

THE IEICE TRANSACTIONS ON INFORMATION AND SYSTEMS (JAPANESE EDITION)

EIC | **電子情報通信学会** **D** | **論文誌**

情報・システム

VOL. J101-D NO. 3
MARCH 2018

本PDFの扱いは、電子情報通信学会著作権規定に従うこと。
なお、本PDFは研究教育目的（非営利）に限り、著者が第三者に直接配布することができる。著者以外からの配布は禁じられている。

情報・システムソサイエティ

一般社団法人 **電子情報通信学会**

THE INFORMATION AND SYSTEMS SOCIETY

THE INSTITUTE OF ELECTRONICS, INFORMATION AND COMMUNICATION ENGINEERS

手技訓練補助のための正常異常状態対で構成される状態遷移モデルによる異常動作の検出

小川 陽子^{†a)} 松尾 直志^{††} 島田 伸敬^{††} 白井 良明^{††}
来見 良誠^{†††} 小森 優^{††††}

Anomaly Detection of Knot-Tying Action in Surgical Training
by Probabilistic Frame Anomalouslyness Estimation

Yoko OGAWA^{†a)}, Tadashi MATSUO^{††}, Nobutaka SHIMADA^{††}, Yoshiaki SHIRAI^{††},
Yoshimasa KURUMI^{†††}, and Masaru KOMORI^{††††}

あらまし 手術手技の訓練は、まず手形状を覚えることから始まる。本論文では、手術手技の自己訓練のための手形状を改善箇所を検出手法を提案する。提案手法では、RGB-D センサから得られた手技動作の手形状が要改善か否かを、点群位置合わせと時系列フィルタリングを用いて確率的に推定する。動作モデルを正常異常状態対から構成される状態遷移モデルとして構築し、手本動作の1フレームに正常と異常の2状態を割り当てることで正しいフレームマッチングと要改善箇所の検出を同時に行う。また、被訓練者の習熟度に合わせて手本手形状との違いが大きい箇所を優先的に指摘するために、入力シーケンスが得られるたび適応的に状態遷移モデルのパラメータを推定し直す。我々は、1名の熟練医及び1名の初心者からの結紮動作から状態遷移モデルを構築し、状態遷移モデルの構築に使用していない2名の初心者の結紮動作に対し入力動作内の要改善箇所の指摘における本手法の有効性を確認した。

キーワード 異常検出, 動作認識, 状態遷移モデル, 手術手技訓練支援

1. ま え が き

外科手術を行う医師は、患者への負担を軽減するため高度な手技技術が要求される。通常、外科訓練医は熟練医師から直接手技を習うが、近年の医師不足により熟練医は自身の職務に忙しく、教育の機会が限られており、訓練医の自学自習の効率化が求められている。そのための一つのソリューションとして、訓練医の手

術手技を自動的に評価しフィードバックするシステムが挙げられる。本論文では、手始めに結紮手技をターゲットとし、訓練医が行う結紮動作を熟練医による手本動作と比較することで要改善箇所を指摘するシステムを提案する。

結紮訓練は一般にボックストレーナーや机に貼った糸を用いて行われる。また、結紮の訓練は表1に示す4段階で構成される。初心者は、まず手本となる熟練医の動作を見て真似をし、自身の動作のうち手本の動作と異なる部分(要改善箇所)を探し修正していくことで上達する。そのため、要改善箇所を自動的に検出し提示することは訓練の効率化に有用である。Oshima

表1 結紮訓練の構成

Table 1 Construction of surgical knot-tying training.

段階	訓練内容	評価基準
1	形を覚える	手の形
2	緩まないように縛る	糸のテンション
3	体組織が千切れないように結ぶ	体組織にかかる力
4	スムーズに結ぶ	結ぶ速度

[†] 立命館大学大学院情報理工学研究科, 草津市

Graduate School of Information Science and Engineering,
Ritsumeikan University, Kusatsu-shi, 525-8577 Japan

^{††} 立命館大学情報理工学部, 草津市

College of Information Science and Engineering, Ritsumeikan
University, Kusatsu-shi, 525-8577 Japan

^{†††} 独立行政法人 地域医療機能推進機構 滋賀病院, 大津市

Japan Community Health care Organization Shiga Hospi-
tal, Otsu-shi, 520-0846 Japan

^{††††} 滋賀医科大学, 大津市

Shiga University of Medical Science, Otsu-shi, 520-2192
Japan

a) E-mail: ogawa@i.ci.ritsume.ac.jp

DOI:10.14923/transinfj.2017PDP0033

らは縫合と結紮の総合的な訓練システムを提案している [1], [2]. Oshima らの方法では, 縫合に要した時間, 皮膚に加わった力, 結紮力, 縫合間隔, 縫合幅 (左右の均等さ), 創離開 (傷口面積) の 6 項目を評価しているが, 一番最初の段階である手の形については評価できない. 本研究では訓練の初期段階を対象に, 手の形状の要改善箇所を検出, 指摘することを目的とする. 図 1 に右手結びにおける要改善箇所と改善不要箇所の例を示す. 初学者の動作では手の形が手本動作と大きく異なる部分が多いことを想定し, 動作全体を見て特に手本と異なる区間を検出する方法を提案する. また, ユーザの動作と手本動作をフレームごとに対応付け, ユーザが目視で解釈しやすいように指摘箇所に対応する手本動作をユーザの動作と並べて提示し動作修正を促す.

改善不要な箇所を正常動作, 要改善箇所を異常動作と捉えると, 要改善箇所の検出は異常動作の検出といえる. 異常動作の検出は大きく 2 通りの方法が提案されており, 一つは異常動作自体の動作モデルを作成する方法, もう一つは正常動作の動作モデルから異常動作を定義する方法である.

Wang らは人のシルエットの特徴を用いていくつかの動作の HMM を構築し, そのうち一つの動作を正常動作とし, 正常動作以外に相当する動作を検出した場合にアラートを上げるシステムを提案している [3]. Foroughi らは, 動画のある一定区間からシルエットや頭の位置の動きを特徴量として, 10 クラスの動作を SVM で識別し, 転倒を検知する方法を提案している [4]. これら異常動作自体の動作モデルを作成する方法は, 典型的な異常動作クラスの知識が利用できる場

合には適用できるが, ある一つのまとまった正常な動作の中に発生する部分的な未知の動作の検出には適用できない.

