

マルチモーダルウォークスルー生体認証を想定した複数領域静脈認証システムの基礎研究

代表研究者 中野 和也 成蹊大学 理工学部 教授

1 はじめに

生体認証はカードなどの持ち物やパスワードの記憶が不要であることからその利便性の高さが注目されている。このような特徴からスマートフォンの普及により、生体認証はより身近な本人確認手法の1つとなった。使用される生体情報としては、顔、光彩、指紋、指や手のひらの静脈が代表的なものとして知られている。このような生体認証は、従来はPCやスマートフォンのログインといった本人確認に用いられることが多かった。この認証はユーザー1人が事前に登録した生体情報テンプレートに対して、認証時に取得した生体情報を用いて本人確認を行う1対1の認証であった。一方、大規模な認証の事例として、インドにおいては個人識別番号制度アドhaar（Aadhaar）の個人認証に生体認証が利用されている。さらに、近年では大規模コンサートや空港等の入退室管理のための顔認証ゲートといったものが登場している。このように、既に社会的な実装も進みつつあり、社会的なインフラとして将来的に広く利用されることが予想される。このような認証は事前に登録された多くのユーザーの生体テンプレートの中から1人の生体情報との間で認証を行う1対N認証である。また、このような認証ゲートでは、混雑防止のために短時間で多くの本人確認をする必要があるため、立ち止まらないウォークスルー認証が望ましいと考えられる。しかし、顔認証システムについては、既に複数の企業により実現されているが、顔画像はカメラで簡単に取得可能であり、本人に気づかれずに撮影されることやWeb上でSNS等に顔画像を公開している場合に悪用される危険性も有する。したがって、顔情報に加えて複数の生体情報を用いるマルチモーダル認証が重要となる。先行研究として、我々はこれまで非接触型指静脈認証に関する基礎的な研究を行ってきた[1]。従来の指静脈認証では、ガイドにより指をレンズのピントに合う位置に固定することでボケのない静脈画像を撮影する。しかし、非接触でガイドがなく手を自由に動かせる状態で静脈画像を撮影する場合、レンズの被写界深度の外に指があると画像にボケが生じる。そこで、手が被写界深度の範囲外にある場合に生じる画像ボケを想定して、深層学習を用いた画像改善について提案した[2]。この研究では、ガウシアンフィルタによりボケを生じさせた画像群を生成し、その画像を入力データ、ボケのない画像から静脈パターンを抽出したものを正解データとするデータセットを作製し、Pix2Pix と呼ばれる敵対的生成ネットワーク（GAN: Generative Adversarial Network）に基づく画像生成アルゴリズムにそのデータセットを学習させた。ボケのある画像に対して学習済のPix2Pixを用いてコントラストの高い静脈パターンを生成することができた[2]。このような先行研究を踏まえて、顔以外に追加するセキュアなモダリティとして手の複数領域の静脈に注目した。静脈は手の内部の情報であり、他人が取得することが困難である。さらに、生体認証に使用されるモダリティとして多くの研究や実装事例がある。従来の静脈人認証では、指などを固定具に置いて近赤外光を用いて静止画撮影していた。この場合、ウォークスルー認証においては、ユーザーの動きを止めることになり、ゲート付近で混雑を引き起こす原因となりうる。このことから、カメラを用いて歩行しながら認証が可能なシステムが必要となる。このシステムは、非接触で撮影した手の画像を用いることから、装置に直接触れることがなく衛生的であるため医療施設等にも利用ができる。このシステムの実現のためには、手に何も接触させずに動きの中で撮影した手の近赤外動画画像から認証に必要な特徴を抽出する必要がある。本研究ではウォークスルー型認証を目指して、近赤外カメラで撮影した動く手の近赤外動画画像から手背（手の甲）、指、掌といった手の複数領域の静脈パターンを抽出する。撮影した動画画像には背景や認証に不要な領域も含まれるため、それらを除くために手の動きに合わせて認証に必要な静脈領域のみを追跡することが重要となる。さらに、本研究では、手を3次元で自由に動かすことができるため、ユーザーに対して手の静脈の情報を取得可能な位置を知らせるガイドソフトウェアが必要となる。本研究では、このような静脈領域に対する動的な生体認証を目指した撮影および認証システムについて研究を実施した。

