に反応する三つのセンサからなると仮定する。 $N_i(\lambda)$ が P からの光を受け取るがその i 番目のセンサのスペクトル検出感度とすれば, センサごとの強度は以下ようになる。

$$r_i = \mathbf{l} \cdot \mathbf{n} \int S(\lambda) \rho(\lambda) v_i(\lambda) d\lambda \quad (1)$$

あるいは、行列形式で次のようになる。

$$\mathbf{r} = M\mathbf{n} \quad (2)$$

ちなみに、 M の (i, j) 番目の要素は次のようになる。

$$m_{ij} = l_j \int S(\lambda) \rho(\lambda) v_i(\lambda) d\lambda. \quad (3)$$

光源が加わると、システムがリニアで各光について $1 \cdot \mathbf{n} \geq 0$ であるならば、各センサの応答は個別の各光源からの反応の総和になる。ただし、式(2)にいたるには、式(4)が成り立つことを確認する必要がある。

$$M = \sum_k M^k \quad (4)$$

M^k は k 番目の光源を示す。各 M^k はランク1であるので、自己遮蔽の場合は、可逆であるためには、少なくとも三つの異なる光が必要であることを意味する。表面に色が付いていれば、 $\rho(\lambda)$ とその結果として M は遮蔽されていないすべての箇所まで一定である。

式(2)は、カラーカメラで計測されるRGB(赤, 緑, 青)画素と画素に投影される点での表面の方向間の1対1マップを成立させる。われわれの戦略はこの1対1マップを逆に使って、変形表面のビデオを法線マップに変換することである。“使いやすい”キャリブレーションツール(図8の左)を採用することで、1対1マップを推定する。パターンは平面の方向を推定できるように特別なマークがついている平面である。パターンの中心に物体を置き、現在の方向で反射する色を計測する。最小二乗法を使ってマップ M に適合する (\mathbf{r}, \mathbf{n}) のペアのシーケンスを獲得する。

4.2 法線からの奥行き

RGB付きの線形マップ M を推定し、表面法線に逆変換することで、色付き照明の下で撮像したビデオシーケンスをビデオ法線マップに変換できる。

暗い部屋の条件では、単純な強度しきい値で、オリジナルビデオはほぼ完全な黒なので、すべてのフレームで背景画素を分割できる。輪郭は奥行きがゼロであるので、すべてのフレームの奥行き情報を得るために各法線マップを統合する。この統合プロセスはよく確立された技術であり、いくつかのアルゴリズムが使える。逐次過緩和法(SOR)¹³が頑健で単純であるので、われわれはSORを使っている。統合プロセスで最終的に、“奥行き情報ビデオを獲得できる”。

図10で、高解像度(メッシュはおおよそ180,000頂点からなる)で、テクスチャマップなしに380フレームから複数のアングルを示す。画像はセーターの高周波の成分である細かい凹凸までくっきり示している。われわれの知るかぎりでは、この方法は、図10に示す詳細さで変形生地を再構成できる唯一の方法である。

アニメーションで生地撮像ができるわれわれの方法の潜在



図10. 変形するセーターの生地再構成結果 — 4.1節で述べた技法を使って、長いビデオシーケンスの1枚のフレームでのマルチスペクトル照度差再構成。左から右にセーターシーケンスの380枚から複数のアングル。



図11. アニメーションキャラクターに獲得した動く生地をはり付けた例 — スムーズスキニングを用いて、モーションキャプチャ(MOCAP)データで人工的な骨格をアニメーションするように、骨格に動くメッシュをはり付けた。撮像して登録したシーケンス(この場合はダンスシーケンス)を再生することで、メッシュが単純にアニメーションされる。

能力をデモするために、動いている生地シーケンスを撮像して、オプティカルフロー¹⁴を使って登録し、人工的な骨格に記憶したメッシュをはり付けた。スキニングアルゴリズムは、例えば参考文献(16)のようにリアルさと複雑さとの多様性を扱う。図11は描画されたシーケンスからのフレーム例を示す。骨格と生地の動きは明白にはそろっていないけれど、制御可能なキャラクターでの生地の動きの視覚効果は魅力的である。このデータ駆動型の生地アニメーションは有用なツールであり、物理的な生地シミュレーションの代替案としても使える。

4.3 要旨

照度差ステレオ法のためのマルチスペクトル照明の理論はずいぶん以前に確立されたが、驚くことに見落とされてきた。われわれはいくつかの新たな障害を見だし、それを解決してきた。静的な生地を撮像する現存の研究に匹敵する撮像方法を開発し、更に、動いて形状が変化する生地を撮像することも可能とした。輪郭を奥行きゼロとする単純な境界条件を設けることで、法線フィールドの統合も既に可能になった。古典的な照度差ステレオ法に対して、奥行き情報の正確さを評価している。表面シーケンスの再生は、スムーズに変化している。高度な詳細さは、法線フィールドにより、表面の曲げ、しわ、及び折りたたみじわさえも撮像している。

