

音声情報処理技術の画像処理分野における転用と実用化

Conversion and practical use in the field of image processing of the speech processing technology

梅崎 太造†‡

Taizo Umezaki†‡

†: 名古屋工業大学大学院, umezaki@nitech.ac.jp ‡: 東京大学大学院

概要： 音声情報処理技術を画像処理分野に転用して、これまで実用化してきた技術・製品について報告する。音声の合成・認識に関する技術を応用したものとして、発話訓練機、指紋照合装置、肌判定システム、ナンバープレート認識装置、監視装置、電子部品の自動分類システム、3次元ボディ計測装置、デジタルホログラフィ式3次元顕微鏡、表情変化の自動生成システムの実用化について述べる。さらに、近々実用化予定である市水中の微生物形状や人の血球形状など肉眼では観測不可能な物体を検知・計測する装置についても紹介する。

1. はじめに

1987年に初めて手掛けた製品が商品化されてからこれまでの28年間、特許取得と実用化を重視した研究開発を目指してきた。この間、知的財産の帰属の明確化と技術移転の流れを活性化するために、1999年最初の大学発ベンチャーを起業した。ベンチャー企業を運営する状況下で協業する企業とのマッチングや定款の違いもあり、さらに4つの大学発ベンチャーを起業した。本稿では、これまでに協業企業と共同で実用化した装置やシステムに関して報告する。

2. 発話訓練装置

最初の研究成果としては、音声認識技術を利用した難聴児用発話訓練ソフトウェアシステムがある。これは大学院生のときに研究テーマとは別に、聾学校の先生から頼まれて企業と共同で製作したものである[1][2]。そのソースプログラムは、数社のメーカーに採用され、その中で、G社が「トーキングトレーナ」という名で実用化して販売した(図1)。当時、この手の装置としてはヒットした方で、文部省の標準教材品目にも指定され、全国の聾学校の授業等で使用された。これにより、初めて産学連携および福祉分野の研究を経験できた。現在は、ゲーム性や三次元的表現を取り入れたもの、さらにはアドベンチャーゲームの要素を組み込んだソフトも開発済みである(図2-図4)。



図1 発話訓練装置

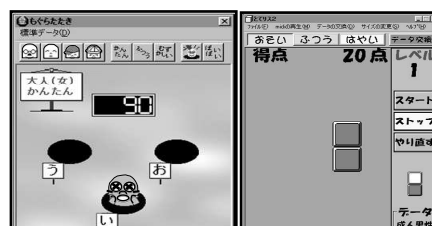


図2 もぐらたたき 図3 音声テトリス



図4 3D感覚を取り入れた訓練ソフト

3. 指紋照合装置

音声認識技術と画像処理技術を併用したアルゴリズム(指紋画像に線形予測分析技術によるスペクトル解析を適用した手法)を用いた指紋照合装置をメーカーと共同で開発して実用化した[3][4]. ケプストラム分析により、指紋画像中に含まれる”指紋らしさ”の成分を除去し、個人性を強める手法を適用している(図5). 本装置は、現在も各方面(官公庁, 企業, 原子力発電所等の重要施設における入退出管理装置やホームセキュリティ装置など)で活用されている[5].

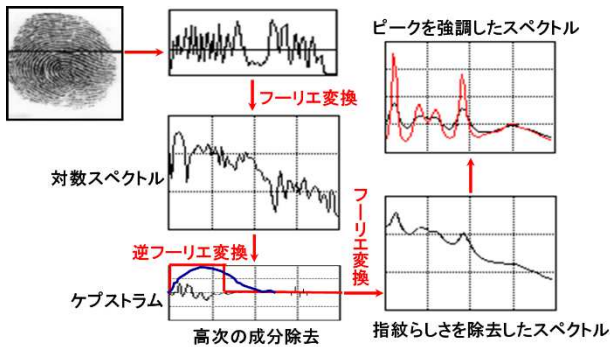


図5 指紋画像中の個人性強調分析例



図6 市販された指紋照合装置

モバイル機器の普及に伴い、小型で場所を取らないラインセンサが組み込まれるようになり、それに特化した照合法を開発した[6]. 音声認識アルゴリズムを組み込むことにより、指を滑らせる速度に関係なく照合可能である(図7, 図8).