2 実験方法及び結果

2-1 撮影システムの構築及び撮影実験

静脈画像を撮影する際の基本事項について説明する．本研究では手の内部に存在する静脈パターンを個人認証のための特徴として使用する．このような静脈パターンを撮影する際には，近赤外光を用いた撮影することが多い．この理由として，近赤外の波長域において血液中に存在するヘモグロビンによる光の吸収が強いことと，生体組織における光の散乱が可視光に比べて弱いことが挙げられる．このような特徴から，この波長域は「生体の窓」と呼ばれている．この「生体の窓」はおよそ 650 から 900 nm の領域の波長域とされる．本研究においても，静脈認証に適切な波長域の LED 光源を選択し，近赤外カメラを用いて手の静脈画像を撮影する．従来の静脈画像の撮影では，指静脈の場合，指を置いてカメラと指の位置を固定するガイドが存在する．この仕組みにより，ほぼ同じ位置の静脈画像を撮影することが可能となる．さらに，光源とカメラの配置により撮影系は大きく次の 2 つタイプに分けることができる．1 つ目は，透過光撮影タイプである．これは，光源とカメラが向かい合う形で配置されており，撮影時にはカメラと光源の間に手を置く．光源からの近赤外光は指などを透過してカメラに入る．そして，指内部の血管は光に吸収により暗く写り，それ以外の組織領域や背景は明るく写ることで静脈パターンを得ることができる．次に，反射光撮影タイプである．これは，光源とカメラが同じ面で配置されており，光源からの光が手の内部において生体や血管中における光の吸収や散乱により拡散反射された光を撮影する．どちらのタイプにおいても長所と短所が存在する．透過光撮影タイプは生体を透過した光を撮影するので，生体内部つまり手の深部に存在する静脈の情報を捉えることができる．その一方，光源の光が直接カメラに入ることにより起こるアーチファクトや手を透過するためには多くの光量が必要となる．さらに，カメラと光源の配置によりシステム全体が大きくなる．それに対して，反射光撮影タイプは，光源とカメラが近い位置に配置されているので，透過型に比べて小型化することが可能である．ただし，皮膚表面からの光の深達度は浅いことから，透過型に比べて画像中の血管領域と組織領域との間のコントラストが低い．本研究では，このようなことを踏まえた上で，コンパクトに非接触で手の静脈を撮影する方法として反射光撮影タイプを採用した．

撮影システムの基本構成を図 1 に示す．必要なものとして，主に近赤外カメラと近赤外光源から成る．近赤外カメラとしては，近赤外光に高い感度を持つデジタルカラーカメラ (DFK33UX462, The Imaging Source 社) を使用した．近赤外光源としては 850nm をピーク波長とする，近赤外光源 (M850LP1, Thorlabs) と面型 LED 照明 (LDL-220X200IR2-850) を使用した．撮影時に背景に何も映らないようにする必要がある．さらに，近赤外光以外の外光がカメラに入ってしまうと静脈画像のコントラストが低下する．したがって，これらのことを考慮して撮影は卓上の光学暗室内で実施した．撮影システムは光学ブレッドボード上で構築した．構築した撮影システムにおいて画角内に手全体が収まるようにカメラとの距離を設定した．撮影した静脈画像に対する前処理として以下の画像処理の検討を実施した．まず，画像中に含まれるノイズの低減を目的として，平滑化処理を行った．平滑化手法の候補としては，メディアンフィルタ，ガウシアンフィルタ，バイラテラルフィルタなどがあるが，本研究では撮影した画像に対して最も有効であったメディアンフィルタを採用した．さらに，背景に対する血管の強調処理が必要であるが，これに対しては血管強調処理としてよく用いられている適応的ヒストグラム平坦化 (CLAHE: Contrast Limited Adaptive Histogram Equalization) を採用した．次に，画像中の認証に必要な領域のみを抽出する必要がある．この領域を関心領域 (ROI: Region of Interest) とする．本研究ではこの ROI の抽出に Google が提供する MediaPipe[3] の中の 1 つである「Hand Landmark Detection」を利用した．Google のオープンソース機械学習ライブラリである MediaPipe の「Hand Landmark Detection」を利用することで手のキーポイントを 21 点得ることができる．実際に撮影した画像でキーポイントを表示した例を図 2 に示す．これらキーポイントは 3 次元座標を持っている為，ROI の指定に用いた．

