

語句が持つ動作の特徴を用いた 身体動作データの類似検索法

湯川 崇^{†1,†2} 玉本 英夫^{†3}

身体動作を記録するためにモーションキャプチャの利用が普及しているが、装置が高価であることや必要な動作を取得するためには演技者や場所が必要となることから、記録済みの動作データを再利用可能にするためのデータベースの構築が望まれている。これまで、舞踊符というモーションキャプチャを使用した身体動作の記録および記録済みデータの再利用を目的とした手法が提案されている。舞踊符は動作データを基本的な動作のまとまり（基本動作）に分割し、その動作の名前を符号として割り当てることにより作成されているが、動作の分割や符号の割当ては手作業で行われるため、大量のデータを扱うことが困難である。本論文では、パターン認識の一手法である部分空間法を用いて、基本動作とその動作の名前が含む語句との関連を定量化し、その結果を基本動作データの類似検索、および基本動作の自動作成の代替として利用する手法を提案する。提案手法により、1) 基本動作データをデータベースへ登録するときの自動インデクシング、2) 動作を表す語句を検索クエリとした基本動作の検索、3) 動作の特徴により作成した基本動作の候補からの基本動作の検索が可能になる。提案手法を評価するために基本動作のテストデータを作成し、基本動作の検索実験を行った。その結果、提案手法が基本動作の自動生成、および動作を表す語句による基本動作の類似検索の有効な手法となることが確認できた。

A Similarity Search Method of Human Motion Data Using Features of Motion Words and Phrases

TAKASHI YUKAWA^{†1,†2} and HIDEO TAMAMOTO^{†3}

It is considered that a database of human motion data recorded using a motion capture system should be constructed. The method named BUYOFU is proposed to record human motion by using a motion capture and to reuse it. BUYOFU is created by dividing the motion data to a unit of basic motion and assigning a name of its motion to it. Because those works are performed manually, it is difficult to make a large quantity of BUYOFU. This paper proposes a method to quantify the relationship between the feature of a human motion and its name using the partial space method that is one of the pattern recogni-

tion techniques. Using the proposed method, 1) indexing can be automatically performed when registering human motion data into the database, 2) we can search targeted human motion data from the database using motion words and phrases as a query, and 3) we can search targeted human motion from automatically created candidates of basic motion. In order to evaluate our proposed method, we created test data sets for basic human motion, and made an experiment. As a result, it turned out that our proposed method would become a promising candidate of an efficient method for searching human motion data from the database.

1. はじめに

近年、映画やビデオゲームの製作過程において人間のキャラクタに現実感のある動きを付けるための身体動作データを生成する手法として、モーションキャプチャが広く利用されている。モーションキャプチャは身体動作を3次元時系列データとして記録するための有効な手法であるが、装置が非常に高価であることや、動作を記録するための演技者および場所が必要であることから手軽に利用することが難しい。そのため、モーションキャプチャを用いて記録した身体動作データをデータベース化して再利用可能にする手法が望まれている^{1),2)}。

身体動作データの記録保存および再利用を目的として、モーションキャプチャを用いた「舞踊符」という手法が提案されている^{1),3)}。これは、記録した身体動作データを舞踊の振りなどの基本となる短時間の動作の単位（基本動作）に分割し、動作の名前（動作名）を符号として割り当ててデータベース化することにより、身体動作の記録や新しい動作データ生成の素材として再利用可能にするための手法である。舞踊符の作成は、基本動作への分割および符号の割当ての2つの手順からなるが、どちらも対象動作に関する専門知識と多大な労力を必要とするため、これまで舞踊符作成の省力化を目的とした研究が行われている⁴⁾⁻⁷⁾。

基本動作の分割については、着目する部位の速さや角度の変化に基づいて動作データを分割する手法^{4),6)} や自己組織化マップにより類似した姿勢をクラスタリングすることで分割を行う手法⁵⁾ が提案されている。これらの手法を用いることにより基本動作の分割作業が