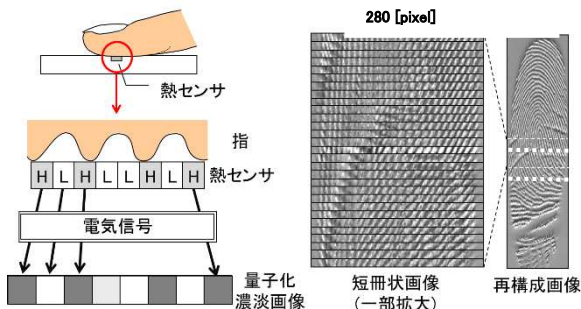


図7 ラインセンサによる指紋画像センシング



図8 ラインセンサを組み込んだ装置

4. 肌のキメ判定

従来、高価な専用の撮影装置を用いて撮影した肌画像に対して、医師や専門家の専門知識により肌状態を評価する手法が用いられていた[7][8]. 肌の状態を鮮明に撮影することが可能であり、水分計や油分計を搭載したものも存在する. そのため利用場所は、病院やサロン・エステ・化粧品販売店などに限られ、施術前後の診断用として使用する場合が大半を占めていた. 故に一個人が手軽に使用できるシステムではない.

そこで、専門知識のない一個人でも手軽に使用できるスマートフォンを用いた肌状態の評価手法を開発した. 肌のキメの度合いを判定するアルゴリズムには、音声情報処理技術である線形予測分析法を用いている(図9)[9]. スマートフォンに2波長LED付きのアタッチメント式レンズを装着することで、肌状態を鮮明に撮影できる[10](図10, 図11).

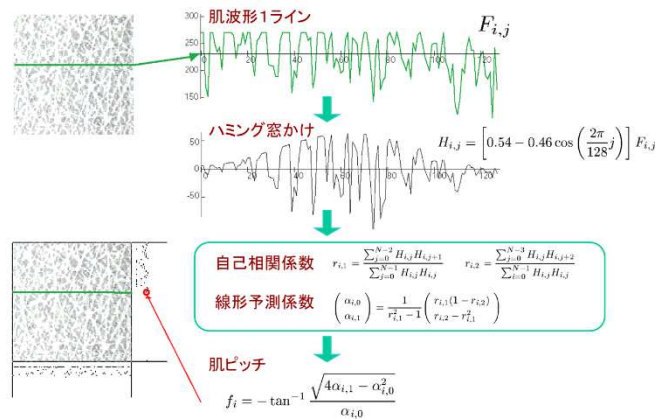


図9 肌のキメ判定アルゴリズム



図 10 アタッチメント式レンズ装着

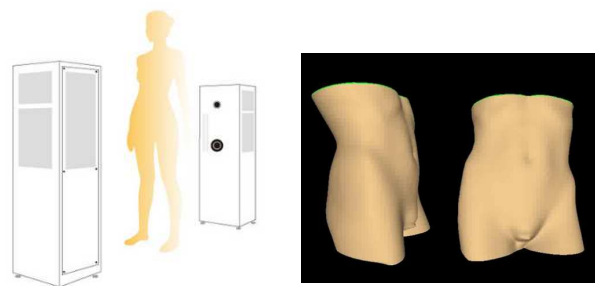


図 13 ボディ形状計測装置と計測例

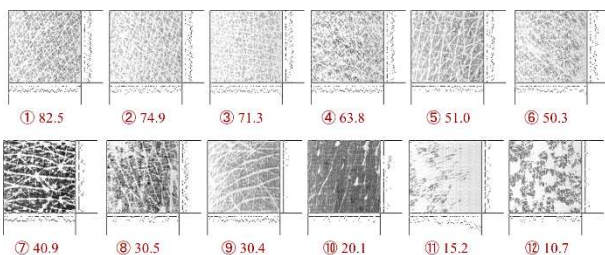


図 11 肌のキメ判定(数値が高いほどキメが細かい)