最後に，画像間の類似度を評価するために，画像中の特徴点の抽出及びマッチングをする必要がある．通常は，位置合わせをして撮影した画像で上記の処理を行うが，本研究では 3 次元空間中で動く手を撮影していることから，手とカメラの距離の違いによりテンプレートと認証時の画像の手の大きさが異なることがあり得る．さらに，照明の変化や手の回転方向の違いもあることから，大きさや回転にロバストな手法が必要である．このような手法として，本研究では AKAZE (Accelerated KAZE) [4] を用いた．AKAZE を用いて得られた特徴点をマッチングすることでマッチングスコアを算出した．本研究におけるソフトウェアは全て，Python と OpenCV を用いて実装した．

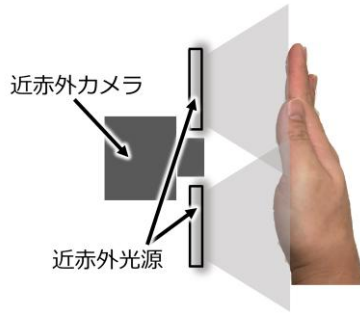


図1 撮影系の構成

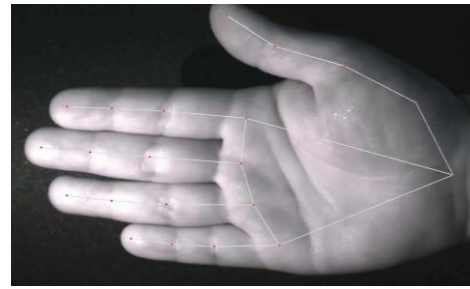


図2 MediaPipeによるキーポイント

AKAZEによる特徴点抽出の例を図3に示す。使用した画像は前処理としてメディアンフィルタ，適応的ヒストグラム平坦化をした後にMediaPipeのキーポイントを用いて切り抜いた手背（手の甲）静脈領域である。さらにキーポイントのパラメータとしてkeypoint_thresholdを変えた。図中の白い点が特徴点を示している。この図が示す通り，適切なパラメータ設定により静脈の領域に多く特徴点が抽出されていることがわかる。次に，テンプレート画像と認証時を想定して撮影した画像で特徴点マッチングをした結果を図4に示す。図4(a)において，左のテンプレートと右の撮影した画像は同じ人の手背である。このとき，左のテンプレートで抽出された特徴点の数が346，右の撮影した画像で抽出された特徴点が896となった。そして，それら特徴点のマッチングの結果は156となった。特に，手背静脈に沿って多くのマッチングが取れていることがこの結果からわかる。次に，異なる手のマッチング結果を図4(b)に示す。図4(b)は左のテンプレートと右の撮影した画像は異なる手背である。このとき，左のテンプレートで抽出された特徴点の数が286，右の撮影した画像で抽出された特徴点が326となった。そして，それら特徴点のマッチングの結果は92となった。これは，同じ手背のマッチスコア156よりも低くなっており，本人識別ができることが示唆された。