†1 ノースアジア大学法学部

Faculty of Law, North Asia University

†2 秋田大学大学院工学資源学研究所

Graduate School of Engineering and Resource Science, Akita University

†3 秋田大学工学資源学部

Faculty of Engineering and Resource Science, Akita University

省力化されるが、人間の判断で行っている手作業による分割を自動化することは困難であり、完全な自動化には至っていない。

基本動作への符号の割当てについては、バレエのように基本動作と動作名との関連が体系的に定義されている動作であれば、動作名を符号として割り当てる作業はそれほど労力を必要としない。しかし、大規模な動作データベースの構築を目的とした場合、手作業ではなく自動化されることが望ましい。また、民俗芸能の舞踊の中には、名前を持たない基本動作があり、そのような動作では基本動作に適切な符号を割り当てるのが困難である。

本論文では、モーションキャプチャを用いて記録した動作データを基にした基本動作データベースの構築の自動化を目指して、動作名に含まれる語句と基本動作の特徴との関連を定量化し、その結果を基本動作のインデクシングおよび検索に利用する手法を提案する。

本論文の構成は次のとおりである。2章で身体動作データベースに関する関連研究と提案手法の概要を述べる。3章で動作データから語句の特徴を抽出する手法について述べ、4章で動作を表す語句を検索クエリとして基本動作を検索する手法、および動作データの特徴を用いて作成した基本動作の候補から、目標とする基本動作を検索する手法について説明する。5章、6章で提案手法の評価実験とまとめを行う。

2. 身体動作データベース

2.1 関連研究

身体動作データなどのマルチメディアコンテンツのデータベースで使われているコンテンツ検索手法は2つに大別できる。1つは、データの内容を表す語句をメタデータとして検索対象データにマークアップしておき、検索対象データそのものの代わりにメタデータを検索に使用する手法である。この方式ではデータ内容を言語化する過程を手作業で行うため時間と労力を要すること、またメタデータ作成者の主観的な判断を必要とするという問題がある⁸⁾。

もう1つの手法は、データの内容そのものを比較することによる検索である。これは内容検索 (content-based retrieval) と呼ばれ⁹⁾⁻¹¹⁾、静止画、ビデオ、3Dモデル、動作データ⁸⁾、音楽⁹⁾ など様々なマルチメディアコンテンツを対象に広く使われている。内容検索では、検索キーとしてデータベース中のコンテンツと同じ種類のデータを使用する。その際、検索キーに誤差を含んだ曖昧なデータが用いられる可能性があるため、内容検索の検索結果は検索キーとの類似度の高い順に与えられる類似検索方式をとることが多い⁹⁾。

類似検索方式は、コンテンツの特徴を抽出する手法と特徴間の類似度を求めるための識

別手法により分類できる。特徴抽出手法には、フーリエ変換¹²⁾ やウェーブレット変換¹³⁾、主成分分析⁸⁾ を用いたものがある。特徴間の類似度の計算には、DTW (Dynamic Time Warping) が多く使われているが、計算に必要な時間と記憶容量が多いという問題があることが知られている²⁾。類似度比較の高速化と柔軟性を実現するために単一スケールを使用する手法が提案されている¹⁴⁾。

一方、特徴抽出と識別を同時に行う検索方式として、パターン認識の一手法である部分空間法がある。部分空間法は、学習パターンが表す特徴の低次元部分空間を求め、その部分空間への入力パターンの特徴ベクトルの射影の長さを類似度とする¹⁵⁾。これまでに、文字や顔画像¹⁶⁾⁻¹⁸⁾ などの静止物のほか、表情やジェスチャー¹⁹⁾ などの動きの認識などへの部分空間法の適用が報告されている。これらの手法では、顔や手の特徴を得るための入力パターンとして画像データが使われることが多い。顔画像を認識するためには基本的に1枚の静止画を用いるが、認識精度の向上や動作を認識するため、複数枚の静止画像や動画画像を使用する手法も提案されている¹⁶⁾⁻¹⁸⁾。