5. 3次元ボディ計測装置

これまでのエステティック産業は、“美容”を中心とするサービスを主体としてきたが、平成19年4月から健康保険加入者へのメタボリック検診及び指導が開始されたことに起因して、“健康”を中心とする消費者ニーズが高い。新たな市場として多様な健康・癒し系サービスが急増しており、高い成長率が見込まれる。我々は、フェイシャル・ボディ・脱毛などの中でも特にボディに着目した。ボディ形状の経月変化を3次元形状計測技術により詳細に提示することで、「安心・信頼」を消費者に提供することを目指す。計測法としては、空間コード化法と位相シフト法を併用することで精度を向上させている[11][12](図12, 図13)。

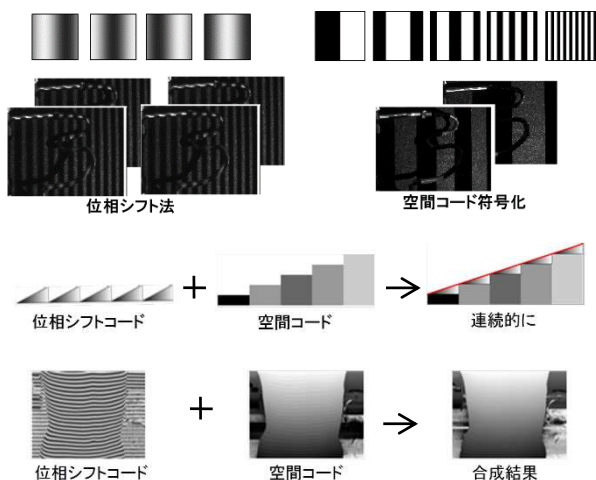


図 12 空間コード化法と位相シフト法の併用

6. デジタルホログラフィ式3次元顕微鏡

既存の高分解能顕微鏡としては、走査型電子顕微鏡および共焦点レーザ顕微鏡が最も使用されている。前者では、観察対象をスキャンしながら物体の微細な構造を高分解能で撮像できる。しかしながら、物体構造の3次元形状は取得できない。一方後者では、計測対象の形状を3次元計測できるが、装置が高価であり、計測精度が1[μm]程度であるため、ナノ構造の物体に対する3次元計測は困難である。そこで、前記問題を解決するために、デジタルホログラフィを用いた安価かつ小型な3次元顕微鏡を開発した(図14)。位相シフトデジタルホログラフィの原理を利用することで、ナノメートルオーダーの高精度3次元計測を実現している[13][14]。さらに、干渉縞の周波数解析法を応用して、ワンショット方式の安定かつ高速な3次元計測手法を開発した。人の赤血球と水中の活性汚泥微生物の3次元形状を計測した結果を図15, 図16に示す。1回の計測に要する時間は約0.1[s], 計測精度は5.5[nm]である。