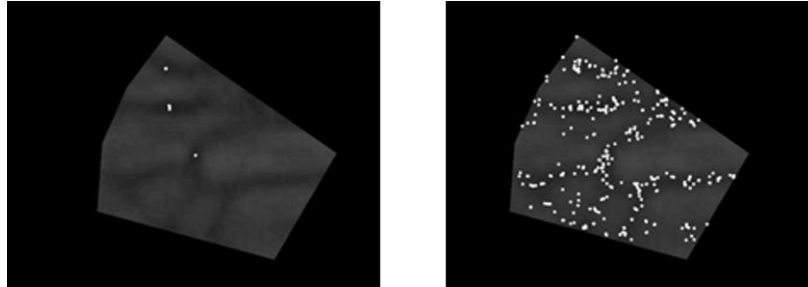
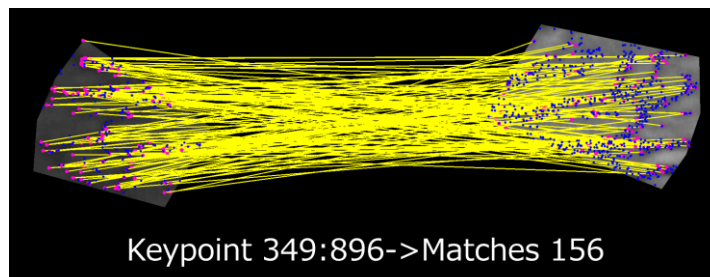
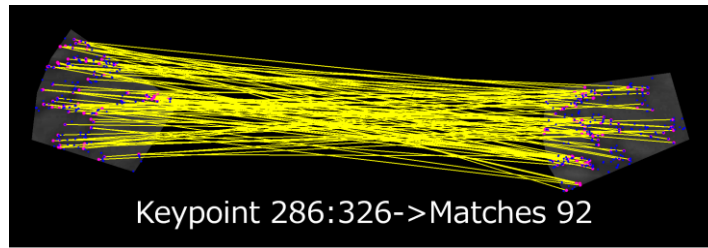


図3 AKAZEによる手背静脈領域の特徴点抽出
(左: keypoint_threshold= 10^{-4} , 右: keypoint_threshold= 10^{-5})



(a)



(b)

図4 AKAZEによる特徴点抽出及びマッチング

2-2 非接触型認証のためのガイドシステム

本研究では、非接触で動く手の画像を撮影するので、従来の接触型の撮影のように手を同じ位置におくガイドが存在しない。この場合、許容レベルの回転や拡大に関してはAKAZEによりマッチングが可能である。しかし、登録したテンプレート画像に対して本人の手であっても手の向きや角度によっては著しくマッチングスコアが低下することが考えられる。図5に想定される手の撮影パターンを示す。例えば、図5(a)のようにカメラに対して平行に手を配置した場合はAKAZEで対応可能であるが、図5(b)や(c)に示すように手の一部がカメラより近くまたは遠くにある場合、同じ手であってもマッチングスコアが大きく低下する可能性がある。特に、反射光を撮影する場合、指先がカメラに近い図5(b)のようなパターンは手に影が生じてしまうことから、MediaPipeによるキーポイント抽出も失敗する可能性が高い。このことから、かざした手の位置が認証可能なレベルであるかを判定してユーザーにフィードバックする仕組みが必要となる。このような仕組みがあれば、ユーザーが手をかざす際のガイドとなる。本研究では、このような判定にMediaPipeで得られる座標を用いることを考えた。MediaPipeで表示されるキーポイントは3次元の座標を持っている。このうち、マッチングスコアに大きく影響するのはカメラからの距離に相当するz座標となる。しかし、MediaPipeはあくまで2次元の画像に対して適用されるものであり、計測した正確な3次元の情報を示しているわけではない。そこで、このキーポイントのz座標に関する実験を行った。最初に手の位置をカメラに平行になる状態を保ったまま、手をカメラに近づけたり遠ざけたりして動画を撮影した。この動画の各フレームに対してMediaPipeにより手のキーポイントを取得し、そのキーポイントのz座標のみを抽出した。その結果を図6に示す。横軸は動画のフレーム番号、縦軸はz座標を示す。この結果から、抽出した全てのキーポイントのz座標が大きく変化していないことがわかった。さらに、各指の位置関係は保たれた状態であることが分かる。次に、手をカメラに対して平行になるようにかざして手首の位置を固定した状態から指先をカメラから近づけたり遠ざけたりして手を動かした動画を撮影する。前述と同様に、各フレームに対してMediaPipeにより手のキーポイントを取得し、そのキーポイントのz座標のみを抽出した。その結果を図7に示す。この結果は動画の各フレームに対するMediaPipeで取得可能な21個のキーポイントのz座標の変化を示している。フレーム番号がおよそ400から600の間でデータが無いが、これはキーポイントが取得できなかったフレームを示している。このときの手は、図5(b)に示すように指先がカメラに近づいている。図7では図6に比べてz軸が大きく変化している。図6と7の結果の違いから次のことが分かる。抽出した各キーポイントは手首を基準とした相対的な座標であることが考えられる。図6では、手首を含めて手全体でカメラに近づけたり遠ざけたりしているので手首を基準とした各キーポイントの相対的な距離は大きく変化しない。その一方で、図7では手首の位置は固定されていて指先をカメラから近づけたり遠ざけたりしているので、手首を基準とした相対的な距離は大きく変化する。このとき、図5(b)のように指をカメラに近づける場合はzの座標は負の値を示し、図5(c)のように指をカメラに遠ざける場合はzの座標は正の値を示した。