異常動作クラスの知識が必要ない方法として, 正常な動作のゆう度がしきい値を下回った場合を異常動作とする方法 [5] や, 異常動作を一様分布若しくは正常動作の逆分布としてモデル化し, そのゆう度比から異常動作を検出する方法 [6] がある. これらは, 十分な学習のために学習サンプルが大量に必要なため, 手本となるサンプルが集めにくい場合には適さない. また, 熟練度に合わせた動的な要修正箇所の指摘は困難である.

そこで我々は, 少ない学習サンプルから構築できて熟練度に合わせた要修正箇所検出も可能な動作モデルとして, 正常と異常の状態対から構成される状態遷移モデルを提案する. 結紮動作には片手結びや両手結びなど幾つかのタイプがあるが, そのタイプごとに代表的な手本動作を 1 本用意し, ユーザによる入力動作の各時刻に対応する手本動作の時刻と, それが要改善か否かを同時に推定する. 更に, 入力動作が得られるたび, 状態遷移モデルのパラメータを推定し直すことで, 入力動作の中で特に手本動作の手形状との違いが大きい箇所を指摘する.

本論文では, 2. で正常状態のみからなる状態遷移モデルを用いた異常動作検出方法をベースラインとして述べ, 3. で本提案の中心となる正常異常状態対で構成される状態遷移モデルを用いた方法を示す. 4. では, 入力シーケンスに合わせた状態遷移モデルのパラメータの適応的推定について述べ, 5. で実験結果を示す.

2. 入力動作のフレームごとの確率的状態推定

本研究では手術手技訓練を対象としているため, 入力される結紮動作のタイプは既知とする. 本節では, ベースラインとして図 2 のような正常状態のみからなる状態遷移モデルを用いた異常動作検出について述べる.

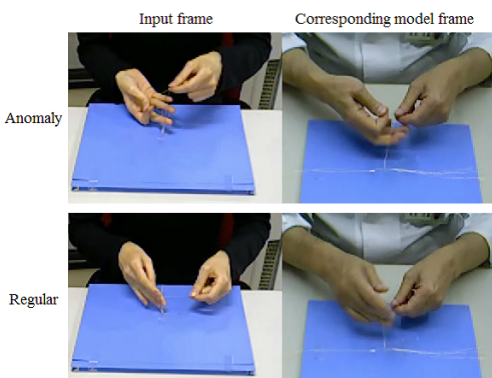


図 1 異常手形状と正常手形状の例

Fig. 1 Examples of anomaly and regular hand shapes.

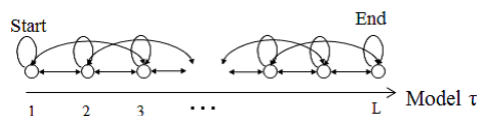


図 2 正常状態のみからなる状態遷移モデル

Fig. 2 State transition model consisting of regular states.

2.1 観 測

本研究では、結紮動作中の手の三次元的な形状を評価するために、RGB-D センサを使用し、手本動作及び入力動作の開始から終了までを一定のフレームレートでキャプチャする。以下では、RGB-D データをキャプチャした各時刻をフレームとし、連続したフレームからなる系列をシーケンスと呼称する。更に、RGB-D センサから得られた三次元点群から手に由来する点群のみを抽出して使用する。RGB-D センサと結紮対象、及び被訓練者の位置関係はおおまかに限定できるものとし、肌色領域の芯線に相当する点から深度に基づいて領域を拡張することで手領域を抽出する。また、幾何的配置と時系列追跡に基づいて左右の手領域に分割する。

2.2 手戻りを許した状態遷移モデルによる対応フレーム推定

結紮動作のタイプごとに手本シーケンス 1 本が得られるものとして、図 2 のように、 L フレームからなる手本シーケンスの各フレームに対し 1 状態が割り当てられた状態遷移モデルを構築し、入力シーケンスのフレーム t が対応する手本シーケンスのフレーム τ_t を推定し、 τ_t を当該手技の時刻 t における進捗度（フェーズ）とみなす。状態間の遷移は手本動作に対する入力動作の速度比 $\dot{\tau}$ に相当する。以下では便宜上 $\dot{\tau}$ を動作速度と呼称する。動作認識の分野では left-to-right の HMM が使用されることが多いが、本研究では対象を初心者の方の手技としているため、必ずしも正しい順に動作が入力されるとは限らず手戻りや停滞が発生し得る。また、熟練者の手技に対して被訓練者の手技は基本的に遅いが局所的には最大 2 倍まで早くなり得ることを想定し、前後 2 状態までの遷移が可能な状態遷移モデルを用いる。すなわち $-2 \leq \dot{\tau} \leq 2$ とする。

ある時系列の入力シーケンス $O_{1:T} = \{o_t\}_{t=1}^T$ が得られたとき、入力シーケンスの各フレームに相当するフェーズを推定するには各時刻の事後確率分布 $P(S_t | O_{1:T})$ を推定すればよい。ただし、 $S_t = (\tau_t, \dot{\tau}_t)$ であり、 $\dot{\tau}_t$ は入力シーケンスの t フレーム目における動作速度を表す。 S_t の一次マルコフ性を仮定し、 $P(S_t | O_{1:T})$ をベイズの定理に基づいて展開すると、

$$\begin{aligned} P(S_t | O_{1:T}) &= \sum_{S_{t+1}} P(S_t, S_{t+1} | O_{1:T}) \\ &= \sum_{S_{t+1}} \frac{P(S_{t+1} | S_t) P(S_t | O_{1:t})}{P(S_{t+1} | O_{1:t})} P(S_{t+1} | O_{1:T}) \end{aligned} \quad (1)$$

となり、 $P(S_t | O_{1:t})$ は更に

$$\begin{aligned} P(S_t | O_{1:t}) &= \sum_{S_{t-1}} P(S_t, S_{t-1} | O_{1:t}) \\ &= \sum_{S_{t-1}} \frac{P(o_t | S_t) P(S_t | S_{t-1})}{P(o_t | S_{t-1})} P(S_{t-1} | O_{1:t-1}) \end{aligned} \quad (2)$$

と展開できる。ゆー度 $P(o_t | S_t)$ と遷移確率 $P(S_t | S_{t-1})$ 、及び 1 フレーム目の事前分布 $P(S_1 | o_1)$ が決まれば式 (2) から式 (1) の順に逐次的に解くことができる [7]。