2.2 提案手法の概要

本研究で対象としている人間の動きは動作のタイミングや位置の変動が大きく、同一人物が1つの動作を演じた場合でも、別々に記録した動作データが一致することはほとんどない。そのため、他のマルチメディアコンテンツと同様に類似検索が有効と考えられる。

動作データを検索クエリとして使用する類似検索では、クエリを作成するための手軽に利用できる動作データの入力デバイスがないため、検索システムの実用化が困難となる。また、誰にでも簡単に利用できる基本動作の検索システムを構築するためには、Webのキーワード検索と同様に語句を検索クエリとした検索インタフェースが望ましいと考える。そこで本論文では、基本動作のインデクシングや検索に、動作を表す語句と動作データとの関連を利用する手法を提案する。提案手法の実現のためには、動作を表す語句と動作の特徴との関連を定量化する必要があるが、これは部分空間法を用いることにより可能になる。

提案手法では、基本動作データを n 点リサンプリングした値を基本動作の特徴ベクトルとして使用する。そして、特徴ベクトル空間中で動作名に含まれる個々の語句が分布する部分空間を求め、インデクシングや検索に利用するために部分空間法を適用する。部分空間法が使用する学習データは動作名が割り当てられた基本動作から構成され、あらかじめ手作業により作成しておくこととする。

提案手法により、メタデータをマークアップする作業や、動作データを入力するための高価なデバイスを必要とせず、動作を表す語句を検索クエリとする基本動作データの検索が

可能になる。

さらに、これまで手作業で行われていた基本動作への分割を自動化する手法を提案する。この手法は、手作業で作成した基本動作の分割点の特徴を調査し、その結果明らかになった特徴を基に動作データを分割することで、基本動作の候補を作成する。また、作成した基本動作の候補から、動作を表す語句と動作データの特徴ベクトルとの関連を用いて、適切な基本動作を検索する手法を提案する。

3. 語句の特徴の抽出

3.1 身体動作データの取得

本研究では、演技者の動作への制約が少ないこと、およびセンサを装着した部位の位置と方向が同時に計測できることを考慮して、Ascension Technology社のワイヤレスにより記録可能な磁気式モーションキャプチャを使用して身体動作データを取得した。

図1にセンサの取付け位置を、図2に本研究で使用する身体モデルの階層構造をそれぞれ示す。演技者に磁気センサを取り付ける位置は、図1中に楕円で示した手首、肘、肩、首、腰、股、膝、足首の各関節で区切られる、合計15カ所の身体の可動部(セグメント)である。楕円の中の黒丸印はセンサを表している。

磁気式モーションキャプチャの1つのセンサにより取得されるデータは、3次元空間でのセンサの位置(x, y, z)とオイラ角で表される方向(azimuth, elevation, roll)との6成分からなる時系列データである。動作データのサンプリングレートは30フレーム/秒とした。

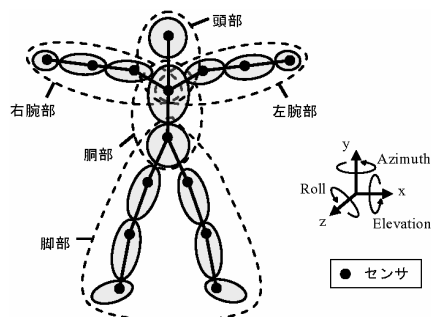


図1 センサの取付け位置と部位の定義
Fig.1 Definitions of sensor positions and body parts.