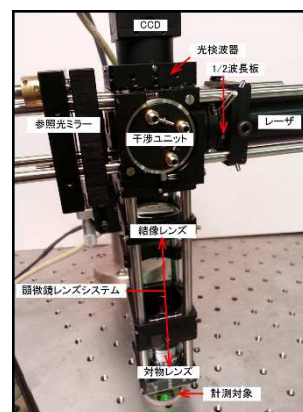


図 14 デジタルホログラフィ式3D顕微鏡

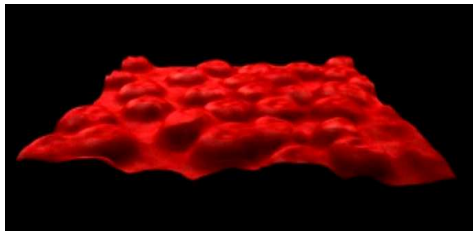


図 15 人の赤血球の計測例

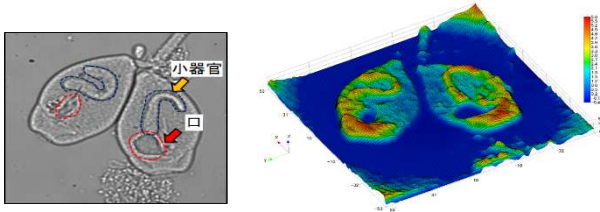


図 16 活性汚泥微生物の計測例

7. 表情変化の自動生成

従来、アニメや映画等で利用されている合成発話表情画像の多くは、一画ずつマニュアル操作により作成されている。半自動的に画像生成できる方法も発表されているが、短時間の音響変化に対応するために大容量の動画データデータベースを持たなければならない。これらの欠点を補う手法として、コホーネン型ニューラルネットワーク (KNN) モデルによる発話表情画像の自動生成法を提案した [15] (図 17)。

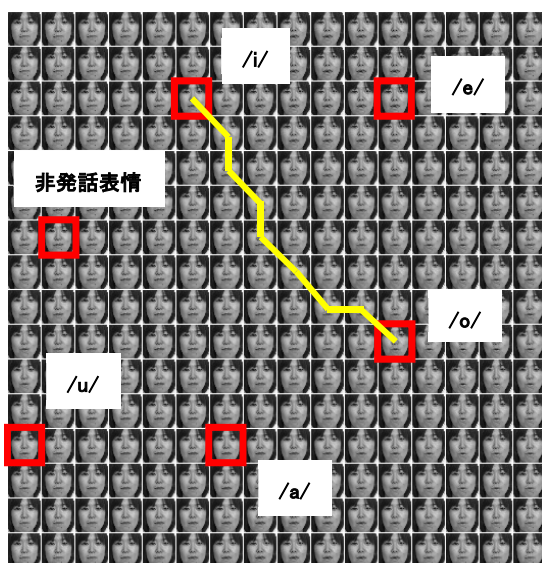


図 17 コホーネンネットワークによる表情作成

一般に日本語の単音節は子音+母音で構成されているが、子音の継続区間長は母音継続区間長に比べて短く前後続母音の影響を受け易いため、発話表情への影響が少ない。この特性を利用して、母音から母音への発話表情系列を KNN により自動生成し、音声と同期させることにより、コンテキストに依存しない発話表情の合成を実現している。最適表情変化経路探索に DP マッチングを用いている。

8. ナンバープレートの自動認識

車両走行中は画像ぶれがあるため、頻りにナンバープレート文字の切り出しミスが起こり、ケタ数も誤る場合が多い。バンパーとプレートの“境目”を切り出して文字“1”と判定してしまう現象は、その典型的な例である。日本のナンバープレートは桁数が既知であり、陸支コードや車種コードなど、文字が小さい部分を除けば数字のみで構成されている。これらは位置情報を用いて補正を行うことも可能であるが、海外では桁数が異なる場合もある。また、英数字の組み合わせでパターンも多く、誤判別しやすい。そこで、前フレームの認識結果を用いて、現フレームの文字切り出し結果を補正する手法を提案した [16]-[18] (図 18)。