2.3 手本手形状の点群に対する入力点群のゆー度

1 状態が 1 点群に相当する状態遷移モデルを用いるので、各状態のゆー度 $P(o_t | S_t)$ は手領域点群を関節角度や特徴記述子のパラメータ空間上の分布として記述することがまず考えられる。画像から手指の関節角度を推定する研究が近年盛んにされているが [8]～[10]、強力な自己隠ぺいを有するフレームなどのイレギュラーな入力に対して大きく間違える場合がある。それに対し、見えそのものを何らかの特徴量で記述する方法も多く提案されているが [11], [12]、データの欠損や隠ぺい、視点変化に対する頑健性が低い。また、各状態が関節角度や特徴ベクトルの多次元空間上の分布をもつと仮定すると、推定するパラメータが状態の数だけ多くなり、必然的に学習用シーケンスも多く必要となる。

本研究では、ICP (Iterative Closest Point) アルゴリズム [13] による手領域点群の位置合わせを行うことで、自己隠ぺいや視点変化による影響を吸収する。ICP アルゴリズムでは、入力点群の各点から最も近い手本手形状の点群の 1 点に対応付け、その対応点間距離の 2 乗平均（ICP 残差）が最も小さくなるように入力点群を剛体変換する操作を繰り返す。各状態のもつ ICP 残差の分布を個別にモデル化するにはそれぞれの手本手形状に対応する手形状のサンプルが多数必要であるが、本研究では学習サンプルが多く得られない場合を想定しているためこれが困難である。しかし、手本手形状がどのような形状であっても入力手形状が正常（手本手形状と同様の姿勢）である場合、ICP 残差は狭い範囲に分布する。更に、ICP 残差が 0 のときに最もゆー度が高く、残差が大きくなるにつれてゆー度は下がるべきであることは、どのような手本手形状においても共通している。これらから、状態を周辺化し

た ICP 残差の分布を一つのパラメトリックな分布としてモデル化し、各状態のゆう度分布として使用する。

ICP 残差のゆう度は、ICP 残差が 0 のときに最も高く、ICP 残差が大きくなるにつれ単調減少するような関数であることが望ましい。手本手形状の点群 q に入力点群 o を位置合わせしたときの K 個の対応点ペアの三次元変位ベクトルの集合を $C_{1:K} = \{\mathbf{c}_k\}_{k=1}^K$ とすると、ICP 残差 d^2 は $d^2 = \frac{1}{K} \sum_{k=1}^K \mathbf{c}_k^T \mathbf{c}_k$ であり、 $C_{1:K}$ の各成分の分散の和になる。ここで、 $C_{1:K}$ の各成分の標準偏差 (s_x, s_y, s_z) を観測とみなし、このゆう度を 0 を中心としたガウス関数を用いて、

$$L(d^2) = \frac{8}{(2\pi\sigma^2)^{\frac{3}{2}}} \exp\left(-\frac{d^2}{2\sigma^2}\right) \quad (3)$$

と定義する (ゆう度モデル)。

以下では入力シーケンスのあるフレームと手本シーケンスのあるフレームのペアをフレームペアと呼ぶ。学習用入力シーケンスの各フレームとその対応する手本シーケンスのフレームのペア (正解フレームペア) の ICP 残差のうち、入力シーケンスのフレームが目視で正常動作とされたものを学習サンプルとし次式で求める。

$$\sigma^2 = \frac{1}{3N} \sum_{i=1}^N d_i^2. \quad (4)$$

N はサンプル数、 d_i^2 は i 番目のサンプルである。更に、二つの点群の変位ベクトル集合は直前の動作速度に依存せず、二つの点群形状のみによって決まるとする

$$\begin{aligned} P(o_t | S_t = (\tau_t = j, \dot{\tau}_t)) &= P(o | \tau_t = j) \\ &= P(o | q_j) := L(d^2) \end{aligned} \quad (5)$$

となる。 q_j は j フレーム目の手本手形状の点群である。これを式 (2) に代入する。

また、基本的な ICP では、入力点群の初期姿勢によって収束する姿勢が異なる場合があり、適切な初期姿勢を与える必要がある。本研究では、初期姿勢として両点群の重心を合わせた姿勢、直前のフレームの推定結果と手本動作から予測した姿勢、及び直前のフレームの推定結果と入力動作から予測した姿勢を用い、回転量と並進量がしきい値未満かつ最も残差が小さくなった変換を採用する。また、姿勢変換により点群の隠れ目が起こり得るので、点群法線とカメラ間の姿勢

から、双方のカメラで観測可能な点のみを用いて位置合わせを行う。

2.4 遷移確率

遷移確率においても学習サンプルが多く得られない場合を想定しており、遷移ごとに遷移確率をモデル化することが困難であるため、状態で周辺化した確率分布で近似する。また、結紮手技の動作は基本的に前へ進み、急激な動作速度の変化は起こりにくいで、動作速度を定常な一次のマルコフ過程と捉え、直前の動作速度に対する現在の動作速度の確率分布をモデル化する。すなわち、 $\dot{\tau}_{t-1} = \{-2, -1, 0, 1, 2\}$ に対する $\dot{\tau}_t = \{-2, -1, 0, 1, 2\}$ の全ての組合せの確率 $P(\dot{\tau}_t | \dot{\tau}_{t-1})$ を学習サンプルから求め、各遷移確率を近似する。

2.5 対応フレームの決定

入力シーケンスの各フレームの対応するフェーズを一意に決定することができれば、被訓練者と熟練者の手形状と動作速度の比較が可能となり、効果的な訓練のために有用である。

各状態のゆう度分布及び各遷移確率が十分な数のサンプルから学習できるならば、Viterbi アルゴリズムを用いて状態と遷移の同時確率が最も高くなる組合せから各フレームの状態を求める。しかし、今回のように学習サンプルが十分に得られない場合には誤った対応付けがされ得る。そのため、 $P(S_t | O_{1:T})$ を τ について周辺化した $P(\tau_t | O_{1:T})$ と速度制限を用いて各フレームの状態を決定する。時刻 t が状態 τ_t であるとしたとき、 $\{\tau_t\}_{t=1}^T$ の同時確率最大となる組合せは Viterbi アルゴリズムにより以下の式で求める。

$$g(1, \tau_1) = -\log P(\tau_1 | O_{1:T}) \quad (6)$$

$$\begin{aligned} g(t, \tau_t) &= \min_{\tau_t} \{ -\log P(\tau_t | O_{1:T}) \\ &\quad + g(t-1, \tau_{t-1} = \tau_t - \dot{\tau}_t) \} \quad (7) \\ -2 &\leq \dot{\tau}_t \leq 2 \end{aligned}$$

$g(t, \tau_t)$ は、時刻 1 から t までの最大同時確率の負の対数ゆう度である。全ての $g(t, \tau)$ を求め、 $g(T, L)$ から逆に辿ることで、同時確率最大となる各時刻の対応する状態を決定する。