3.2 基本動作データの作成

本研究で構築する基本動作データベースは様々な目的への応用を想定し、身体動作を個々の部位の動作の集まりとして扱う。そのため、各センサから取得した動作データを、図1中に破線で示した頭部、右腕部、左腕部、胸部、脚部の5つの部位に分割し、それぞれの部位ごとに基本動作データを作成する。本研究では「うなづく」「手を振る」「座る」などの短時間の単純な動作を基本動作とする。

動作データは、演技者の体型や演技者へセンサを取り付ける位置や角度の違いにより、1つの動作を取得しても異なるデータとして記録される。基本動作としてデータベースに登録するためには、基本動作が演技者の体型や動作取得環境と無関係であることが望ましい。そこで、モーションキャプチャにより取得したデータに対して、位置、角度、およびスケールの標準化を行う。

位置の標準化 センサを取り付ける位置は同一人物でも収録ごとに变化するが、関節の位置は同一人物であれば変化することはない。そこで、標準化の最初のステップでは、取得データから各関節の位置を推定することにより動作取得時のセンサの取付け位置の影響を除く。関節位置の推定は、その関節をはさむ2つのセグメントに取り付けたセンサを i, j (腰に近い方を i とする)、フレーム k でのそれぞれの位置を p_{ik}, p_{jk} とすると、関節の位置は関節の角度が変化しても p_{ik}, p_{jk} からの距離がつねに一定であることを利用する²⁰⁾。すなわち、回転行列で表した i, j の方向をそれぞれ R_i, R_j とし、 i, j から見た関節の位置をそれぞれ c, l とする。ここで

$$d_k = -(p_{ik} - p_{jk}) \tag{1}$$

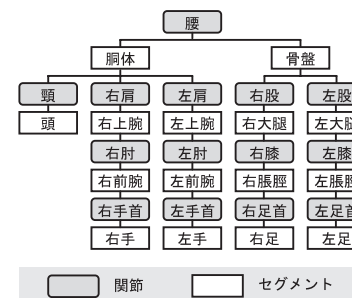


図2 身体モデルの階層構造
Fig.2 Hierarchical structure of the human body model.

$$Q_k = [(R_{ik}) (-R_{jk})] \quad (2)$$

とおくと、すべてのフレーム ($k = 0..n - 1, n > 2$) について、

$$\begin{bmatrix} Q_{i0} \\ \vdots \\ Q_{ik} \\ \vdots \\ Q_{in-1} \end{bmatrix} \begin{bmatrix} c \\ l \end{bmatrix} = \begin{bmatrix} d_{i0} \\ \vdots \\ d_{ik} \\ \vdots \\ d_{in-1} \end{bmatrix} \quad (3)$$

が成り立つので、この式を解くことにより c と l が得られる²⁰⁾。フレーム k での関節位置 P_k は c と l を用いて次式により得ることができる。

$$P_k = ((R_{ik}c + p_{ik}) + (R_{jk}l + p_{jk}))/2 \quad (4)$$

角度の標準化 次に、推定した関節の位置を基に関節の角度を求める。図 2 の身体モデルにおいて、親子関係にある 2 つの関節を m (親) と n (子)、それぞれの位置を P_m, P_n とするとき、関節 m の角度を次のように定める。

関節 m の方向が初期状態 ($azimuth, elevation, roll = (0, 0, 0)$) のときに P_m から P_n へ向かう単位ベクトルを $q = P_n - P_m$ とする。また、位置の標準化で求めた動作中の関節 m, n の位置をそれぞれ P'_m, P'_n とし、 P'_m から P'_n へ向かうベクトルを $v = P'_n - P'_m$ とする。このとき関節 m の角度は、回転の中心軸 n を

$$n = v \times q \quad (5)$$

その周りの回転角度 θ を

$$\theta = \arccos \frac{v \cdot q}{|v||q|} \quad (6)$$

とする回転量として得られる²¹⁾。

ここで、回転量をオイラ角や回転の中心軸とその周りの回転角などの形式で表すと、360 度以上回転したときに 360 度と 0 度との間で値が不連続となる。これを避けるため、動作データの回転量を 3×3 の回転行列に変換し、その行列の 9 つの成分を基本動作の回転の特徴を得るために使用する。

スケールの標準化 位置と角度について標準化した後、平均的な体型の標準モデルを定義し、標準モデルの身体各部のサイズに適合するようにデータのあてはめを行う。標準モデルの各部位のサイズには H-Anim²²⁾ で提案されている値を採用する。