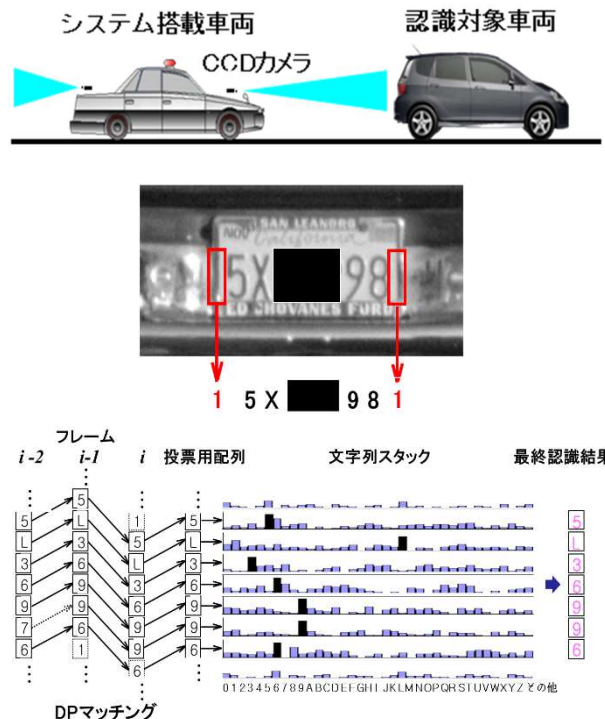


図 18 ナンバープレート文字の切り出し

本手法では、認識対象とする文字列の桁数は未知と仮定し、前フレームの切り出された文字列と現フレームの文字列の対応点を求めて、切り出し位置のずれを吸収することで誤りを補正する。対応点の算出には、DP マッチングを用いる。

9. 動体認知型監視カメラ

画像処理技術を用いた動体検出における従来手法の多くは、画像中に存在する移動物体を検出するもので、移動物体全体が状況により画像に含まれない場合や障害物等で隠蔽される場合、認識が困難となる。そこで、画像領域として短冊領域(オレンジ部分)のみを検出領域に選び、通過する移動物体を時系列処理により検出する手法を提案した[19](図 19)。これにより、移動物体の検出に利用可能な画像領域が限定される場合においても検出可能となる。短冊部分の重なり幅とその変化量により動体の速度・加速度も求まる。動体の速度により、短冊連結長は変化するが、DP マッチングにより速度に関係なく動体を特定できる。

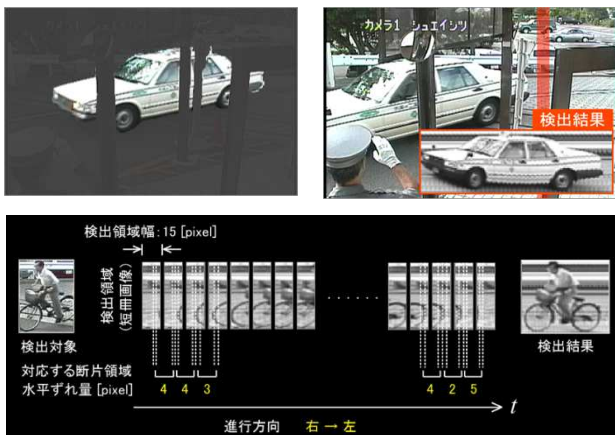


図 19 短冊フレームによる動体検知と認識

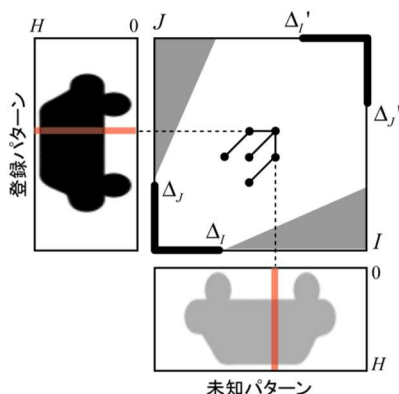


図 20 動体識別用 DP マッチング

検出した移動物体に対してさらに、DP マッチングによる識別実験を実施した結果、約 98[%]の物体識別率を得た(図 20)。動体が通過したときの画像データのみ記憶するため約 50 倍記憶可能である。また、1 短冊だけでも検知できるため、カメラを隠すことも可能となる。本手法は、指紋ラインセンサにヒントを得たものである。