2.6 ゆう度に基づく異常フレームの検出

対応する状態を決定したのち、そのゆう度がしきい値以下となるフレームを異常フレームとして検出すると、図 3 のようになる。図 3 左は異常を含まないシー

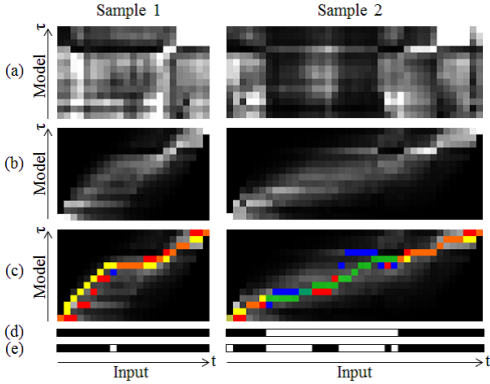


図 3 結果例：(a) ゆう度, (b) 事後確率分布, (c) 結果と真値, (d) 異常動作区間の真値, (e) 検出された異常動作区間. (a) (c) では, 白いセルほど確率が高いことを意味する. (c) 中の赤と青のセルは計算により求められた正常区間と異常区間, 黄と緑のセルは正常区間と異常区間の真値である. (d) と (e) の白い箇所は異常区間を表す. Sample 1 と Sample 2 はそれぞれ異常を含まないシーケンスと異常を含むシーケンスである.

Fig. 3 Baseline result: (a) likelihood, (b) posterior probability distribution, (c) results and ground truth, (d) ground truth of anomaly section, (e) detected anomaly section. In (a) (c), Whiter cells mean they have higher probability. In (c), red and blue cells are calculated results of regular and anomaly. Yellow and green cells are ground truth of regular and anomaly. White lines in (d) and (e) mean anomaly sections. Sample 1 consists of only a regular section, and Sample 2 contains an anomaly section.

ケンスを入力とした場合, 図 3 右は異常を含むシーケンスを入力とした場合である. しきい値は, 学習用のシーケンスに対し, 0.1 から 0.5 まで 0.1 刻みに 5 通り試行して F 値が最大となるしきい値 (0.4) を用いた. (a), (b) はそれぞれゆう度, 事後確率分布であり, 縦軸が手本シーケンスのフレーム, 横軸が入力シーケンスのフレームである. 白いほど高いことを表す. (c) は (b) 上に対応付け結果と真値を描画したものである. 赤が正常状態に対応付けられたフレーム, 青が異常状態に対応付けられたフレーム, 黄色が正常フレームの真値, 緑が異常フレームの真値である. (d), (e) はそれぞれ正常動作区間と異常動作区間の真値と検出結果を示したものであり, 黒が正常動作区間, 白が異常動作区間である. 図 3 右の 14, 19 フレーム目の照合結果を図 4 に示す. 異常と判断されたフレームには入力フレームに赤枠と異常を示すアイコンを入力フレーム下部に描画している. 14 フレーム目は異常と判断す

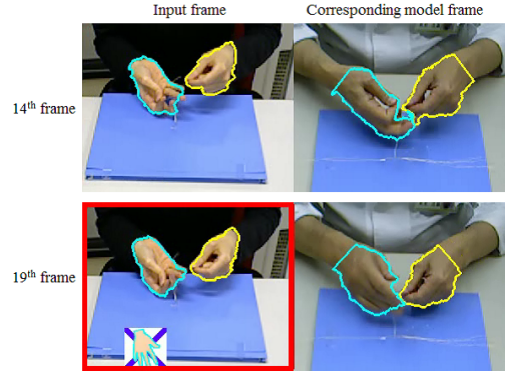


図 4 ユーザに提示される結果画像例：図 3 の sample 2 の 14 番目と 19 番目のフレーム.

Fig. 4 Shown result images for user: 14th and 19th frames of sample 2 of Fig. 3.

べき区間内で正常と判断されたフレームであるが, 手本と入力の手を見比べると大きく異なるポーズをしているにもかかわらず正常であると判断されている. 19 フレーム目はフレーム対応付けを大きく間違えたフレームである. 入力に結び目を作る途中であるが手本は糸を引っ張る動作へ移行する段階である. このように, 対応状態を先に決定してから異常か正常かを判別すると, 正常か異常かの判断のみならずフレーム対応付け自体も間違いやすくなる.

3. 異常動作検出のための結紮動作モデル

2. で述べた方法では, 対応する状態を決定してからそのゆう度のみに基づいて正常であるか異常であるかを判別しているため, 1 フレームごとに正常と異常が切り替わる箇所が散見されたが, 異常フレームと正常フレームはどちらも急に発生するものではなく前後のフレームに対して独立でない. そのため, 本質的には入力の各フレームの対応するフェーズと正常か異常かは同時に推定されるべきである. また, ユーザの視点に立つと, 1 フレームごとに異常であるか否かを指摘されるより, まとまった区間で指摘される方が改善の方向性を解釈しやすい.

そこで, 我々は異常状態を陽に組み込んだ状態遷移モデルを用いて, 正常状態と異常状態間の遷移, 異常状態のゆう度を定義することで前後のフレームが異常か否かを考慮した異常動作の検出方法を提案する. 本節では, 2. の内容を元に改善箇所について述べる.

3.1 異常状態を陽に組み込んだ状態遷移モデル

2.2 では, 手本シーケンスの各フレームに一つの状

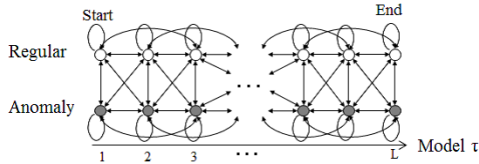


図 5 正常異常状態対からなる状態遷移モデル

Fig. 5 State transition model with regular and anomaly states.