ここまでの作業を標準化と呼ぶ。基本動作データベースには標準化を行った動作データを

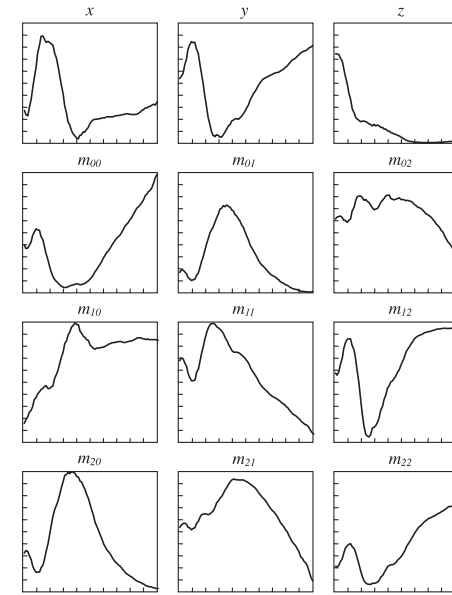


図 3 基本動作データの例
Fig. 3 Examples of basic motion data.

登録する。

標準化の次に、基本動作データから特徴を抽出するための前処理として、動作開始時の体の向きの違いによる影響を除くために正規化を行う。正規化は、頭部、腕部、胸部は腰の位置がつねに正面 (z 軸正方向) を向くように、脚部は腰の $x - z$ 平面での位置が原点で正面を向いた姿勢から基本動作が開始されるように、各部位の位置と角度を変換する。

図 3 に 1 つのセンサが記録した基本動作データの標準化および正規化後の位置成分 (x, y, z) と 3×3 の回転行列に変換した回転成分 ($m_{00} \sim m_{22}$) の例を示す。

3.3 基本動作の特徴ベクトル

本研究では、基本動作が短時間の単純な動作であると想定しているため、図 3 に示すように基本動作データの各成分の変化も単純であると考えられる。また、同じ種類の動作であれば、演じる速さにかかわらず互いに類似した動作として認識できるようにする。そのため、基本動作データの特徴ベクトルとして、基本動作データの各成分を等間隔に n 点でリサンプリングした各点の値を使用する。

表 1 各部位の特徴を得るために使用する関節部位
Table 1 Joint names used to obtain a feature of each part.

部位	使用する関節	ベクトル長
頭部	顎	$12n$
左・右腕部	左・右手首, 左・右肘, 左・右肩	$36n$
胸部	腰	$12n$
脚部	左・右足首, 左・右膝, 腰	$60n$

腕部は左右を別々に扱うが, 脚部は両脚を 1 つの部位とする.

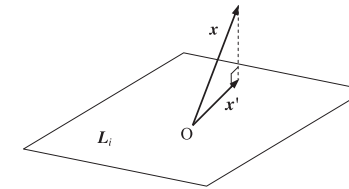


図 4 部分空間法

Fig. 4 Subspace method.

これにより 1 つのセンサで記録した動作の特徴は, 位置成分数を 3, 角度成分数を 9, リサンプリング点数を n として $(3+9)n = 12n$ 値で表すことができる. 各部位の特徴量を求めるために使用する関節名と特徴ベクトル長を表 1 に示す.

3.4 基本動作名と動作語対の作成

本論文では基本動作をある姿勢から他の姿勢への移り変わりであるとし, 基本動作の開始時, 終了時の姿勢の名前をそれぞれ動作開始姿勢名, 動作終了姿勢名と呼ぶ. また, 基本動作名を

動作開始姿勢名 → 動作終了姿勢名

と表記する. これらの姿勢名は, 姿勢を表す語句のほか, 動作に用いる部位, 動作の方向, 動作終了時に部位が到達する位置, 動作の種類, のいずれかを表す語句の 1 つ以上の組合せから構成されていると考える. 動作を表すために使われるこれらの語句を動作語と呼ぶ. 動作名は動作開始姿勢名, 動作終了姿勢名にそれぞれ 1 語以上の少なくとも 2 語の動作語から構成されていることとする.