10. 電子部品の自動分類

電子部品は、実装機(チップマウンタ)によりプリント基板上に実装されるが、近年部品の複雑化や生産形態の変化(多品種少量生産)に伴い、部品種別の分類や形状データ作成に時間とコストが増大している。そこで、電子部品の分類作業を自動化する手法を提案した[20]。一般に、画像分類の研究においては、形状特徴が用いられる。しかし、電子部品における分類は、部品の大きさやリード数が異なる場合においても同じ部品であると認識する必要がある。そこで、DP マッチングを適応した(図 21)。BGA 系に対して分類実験した結果、95.7[%]の分類率が得られた(図 22)。クラスタリングや判別分析法と組み合わせることで良好な検出結果を得ている(図 23)。

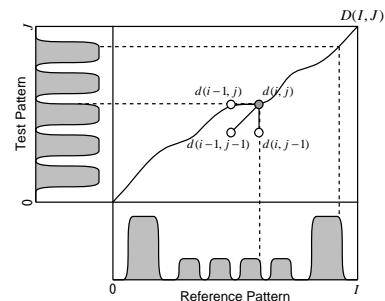


図 21 部品画像のゼロクロス数による分類

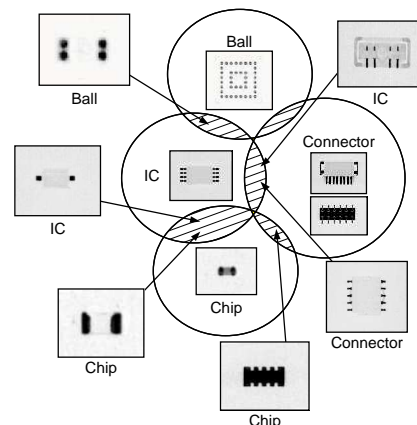


図 22 BGA 系部品のクラス間境界例

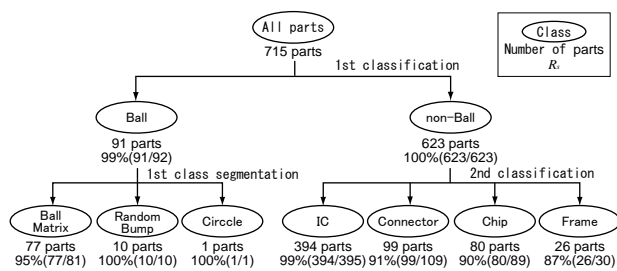


図 23 全部品の自動分類ツリー

11. むすび

平成 13 年経済産業省の「大学発ベンチャー1000社創出計画」が出されてから、大学発ベンチャー起業の数は、目標を達成した平成 16 年度以降上昇気配であったが、平成 20 年度以降は 1800 社相当の横這い状態である。これまで、一握りのベンチャーが成功したとの話が雑誌やテレビ等で話題にはなったが、本当のところ、その経営実態はどうであろうか。正直なところ、「成功した」という話を聞くことは私の周りではほとんど無い。大学発ベンチャー起業を実際に経験してみると、事務処理、契約や特許処理などの法務処理、利益回収までの資金調達、利益相反や人件費問題など解決しないといけない課題が山積みである。大学内における起業に対する賛成派が少ない中、それでもチャレンジする研究者の創出に今後期待したい。

謝辞 本稿で紹介した装置の開発に協力していただいた大学研究者や企業の方々に深謝する。

参考文献

- [1] 梅崎太造, 倉谷和彦, 藤吉弘亘: “ネットワーク環境を利用した聴覚障害児用発話訓練支援システム”, 電子情報通信学会(D-II), J80-D-II, 4, pp.925-932, 1997.
- [2] 梅崎太造, 間瀬寛子, 江川敦史: “ゲーム性を取り入れた聴覚障害児用発話訓練ソフトウェアの開発”, 日本福祉工学会誌, Vol.2, pp.31-36, 2000.
- [3] 梅崎太造, 板倉文忠: “全極型フィルタの群遅延スペクトルによる音声分析とその音声認識用スペクトル距離尺度への応用”, 信学論, J72-D-II, 8, pp.1141-1150, 1989.
- [4] 松本憲幸, 佐藤省三, 藤吉弘亘, 梅崎太造: “LPC 分析に基づく指紋照合法”, 電気学会論文誌, Vol.122-C, 5, pp.799-807, 2002.
- [5] <http://www.chkk.co.jp/products/cablex/tech.html>
- [6] http://www.dds.co.jp/product/ubf_unit/