態を割り当て、入力シーケンスと手本シーケンスのフレーム間マッチングを行った。ここで、手本シーケンスの各フレームにおける正常手形状と異常手形状をそれぞれそのフレームにおける排他的な状態として捉え、図 5 のように手本シーケンスの各フレームに相当する正常状態と異常状態を対として並べ接続することで、あり得る全ての状態をモデル化する。状態数は $M = 2L$ となる。以下では、便宜上正常状態と異常状態を交互に番号付け、そのインデックスを τ で表す。入力シーケンスの各フレームの対応する状態を 2.2 で述べた方法で推定する。

3.2 異常状態に対する入力点群のゆう度

正常状態のゆう度は 2.3 と同様に式 (3) を用いる。ある手本手形状に対して異常といえる手形状は無数にあるため、異常状態のゆう度モデルとしてどんなものが適切であるかを一意に決めることが困難である。しかし、残差 0 付近では異常状態のゆう度は正常状態に比べて低くなるべきであることは、どのような手本手形状においても共通している。ここで、分布パラメータ σ^2 が正常状態よりも大きいガウス関数であれば上記を満たすことができるため、正常状態と同様に、異常状態で周辺化した ICP 残差の分布を式 (3) のガウス関数で表す。

異常状態の ICP 残差の分布パラメータ σ^2 は、入力手形状が異常と目視で判断した正解フレームペアから最ゆう推定する。

3.3 正常異常間を含む遷移確率

異常状態を陽に組み込んだ状態遷移モデルの場合、正常状態から異常状態、異常状態から正常状態の遷移を含めてモデル化する必要がある。目視でフレーム対応付けと正常か異常かのラベル付けを行った学習用のサンプルから、2.4 と同様に $\hat{\tau}_{t-1}$ に対する $\hat{\tau}_t$ の全ての組合せの確率 $P(\hat{\tau}_t | \hat{\tau}_{t-1})$ 推定する。

3.4 対応フレームの決定と異常区間の検出

2.2 と同様に各入力フレームの対応する状態を事後

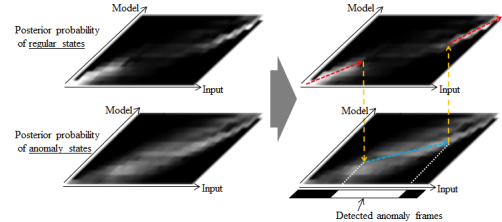


図 6 フレーム間マッチングと異常フレームらしさの同時推定の概要

Fig. 6 Outline of simultaneous frame-wise matching and anomaly detection.

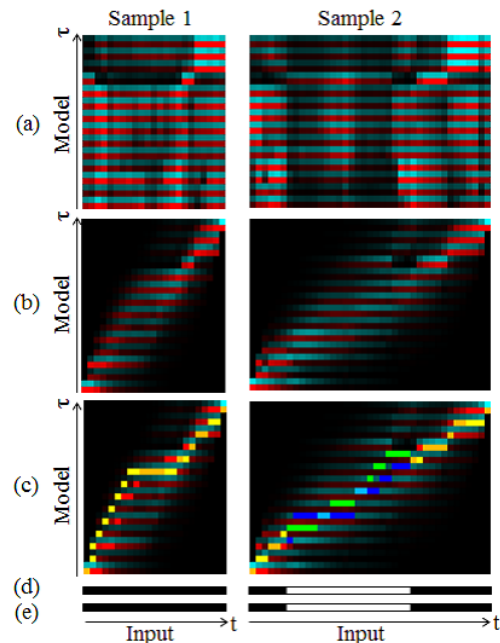


図 7 正常異常状態対からなる状態遷移モデルによる異常検知の結果

Fig. 7 Result of anomaly detection with regular and anomaly state transition model.

分布として推定し、得られた S_t の事後分布を速度 $\hat{\tau}_t$ について周辺化すると、図 6 左のようになる。上が正常状態、下が異常状態の事後分布である。入力シーケンスの序盤と終盤は正常状態の確率が高く、中盤では異常状態の確率が高くなっているが、その変化は滑らかになる。2.5 の手法で対応フレームを決定すると、図 3 右のように頻繁に正常と異常が切り替わるようなことは見られず、図 6 右に示すように動作の中盤において連続的な異常区間として検出される。図 3 と同じ入力シーケンスに対する推定結果を図 7 に示す。(a)、(b) はそれぞれゆう度、事後確率分布であり、縦軸が

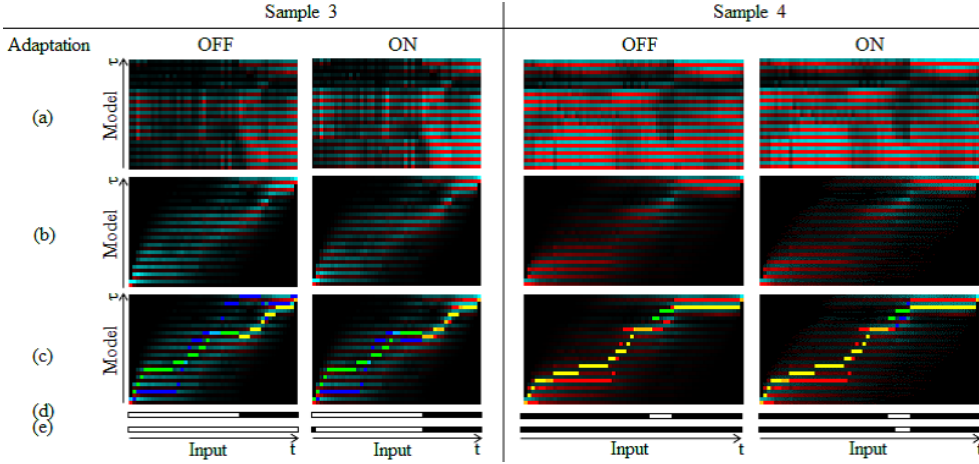


図 8 パラメータの適応的推定の結果
Fig. 8 Result of adapting parameters.

状態遷移モデルの状態、横軸が入力フレームである。正常状態には赤、異常状態にはシアンで色付けしている。明るいほどゆう度または事後確率が高いことを表す。(c) は (b) 上に対応付け結果と真値を描画したものである。図 3 と同様に、赤が正常状態に対応付けられたフレーム、青が異常状態に対応付けられたフレーム、黄色が正常フレームの真値、緑が異常フレームの真値である。(d), (e) も図 3 と同様に真値と検出結果を見やすくしたものである。図 3 に比べ、対応付いたフレームも正常か異常かの判定も真値に近づいていることが見られる。

4. 適応的ゆう度分布の推定

学習データの演技者と被訓練者の間では手技の熟練度は一般に異なる。そこで、入力シーケンスが得られるごとに適切な分布パラメータを再推定することで被訓練者に適応させる。具体的には正常状態のゆう度分布の分散 σ^2 と異常状態のゆう度分布の分散 $\bar{\sigma}^2$ を再推定する。熟練度が低いときは正常とする範囲を広げ、熟練度が上がるにつれて正常の範囲を狭め、訓練を繰り返すごとに指摘される箇所が変わっていくことが理想である。事後分布 $\gamma(\tau_t) = P(\tau_t | O_{1:T})$ と各フレームペアの ICP 残差 $d_{t,\tau}^2$ 、対応点数 $k_{t,\tau}$ が得られたとき、EM アルゴリズムに基づき

$$\sigma_{new}^2 = \frac{3\lambda\sigma^2 + \sum_{t=1}^T \sum_{i=1}^L \gamma(\tau_t = 2i-1) d_{t,\tau_t}^2}{3\left(\lambda + \sum_{t=1}^T \sum_{i=1}^L \gamma(\tau_t = 2i-1)\right)} \quad (8)$$

$$\bar{\sigma}_{new}^2 = \frac{\lambda\bar{\sigma}^2 + \sum_{t=1}^T \sum_{i=1}^L \gamma(\tau_t = 2i) d_{t,\tau_t}^2}{3\left(\lambda + \sum_{t=1}^T \sum_{i=1}^L \gamma(\tau_t = 2i)\right)} \quad (9)$$