(a)

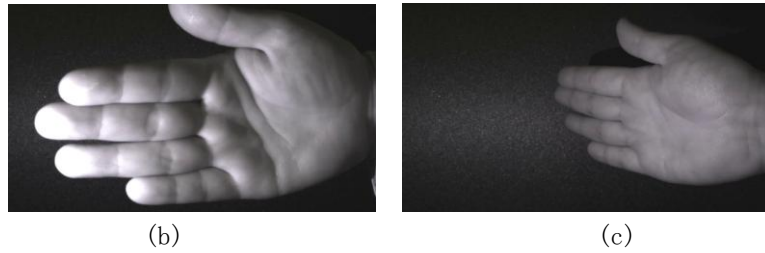


図5 想定される手の撮影パターン

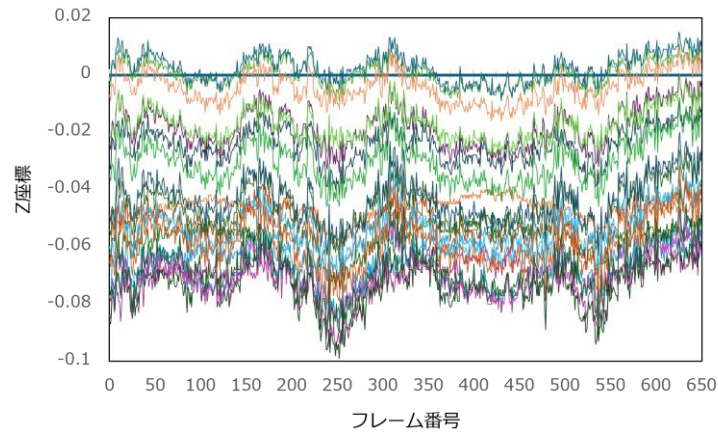


図6 手の移動によるキーポイントの座標変化

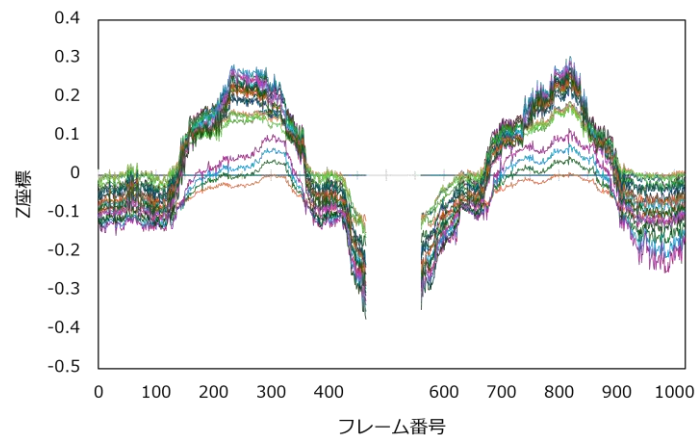


図7 指の移動によるキーポイントの座標変化

MediaPipe で取得されるキーポイントの z 座標と 2 枚の手の画像におけるマッチングスコアの関係が明らかになれば、 z の値またはそれに準ずる値を用いて、認証という観点でカメラに対する手の向きを評価できると考えた。本研究では、キーポイントの座標情報を用いてカメラに対する手の傾きの角度を算出し、AKAZE を用いてマッチングした結果を調べた。角度算出には手首を示すキーポイントを点 A ($z=0$)、指先のキーポイントを点 B とした。さらに、点 B と同じ xy 座標で z が点 A と同じ点を点 C とした。このときの 2 つのベクトル、ベクトル AC とベクトル AB が成す角を内積から求めた。非接触認証では指は自由に動かせるため、撮影した画像中で曲がっていない指もあれば大きく曲がっている指があることも想定される。したがって、手全体だけでなく、指ごとに角度を算出し評価する必要があるため、各指の角度を算出した。カメラに対する指の角度とマッチングスコアに関する結果を図 8 に示す。横軸がテンプレートとする画像と各フレームの画像との間で算出した中指の角度、横軸は算出した中指のマッチングスコアである。中指の角度が変わるについてマッチングスコアも変化した。マッチングスコアのピークの角度はゼロではなく、僅かに負の角度にシフトした。これについては、今後、キーポイントの z 軸の値が実測値とどのくらい相異しているかについて