- [7] 竹前嘉修, 斉藤英雄, 小沢慎治: “皮膚表面画像を用いた肌システム”計測自動制御学会論文集 Vol.37, No.11, pp.1097-1103, 2001.
- [8] 小林宏, 橋本卓弥, 山崎和広, 平井義和: “画像解析による肌のきめの定量的指標の提案と実用化”日本機械学会論文集(C編), Vol.76, No.764, pp.922-929, 2010.
- [9] H. Takeuchi, M. Hoguro, T. Yoshimine, T. Umezaki: “Skin Diagnosis Algorithm for Analysis of Skin Images Captured from a Fingerprint Sensor”, International Journal of Computer Applications in Technology, Inderscience Publishers, Vol.34, No.4, pp.335-342, 2009.
- [10] <http://hadamoreinfo.maxell-sc.co.jp/hadamore/>
- [11] S. Zhang, P. S. Huang: “High-resolution, real-time three-dimensional shape measurement”, Optical Engineering, Vol.45, No.12, 2006.
- [12] <http://style-scan.com/stylescanx2/include/>
- [13] I. Yamaguchi, T. Zhang: “Phase-shifting digital holography”, Opt. Lett., Vol. 22(16), pp. 1268-1270, 1997.
- [14] T. Zhang and I. Yamaguchi: “Three-dimensional microscopy with phase-shifting digital holography”, Opt. Lett., Vol. 23(15), pp. 1221-1223, 1998.
- [15] 木村成宏, 佐藤省三, 梅崎太造: “発話表情の自動生成”, 日本福祉工学会誌, Vol.4, No.2, pp.21-28, 2002.
- [16] 藤吉弘亘, 今村友彦, 梅崎太造, 金出武雄: “ニューラルネットワークによるナンバープレートの位置検出”, 電子情報通信学会論文誌, Vol.J80-D-II, No.6, pp.1627-1634, 1997.
- [17] 佐藤省三, 藤吉弘亘, 梅崎太造, 金出武雄: “ナンバープレート内の一連番号の切り出しと認識”, 電気学会論文誌(C), Vol.121, No.8, pp.1354-1361, 2001.
- [18] 保黒政大, 岡本直樹, 能勢隆, 梅崎太造: “車両搭載型ナンバープレート認識システム”, 電気学会論文誌(C), Vol.126, No.5, pp.589-595, 2006.
- [19] 保黒政大, 井上祐貴, 梅崎太造, 説田武文: “短冊フレーム画像を用いた移動物体検出”, 電気学会論文誌(C), Vol.128, No.8, pp.1277-1285, 2008.
- [20] 江崎弘健, 鍵井清幸, 梅崎太造, 堀越哲美: “電子部品の階層型自動分類システムの開発”, 電気学会論文誌 C, Vol.126, No.12, pp.1447-1453, 2006.

梅崎太造: 1982 年豊橋技科大情報工学科卒。1984 年同大大学院修士課程了。1987 年名大大学院博士後期課程了。工学博士。同年 4 月同大助手。1990 年中部大講師。1992 年同大助教授。1993~1994 年カーネギーメロン大客員研究員。1999 年中部大教授。2003 年名工大大学院教授。2015 年東大大学院特任教授。音声・画像情報処理, 三次元計測, 福祉ロボットの研究に従事。2003 年日本福祉工学会論文賞, 2006 年グッドデザイン賞, 2008 年産学官連携功労者表彰科学技術政策担当大臣賞受賞。電気学会, 情報処理学会, 日本音響学会, 電子情報通信学会, 日本福祉工学会, ヒューマンインターフェイス学会, 生体医工学会, 日本ロボット学会, 精密工学会各会員。