として適応後の σ_{new}^2 と $\bar{\sigma}_{new}^2$ を求め、事後分布の推定と分散再推定の操作を収束するまで繰り返すことで入力シーケンスに応じたゆう度分布のパラメータを得る。このとき、過適応を防ぐため初期の分散 σ^2 , $\bar{\sigma}^2$ にフレームペア数に相当する重み λ をかけて混合している。

ゆう度分布の再推定ののち状態の対応付けを行った結果の 2 例を図 8 に示す。それぞれ左が適応的ゆう度分布の推定なし、右が適応的ゆう度分布の推定有り、(a)~(e) は図 7 と同様である。適応的ゆう度分布の推定なしでは全てのフレームが要改善として検出されてしまっている (図 8 左の (e)) が、適応的ゆう度分布の推定有りではフレーム対応付け精度が向上し、手技の前半のみが要改善として指摘され、真値とほぼ一致している結果を得ている。図 8 右は、学習時のサンプルより平均的な残差が小さい (熟練度が上がった) 被訓練者のシーケンスである。適応的ゆう度分布の推定なしでは検出漏れしていた区間が適応的ゆう度分布の推定によって検出されるようになっている。

5. 実験

2. で述べた正常状態のみからなる状態遷移モデルを用いた場合をベースラインとして、3. で述べた正常異常状態対からなる状態遷移モデルを用いた場合を提案手法 1、提案手法 1 に 4. の適応的ゆう度分布の推定を

表 2 検出区間評価: (i) 正常状態のみ, (ii) 提案手法 1, (iii) 提案手法 2

Table 2 Validation of results: (i) is only regular states, (ii) is proposed method 1, (iii) is proposed method 2.

	TN	FP	FN	TP	正確度	再現率	精度	F 値
(i)	1891	380	62	705	0.86	0.92	0.65	0.76
(ii)	1889	313	34	802	0.89	0.96	0.72	0.82
(iii)	2029	215	46	748	0.91	0.94	0.78	0.85

適用した手法を提案手法 2 として比較する。

手術用の縫合糸を台紙にセロハンテープで留め、医師 1 名の手本動作、初心者 3 名（うち 1 名はある程度練習済み）の入力動作を 20fps 程度のフレームレートで撮影した。初心者 3 名のうち練習済みの 1 名のデータをゆう度分布と遷移確率のパラメータ学習に用い、残り 2 名のデータを評価用に用いた。評価用データはシステムによる指摘前後のものを含む。RGB-D センサには ASUS Xtion Pro Live を用いた。結び動作は左手結びと右手結びの 2 通りかつ既知とし、左手結びならば左手、右手結びならば右手を検証対象とした。手本動作には熟練医の手技を、学習サンプルにはある程度慣れた被訓練者の手技を用いる。教師ラベルは全てのデータに対し筆頭著者が日を開けて複数回目視評価を行い、可能な限り身体や精神の状態による影響を除外するよう配慮した。一部の目視評価結果は熟練医の確認を取った。

異常と正常の区間は連続的に遷移するものであり、区間境界の多少のずれはペナルティを低くしたい。全体的に長いシーケンスの場合、正常と異常の区間の間で中間的なフレームが多く観測されるので、シーケンス長に応じて検出位置のずれに許容範囲を設ける。今回は、シーケンス長の 1 割を区間境界のずれの許容範囲とする。また、異常区間であるべき範囲内の正常区間は区間境界にかかわらず全て検出漏れとし、正常区間であるべき範囲内の異常区間は全て誤検出として評価する。

状態遷移モデルが正常状態のみの場合、提案手法 1、及び提案手法 2 の検出区間評価結果の比較を表 2 に示す。真陰性と偽陽性の和、及び偽陰性と真陽性の和に違いが生じているが、これは前述のとおり境界のずれを許容したためである。学習データは 13 シーケンス（421 フレーム）、評価用データは 72 シーケンス（3038 フレーム）を用いた。評価用データのうち、目視で要改善箇所とされたフレームを含むシーケンスは 40 本であった。表 2 から、正確度、再現率、精度、F

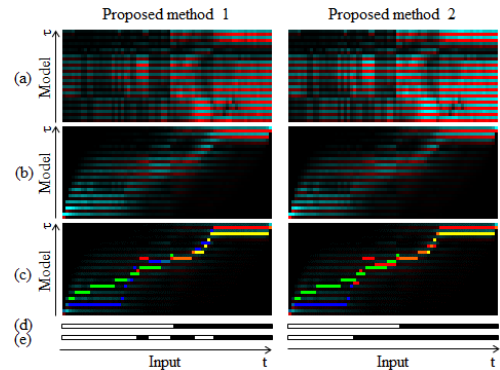


図 9 シーケンスの半分以上が目視で要改善箇所とされた例

Fig. 9 Result of an input sequence that is consisted mostly of anomaly motion.

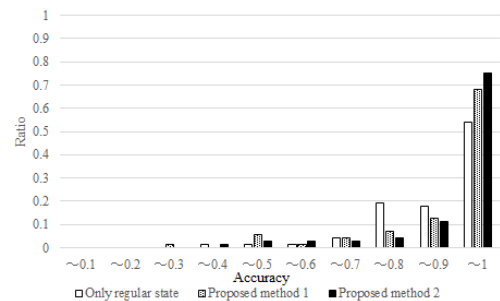


図 10 正確度の頻度分布

Fig. 10 Histogram of accuracy.