最後に、リッチな3次元生地姿勢のストリームへアクセスすることで、データを用いるだけで簡単に、洋服を着たアバター(分身)のリアルなキャラクターアニメーションを創造的に生成できることも示した。

5 コンピュータビジョンを使った ジェスチャ ユーザーインターフェース

コンピュータビジョンは手のジェスチャで、触れることのない入力を可能にする。東芝 研究開発センターとともに、頑健で効率的で、直感的な方法を研究している。

マウスとキーボードは多くのパソコン (PC) の標準的な入力デバイスとなっている。ポインティングデバイスとしてのマウスには多くの限界がある。まず第1に、いくつかのアプリケーションでは、マウス入力の2自由度より多くの自由度があれば、便利である。画面上の物体の操作 (例えば引伸し) のようなあるアプリケーションの制御では、両手が使えると実用的である。人間工学的見地から、キーボードとマウスを切り替える入力は、煩雑であり、マウスの繰り返し使用は精神的に苦痛である。手書きやお絵かきのアプリケーションにとっては、自然な入力デバイスではない。デザイナーはそのようなタスクには正確であることからお絵かきパッドを使うが、特別な投資であり、初期学習も必要である。マウスもお絵かきパッドも特別な机上のスペースを必要とする。しかしながら、最近では、コンピュータにCCD (電荷結合素子) カメラが付いていたり、テレビ電話やテレビ会議アプリケーションが付いていたりする。

5.1 画像位相幾何学の変化検出による ポインティングインターフェース

この節では、カメラがキーボードの方向に向けて画面のトップに設置されている、特別あつらえのPCへの直感的で視覚的なポインティングインターフェースを紹介する。ユーザーは簡単にキーボードとジェスチャ入力を切り替えることができる。更に、ジェスチャ入力モードは伝統的なコンピュータマウスより多くの自由度を持っている。ジェスチャ入力は、手書きや、イメージ操作、視覚的ナビゲーションのようなタスクにとって自然なインターフェースを提供する。

手は前景物体として見えるので、色分割により背景から分離される。方法としては、背景の形状位相幾何学の変化を検出する接続部位計算アルゴリズムを用いる。例えば人さし指と親指のタッチのような2本の指先のタッチによって、背景分離の数は一つ増加する。この新しい形に楕 (だ) 円形を合わせることで、手の位置、方向と大きさを決定できる。加えて、手の向きを推定し、指先を近づけることで点を制御できる。

図12はシステム設定を示す。

- (1) 形状分析 第1の処理ステップとして、システムは前景分離を行う。手検出器^{19), 20)}を使ってカラーモデルを初期化して前景分離を行う。図13に示すように、皮膚の色は検出された領域中心の楕円から抽出され、近傍は背景として扱われる。計算された分散は、各画素の皮膚色尤 (ゆう) 度の計算に用いられる。この値は2値分離アルゴリズムの入力として使われる。形態的な開閉操作に伴うヒ



図12. システム設定 — カメラは、ユーザーの手に向けてモニタの上に設定されている。ポインティングデバイスは、前景分離と形状分析により実現される。



図13. カラーモデルの初期化 — 左: 開いた平らな手 hands が手検出器を使って見つかった。皮膚の色分布は検出領域の中心の値から学習する。中央: 皮膚色の尤度。右: 後処理された2値マップ。

ステリシスしきい値処理により、画素ノイズに起因する誤分類を除くことができる。典型的な結果は、図13の一番右側に示す。木製テーブルと同様に、画面での手の反射により誤分類画素が生じることが留意すべきである。

次のステップは、前景マップの形状の分析である。背景は通常、1接続部位構成と仮定する。例えば、親指と人さし指が輪の形になっているように、前景がグループとなっているときは、この仮定に該当しない。この原則は、ポイントとクリックインターフェースの実装に使っている。接続部位計算アルゴリズムは、接続分割の数のカウントに用いる。二つ以上の分割があるときには、ポインティングインターフェースの入力領域として扱う。第1に、各領域に楕円を当てはめる。四つの楕円の変数は、インターフェースの入力として直接に使われる。従来のマウスの二つの変数に加えて、方向とスケールを扱うことができる。