調査する必要がある．手をかざすユーザーに対してフィードバックする情報としては，指の角度とマッチスコアがある．ここで，マッチングスコアをそのまま提示するよりも，指画像の品質が認証に耐えうるかどうかに関する情報を提示する方がよい．そこで，各指の角度を表示するとともに，指のマッチング数の閾値（本研究では 20 とした）を設定し，その閾値を超える場合は認証可能として「OK」と表示し，閾値以下の場合を「NG」と表示するようにした．図 9 にフィードバック情報を表示した画面を示す．図 9(a) では，カメラに対して平行に手をかざしているので，各指のマッチスコアも高く，すべての指に対して「OK」と表示されている．次に図 9(b) では，指先をカメラに遠ざけると，5 本指のうち 3 本の指（中指，薬指，小指）が「NG」と表示された．さらに傾けると，図 9(c) に示すように，全ての指において「NG」と表示された．また，図 9(d) が示すように，指がカメラに近づくように傾けた場合は，キーポイント自体が取得されなかった．これは，指に影が生じたことによりキーポイントを得られなかったことが考えられる．認証時にこのような指の角度と認証に使用する画像としての採否を表示することで，カメラに手をかざす際のガイドの 1 つとなる．

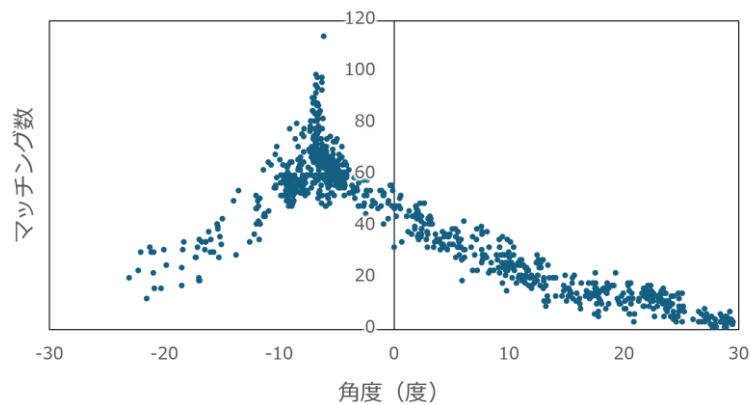
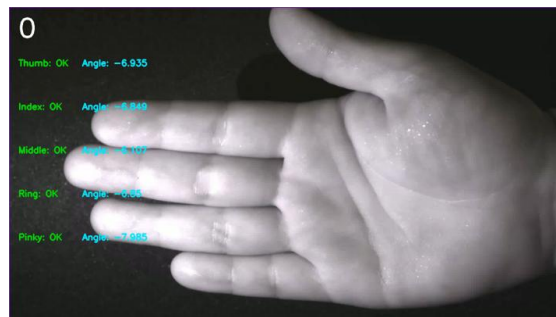
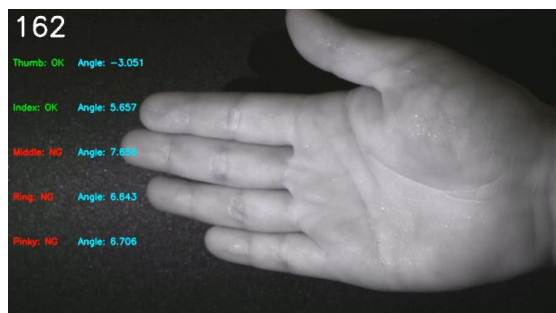