値ともに提案手法が高くなっている。

真値として用いた目視評価結果は、手本と手形状が合っているか否かの絶対的な評価結果となっている。それに対し、提案手法 2 では、シーケンス全体を見て正常状態及び異常状態の分布パラメータを適応的に推定することで、特に手本との差が大きい箇所を選択的に指摘する相対評価となっている。そのため、シーケンスの半分以上が目視評価で要改善箇所とされた例において提案手法 2 では選択的な指摘となり、見かけ上再現率が低下する。具体的な例を図 9 に示す。また、これらのシーケンスごとの正確度の頻度分布を図 10 に示す。ベースラインに比べて提案手法は正確度が 0.9 を超える割合が高くなっている。また、0.5 近辺の割合がやや上昇するが、これは目視評価時に手形状が似ており改善不要とされたフレーム（図 14 参照）において点群形状の違いが大きく要改善箇所として検出された例が含まれるためである。目視による要改善箇所を含むシーケンスのみでは図 11 に示す分布となってお

り、こちらもベースラインに比べて 0.9 を超える割合が高くなっている。対応フレーム誤差の平均は正常状態のみの場合に 1.4 フレーム、提案手法 1 で 1.3, 提案手法 2 で 1.2 となり、正常状態のみの場合と比較して提案手法 2 で 0.2 フレームほどの向上が見られた。また、シーケンスごとの平均対応フレーム誤差の頻度分布を図 12 に、目視による要改善箇所を含むシーケンスのみでは図 13 に示す分布となった。これらの結果から、目視による要改善箇所を含む場合により提案手法が有効であったといえる。

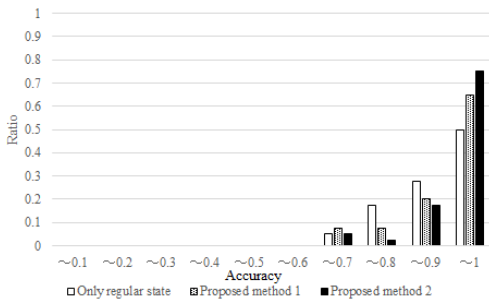


図 11 正確度の頻度分布（目視による要改善箇所を含むシーケンスのみ）

Fig. 11 Histogram of accuracy with anomaly sequence.

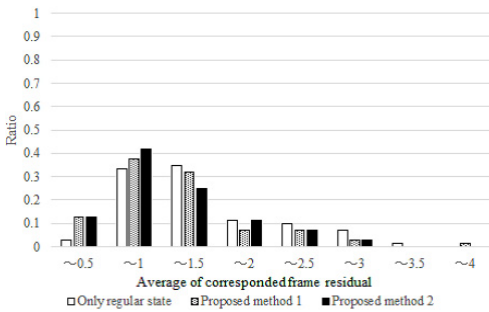


図 12 平均対応フレーム誤差の頻度分布

Fig. 12 Histogram of average corresponded frame residual.

システムが要改善箇所を指摘したのち、被訓練者は指摘された箇所を修正した動作を再びシステムに入力することで、次の要修正箇所があれば更に指摘されることが理想である。図 8 の Sample 3 は訓練初期、Sample 4 はシステムによる要改善箇所指摘後の動作であるが、システムによる指摘によって動作前半が修正され、後半の一部が指摘されるように変化していることが見て取れる。図 15 に指摘前後で同じ手本フレームに対応づいた入力フレームを示す。これから、指摘された箇所の手形状が手本に近づいていることが分かる。

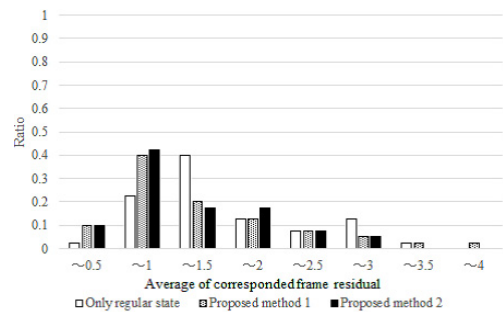


図 13 平均対応フレーム誤差の頻度分布（目視による要改善箇所を含むシーケンスのみ）

Fig. 13 Histogram of average corresponded frame residual with anomaly sequence.



図 14 目視評価時に改善不要とされたが要改善箇所として検出されたフレーム

Fig. 14 Example of confused frame.

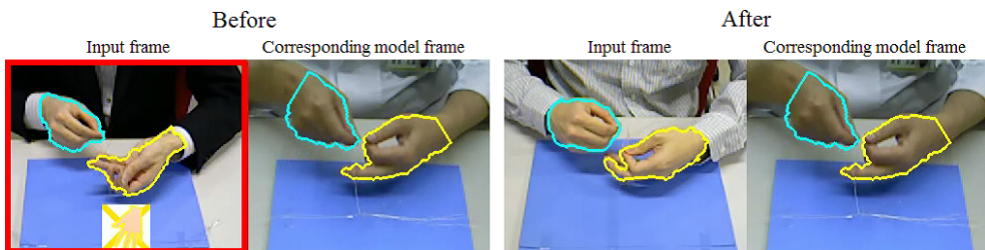


図 15 要改善箇所指摘前後の比較

Fig. 15 Comparison of before and after pointing out anomaly sub-motion.

6. む す び

本論文では、正常状態と異常状態の2種の状態で構成される状態遷移モデルとして動作モデルを構築し、手本と入力シーケンス間のフレームマッチングと同時にフレームの異常性を推定する方法を提案した。被訓練者の結び動作における手の形状が正常か異常かを点群位置合わせの残差とフィルタリング理論に基づいて確率的に推定した。また、入力シーケンスが得られるたび適応的にゆう度分布を推定し直すことで、被訓練者の習熟度に合わせて手本動作との違いが大きい箇所を優先的に指摘した。実験では、1名の熟練医及び1名の初心者の結紮動作から状態遷移モデルのパラメータを学習し、学習に使用していない2名の初心者の結紮動作に対し入力動作内の要改善箇所の指摘における本手法の有効性を確認した。本手法は、モデルに用いた手本を被訓練者に知らせ、被訓練者がそれを真似して手技を行う限りにおいて適用が可能である。モデルに用いた手本とは異なる流派で事前に訓練を行っている手形状の見えがシーケンス全体で全く異なる場合があるが、その場合、テンプレートとなる手本シーケンスを差し替えることで適用可能である。また、今回は左手結びと右手結びの2通りのみを用いたが、結び方には幾つか種類がある。他の結び方を訓練する際にも手本シーケンスの差し替えのみで本手法が適用可能である。それぞれの手本シーケンスに対する初心者の学習データが用意できれば、ゆう度及び遷移確率がより確からしく推定できる。

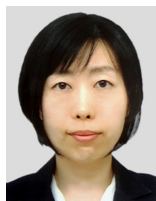
今後は、手技の速度の評価やユーザへの評価結果の提示方法、訓練の記録などユーザビリティの改善などが考えられる。また、ICPを用いるため処理が重く実時間でのフィードバックは実現できていないため、実用化のためにはGPUや分散処理などを用いたICPの高速化が必要である。