- (2) 指先検出 相補的形状分析技術で指先領域を検出する。ここで、効率的だが、比較的明瞭な前景分離を得られる単純なアルゴリズムを紹介する。まず、前景輪郭を計算し、輪郭上に指先が載っている点があると仮定する。すべての輪郭点の近傍で前景/背景比率を計算し、二つのしきい値にある比率だけを保持する。1本の指先候補の周りのたくさんの検出候補の中で、もっとも小さな比率を持つものが出力位置として選択される。この方法の利点は、①評価が高速である、②影などの影響でノイズな輪郭形状に比較的耐性がある、ことである。一方、こ



図14. 指先検出例 — 第1段目では、指の輪郭あるいはその一部のポイントが見えているときに限るが、すべての指が同時に検出されている。第2段目では、手の回転は2次元回転として正確に記述できる。第3段目では、スケール操作は指をドラッグすることで実現できる。

の方法は、比率を計算する局所領域を決定するために、大ざっぱなグローバルスケールを必要とする。

- (3) 実験結果 指先検出アルゴリズムは大変効率的で、様々な入力モードに潜在性がある。図14に示すように、イラストで示している。個々の指先は検出できるが、指が互いに触れると、前景輪郭形状が変化してもはや検出できない。

イメージ位相幾何学に基づく提案アルゴリズムは、リアルタイムで稼働する。ポインティングアルゴリズムの結果を二つ示す。図15は、片手あるいは両手を使ったポインティングのシーケンス例である。マウスポインタのように使うことができる。指が閉ループを形作らなければ、カーソルは動かなくなる。位置に加えて、方向とスケールも確度高く検出できる。簡単な手書きアプリケーションが図16に示されている。手はペンを持っているように動かす。指が置かれると、描画モデルが活性化され、線分が画面上に描かれる。

このシステムの利点は、スケールと方向の推定による自由度の追加だけでなく、タイピングとポインティングとを切り替えられることである。更に、複数検出はシームレスに扱われる。こ

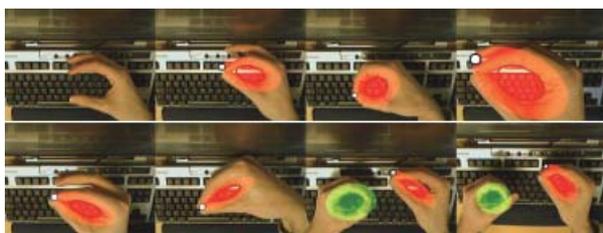


図15. シーケンス例 — 入力シーケンスのフレームである。ポインタは、親指と人さし指で閉ループを作ると活性化される。一つあるいはそれ以上の領域で、位置、方向、スケールが定義され、入力インタフェースとして使える。

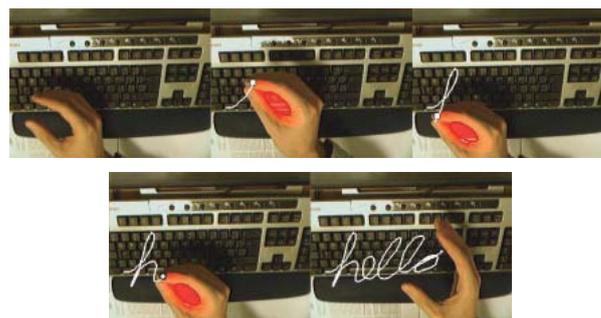


図16. 手書きアプリケーション — 手書きは自然なインタフェースである。

の入力方法は、いくつかのケースで描画パッドに置き換わるかもしれない。指先検出は正確な入力としての潜在性があるが、いくつか解決すべき課題がある。その一つが、検出精度の向上と、指がほかの指にタッチしていることの検出である。今後これらを研究していく。

6 結論

ここでは、東芝欧州研究所 ケンブリッジ研究所で実行されているプロジェクトのうち、四つについて非常に簡単に概要を述べた。制御された屋内環境でのキャリブレーションなしの画像からの3次元形状再構築分野では、特筆すべき進展を成し遂げた。次なる大きな挑戦は、照明に対するより少ない制御で屋外環境でのモデルを構築することと、より効率的なアルゴリズムを確立することである。

手の検出と追跡については、より信頼のおけるアルゴリズムのために、より多くの研究が必要である。東芝欧州研究所 ケンブリッジ研究所と東芝 研究開発センター マルチメディアラボラトリーとで開発した既存の技術は機が熟し、コンピュータやテレビの簡単なジェスチャインタフェースとして活用されている。しかしながら、コア技術はほぼでき上がっているとはいえ、簡単に直感的に、有用でかつ使って面白いインタフェースとするためには、多くの慎重なHCI (Human Computer Interaction) デザインが必要である！

文 献

英文本文の文献 (p.10) を参照。

和 訳

土井 美和子
研究開発センター技監