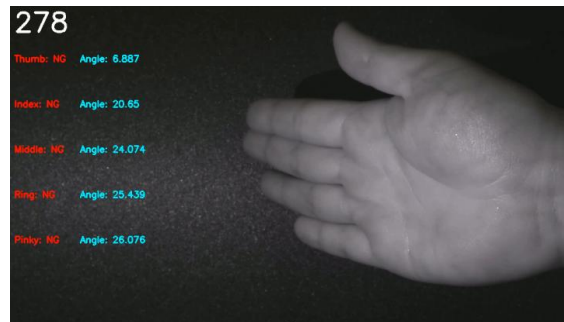
図 8 推定角度とマッチングスコアの関係



(a)



(b)



(c)



(d)

図9 非接触型認証のためのガイドシステム
(左側の表示で緑色がOK, 赤色がNGを示す. 右側の水色の表示は算出した角度を示す)

3 まとめ

本研究では, ウォークスルー認証を想定した静脈認証の基礎的な研究を行った. 従来の静脈認証では, 指を特定の位置に固定するためのガイドが必要であった. 認証時にユーザーがこのガイドに指を接触させて静脈画像を撮影していた. この仕組みはなるべく同じ位置で静脈を撮影することで安定的に静脈画像を得ることができるが, 不特定多数で認証を行う場合衛生面からも適切とは言えない. さらに, 近年, 顔認証入退室ゲートといった大人数の認証を短い時間で行うといったニーズも出てきた. そこで, 本研究ではセキュリティ面も考慮して静脈を用いたウォークスルー認証に関する基礎的な実験を行った. まず, 近赤外光源と近赤外カメラを用いて反射光型の撮影系を構築し手の静脈画像を撮影した. 撮影した画像に対してMediaPipeによるキーポイントを求めてROIの抽出を行った. 抽出した手背領域の画像に対して前処理を行い, AKAZEによるマッチスコアを算出した結果, 本人と他人を識別できることが示唆された. さらに, 非接触認証のための手のガイドシステムも構築した. このシステムではMediaPipeのキーポイントの座標から指の角度を求めてユーザーにフィードバック情報として提示する. さらにAKAZEによるマッチングスコアを算出し, 事前に決めたマッチングスコアの閾値を超えれば「OK」と表示し, 越えなければ「NG」と表示する. この仕組みにより, ユーザーは全ての指が「OK」と表示されるように手や指の位置を調整することができる. 今後は, 多くの被験者による認証精度の評価や手以外の認証についても検討する必要がある.

【参考文献】

- [1] Hiroyuki Suzuki, Narissa Ditthapakdijanya, Takashi Komuro, Keiichiro Kagawa, Kazuya Nakano, Takashi Obi "Finger region extraction using color of skin for hand-waving finger vein authentication", Proc. SPIE 11925, Biomedical Imaging and Sensing Conference 2021, 119250X
- [2] 中野和也, 石川雅浩, 鈴木裕之, 非接触型指静脈認証を想定した深層学習による指静脈画像改善に関する研究, バイオメトリクス研究会 (BioX) 三重大学三翠ホール, 信学技報, vol. 123, no. 28, BioX2023-16, pp. 74-76, 2023年5月.
- [3] Camillo Lugaresi and Jiuqiang Tang and Hadon Nash and Chris McClanahan and Esha

Uboweja and Michael Hays and Fan Zhang and Chuo-Ling Chang and Ming Guang Yong and Juhyun Lee and Wan-Teh Chang and Wei Hua and Manfred Georg and Matthias Grundmann, "MediaPipe: A Framework for Building Perception Pipelines," arXiv preprint arXiv:1906.08172 (2019).

- [4] Alcantarilla, Pablo Fernández, Nuevo, Jesús and Bartoli, Adrien. "Fast Explicit Diffusion for Accelerated Features in Nonlinear Scale Spaces.." Paper presented at the meeting of the BMVC, 2013.

〈発 表 資 料〉

題 名	掲載誌・学会名等	発表年月
ウォークスルー認証に向けた静脈領域抽出及びマッチングに関する基礎研究	2026 年度精密工学会春季学術講演会	2026 年 3 月