基本動作名に含まれる動作語と基本動作データの特徴ベクトル成分との間には, 何らかの関連があると考えられる. たとえば基本動作名「気を付け → イスに座る」には動作開始姿勢名に「気を付け」の 1 語, 動作終了姿勢名に「イスに」と「座る」の 2 語の動作語が含まれる. 動作語「気を付け」と「イスに」の特徴は動作実行前後の腰の y 座標の値として, 動作語「座る」の特徴は動作中の腰の y 座標の値の変化として, 特徴ベクトルの成分中に現れる. したがって, 姿勢を表す動作語は基本動作の特徴ベクトルの特定の成分の値と, 動きを表す動作語は特徴ベクトルの特定の成分の値の変化と, それぞれ関連付けることができると考えられる.

動作語と基本動作の特徴ベクトルの成分を関連付けるために, 手作業により動作名を割り当てた基本動作から構成される学習データをあらかじめ用意する. その学習データ中の基本動作名に含まれるすべての動作語について, それぞれの動作語を含む基本動作の集合を作成

する. この集合は部分空間法で利用する.

実際の動作名では必ずしも 1 つの語句が 1 つの動作に対応せず, 同じ表記で異なる動作を指示する語句 (同表記異義語) が存在することがある. これは, 動作語の特徴抽出を阻害する要因となる. これを避け, 同表記異義語を区別するために, 同じ動作名で一緒に使われる他の語句との 2 語の組を作り, その組合せについて特徴ベクトルに共通する特徴を調べる. 以下, 基本動作名に含まれる動作語の 2 語の組合せを動作語対と呼ぶ. 各動作語対は互いに異なる動作の特徴を表すと仮定する.

前述の基本動作名の例には「気を付け」「イスに」「座る」の 3 語の動作語が含まれるので, 2 語の組合せを作成すると「気を付け, イスに」「気を付け, 座る」「イスに, 座る」の 3 組が動作語対となる.

3.5 部分空間法による識別辞書の作成

提案手法において使用する部分空間法について述べる. 図 4 に部分空間法によるパターン認識の概念図を示す. 図中の L_i はクラス w_i に属する学習データの特徴ベクトルの部分空間を, x は入力パターンの特徴ベクトルを表す. このとき, 部分空間法では特徴ベクトル x の部分空間 L_i への射影長 $|x'|$ を入力パターンとクラス w_i との類似度とする¹⁵⁾.

基本動作名に含まれる動作語は, その各々が基本動作の特徴を部分的に表している. そのため, ある動作語が異なる基本動作名に使われていても, その特徴は特徴ベクトル空間の特定の部分空間に分布していると考えられる. 部分空間法を用いることにより, 動作名に動作語対 w_i を含む学習基本動作の特徴ベクトルの部分空間 L_i を求め, 入力基本動作の特徴ベクトル x と動作語対 w_i との類似度を得ることができる.

現在, 部分空間法として様々に改良された手法が提案されているが, 本研究では, 主成分分析により学習パターンの部分空間を求める CLAFIC 法¹⁵⁾ を用いて, 基本動作と動作語対との関連を表す度合い (類似度) を求める. その手順を次に示す.

動作名に動作語対 w_j を含む基本動作の特徴ベクトルの集合を χ_j , その特徴ベクトルの数と次元数をそれぞれ n_j, d とする . ここで , 動作名に動作語対 w_j を含む基本動作の特徴ベクトル x の自己相関行列

$$R_j = \frac{1}{n_j} \sum_{x \in \chi_j} x x^t \quad (7)$$

の固有値のうち , 大きい方から k 番目の固有値 λ_{jk} に対応する固有ベクトルを u_{jk} とする . 大きい方から m 個の固有値を使用したときの u_{jk} の累積寄与率を

$$a(m) = \frac{\sum_{k=1}^m \lambda_{jk}}{\sum_{k=1}^d \lambda_{jk}} \quad (8)$$

とするとき , すべての動作語対に共通なパラメータ κ を設定し , 動作語対ごとに $a(d_i) \leq \kappa \leq a(d_i + 1)$ となる値 d_i を選択する .