謝辞 本研究はJSPS科研費 24500224, 15H02764及び文部科学省私立大学戦略的研究基盤形成支援事業S1311039 (2013-2016)の助成を受けたものである。

文 献

- [1] N. Oshima, J. Solis, Y. Ogura, K. Hatake, and A. Takanishi, "Development of the suture/ligature training system WKS-2 designed to provide more detailed information of the task performance," *Intelligent Robots and Systems, IEEE/RSJ Int. Conf. on. IEEE*, pp.58-63, 2007.
- [2] N. Oshima, J. Solis, H. Ishii, N. Matsuoka, K. Hatake, and A. Takanishi, "Integration of an evaluation function into the suture/ligature training system WKS-2R," *Robotics and Automation, IEEE Int. Conf. on. IEEE*, pp.1094-1099, 2008.
- [3] Y. Wang, K. Huang, and T. Tan, "Abnormal activity recognition in office based on R transform," in *Proc. IEEE Int. Conf. Image Process.*, pp.I-341-I-344, 2007.
- [4] H. Foroughi, A. Rezvanian, and A. Paziraei, "Robust fall detection using human shape and multi-class support vector machine," in *Proc. 6th Indian Conf. Comput. Vision, Graphics Image Process.*, pp.413-420, Dec. 16-19, 2008.
- [5] S. Luhr, S. Venkatesh, G. West, and H.H. Bui, "Duration abnormality detection in sequences of human activity," *Technical report, Department of Computing, Curtin University of Technology*, pp.117-132, May 2004.
- [6] T.V. Duong, H.H. Bui, D.Q. Phung, and S. Venkatesh, "Activity recognition and abnormality detection with the switching hidden semi-Markov model," in *Proc. IEEE Comput. Soc. Conf. Comput. Vision Pattern Recog.*, vol.1, pp.838-845, 2005.
- [7] AH Jazwinski, *Stochastic Processes and Filtering Theory*, Academic Press, New York, 1970.
- [8] I. Oikonomidis, N. Kyriazis, and A.A. Argyros, "Full dof tracking of a hand interacting with an object by modeling occlusions and physical constraints," *Computer Vision (ICCV)*, 2011 IEEE Int. Conf. on. IEEE, pp.2088-2095, 2011.
- [9] I. Oikonomidis, N. Kyriazis, and A.A. Argyros, "Tracking the articulated motion of two strongly interacting hands," *Comput. Vis. Pattern Recognit. (CVPR)*, 2012 IEEE Conf. on. IEEE pp.1862-1869, 2012.
- [10] A. Tagliasacchi, M. Schroder, A. Tkach, S. Bouaziz, M. Botsch, and M. Pauly, "Robust Articulated - ICP for Real - Time Hand Tracking," *Computer Graphics Forum*, vol.34, No.5, pp.101-114, 2015.
- [11] K.K. Biswas and S.K. Basu, "Gesture Recognition using Microsoft Kinect@," *Automation, Robotics and Applications (ICARA)*, 2011 5th Int. Conf. on. IEEE, pp.100-103, 2011.
- [12] Y. Li, J. Ohya, T. Chiba, R. Xu, and H. Yamashita, "Study of Recognizing Hand Actions from Video Sequences during Suture Surgeries Based on Temporally-Sectioned SIFT and Sliding Window Based Neural Networks," *IEICE Tech. Rep.*, vol.113, no.493, pp.151-156, March 2014.
- [13] P.J. Besl and N.D. McKay, "Method for registration of 3-D shapes," *Robotics-DL tentative, Int. Soc. for Optics and Photonics*, pp.586-606, 1992.

(平成 29 年 6 月 8 日受付, 10 月 4 日再受付,
12 月 4 日早期公開)



小川 陽子

2006 年立命館大学理工学部情報学科卒業。2008 年同大学院理工学研究科博士前期課程修了。同年(株)構造計画研究所に入社し、2013 年秋より立命館大学大学院情報理工学研究科博士後期課程在籍。日本ロボット学会会員。



松尾 直志 (正員)

2001 年京都工芸繊維大学電子情報工学科卒。2006 年同大学院工学科学研究科博士課程了。2006 年立命館大学総合理工学研究機構研究員。2012 年同情報理工学部知能情報学科助手。2013 年より同情報理工学部知能情報学科特任助教。画像認識

の研究に従事。電子情報通信学会会員。



島田 伸敬 (正員)

1992 年大阪大学工学部電子制御機械工学科卒業。1997 年同大学院博士後期課程修了。博士(工学)。同年大阪大学大学院工学研究科助手。2003 年同助教を経て 2004 年より立命館大学情報理工学部知能情報学科准教授。2012 年より同学部教

授。2007 年より 1 年間米カーネギーメロン大学ロボティクス研究所客員准教授。コンピュータビジョン、ジェスチャインタフェース対話ロボットの研究に従事。電子情報通信学会、情報処理学会、人工知能学会、IEEE 各会員。



白井 良明 (正員)

1964 年名古屋大学工学部機械工学科卒業。1969 年東京大学大学院工学系博士課程修了。工学博士。同年電子技術総合研究所入所。1988 年大阪大学工学部教授。2005 年立命館大学理工学部教授。2012 年より同総合科学技術研究機構客員教授。コン

ピュータビジョン、知能ロボットなどの研究に従事。人工知能学会、情報処理学会、日本機械学会、電子情報通信学会、IEEE など各会員。



来見 良誠

1981 年国立滋賀医科大学医学部医学科卒。1987 年同助手。2000 年同講師。2002 年立命館大学理工学部客員教授兼務(～2015 年)。2005 年滋賀医科大学外科学講座助教授。2011 年滋賀医科大学総合外科学講座教授。2015 年独立行政法人地域医

療機能推進機構滋賀病院長。外科手術(消化器外科・内視鏡外科領域)の診療と研究に従事。日本外科学会、消化器外科学会、内視鏡外科学会、コンピュータ外科学会など各会員。



小森 優

1987 年大阪市立大学大学院工学研究科修了。工学博士。1986 年京都大学医学研究科分子医学系遺伝医学講座助手。UCLA 在外研究員。京都大・医・附属病院医療情報部講師を経て 2001 年滋賀医科大学生命

情報学教授。現在に至る。医用 VR、医用画像処理に関する研究に従事。日本エム・イー学会、VR 学会、VR 医学会、コンピュータ外科学会、日本医療情報学会など各会員。