このとき , 特徴ベクトルを x_i とする基本動作と動作語対 w_j との類似度 $S_j(x_i)$ は次式により求められる .

$$S_j(x_i) = \sum_{k=1}^{d_j} (x_i^t u_{jk})^2 \quad (9)$$

データベースに登録する基本動作の特徴ベクトルを $x_1 \dots x_p$, 学習データの基本動作の動作名に含まれる動作語対を $w_1 \dots w_q$ とするとき , 登録するすべての基本動作について動作語対との類似度を計算することにより作成した行列

$$S = \begin{pmatrix} S_1(x_1) & \dots & S_q(x_1) \\ \vdots & \ddots & \vdots \\ S_1(x_p) & \dots & S_q(x_p) \end{pmatrix} \quad (10)$$

を識別辞書とする . この行列の (i, j) 要素は特徴ベクトルを x_i とする基本動作と動作語対 w_j との類似度を表している .

特徴ベクトルを x_i とする新しい基本動作をデータベースに登録するときには , その基本動作と動作語対 $w_1 \dots w_j \dots w_q$ との間で類似度 $S_j(x_i)$ を計算して識別辞書に追加する .

4. 動作語を検索クエリとする基本動作の検索

4.1 動作語対と基本動作の合致度の計算

識別辞書を調べることにより , ある動作語対の特徴を多く持っている基本動作はどれかを知ることができるので , 動作語を検索クエリとして基本動作を検索するために識別辞書を利用できる . 検索結果は検索クエリが表す基本動作に合致している度合い (合致度) の高い順にデータベース中の基本動作データを整列した順位として与えられる . その手順を次に述べる .

はじめに , 検索クエリとして与えられた動作語中の 2 語を組み合わせ , 動作語対を作成する . このとき , 式 (10) で求めた識別辞書 S 中の動作語対に対応する列の類似度の合計が大きい基本動作が , 検索対象の基本動作に類似したものとなる .

ただし , 本研究で作成した識別辞書では , よく使われる動作語対よりもあまり使われない動作語対の方が高い類似度を持つ傾向があるため , 類似度の合計を求める際に , 学習データから求めた動作語対の出現頻度による重み付けを行う . 動作語対の集合を W , 集合中の動作語対の数を $|W|$, 動作語対 $w_j (\in W)$ の出現頻度を $P(w_j)$, 識別辞書 S 中の特徴ベクトルを x_i とする基本動作と動作語対 w_j との類似度を S_{ij} とするとき , W と特徴ベクトルを x_i とする基本動作との合致度 Q_i を次式で定義する .

$$Q_i = \frac{1}{|W|} \sum_{w_j \in W} -\log P(w_j) \cdot S_{ij} \quad (11)$$

この Q_i を識別辞書中のすべての基本動作について計算する . Q_i の値の大きい基本動作が検索クエリの特徴をよく表すものとなるため , Q_i の値の降順に並べ替えた基本動作を検索結果とする .

4.2 基本動作の自動生成と類似検索⁶⁾

これまでの説明では , 基本動作の分割があらかじめ行われていることを前提として , 動作語を用いて基本動作を検索する手法について述べてきた . しかし , 実用化のためには , 動作データを分割して基本動作を作成するコストが障害となるため , 基本動作への分割も含めて自動化できることが望ましい .

次に , 動作データを自動的に分割して基本動作の候補を作成する手法 , および式 (11) の合致度を用いた , 動作語をクエリとする基本動作の検索手法について述べる . これまでは動作語をクエリとした検索を行うために , 手作業で動作名を割り当てることにより基本動作を

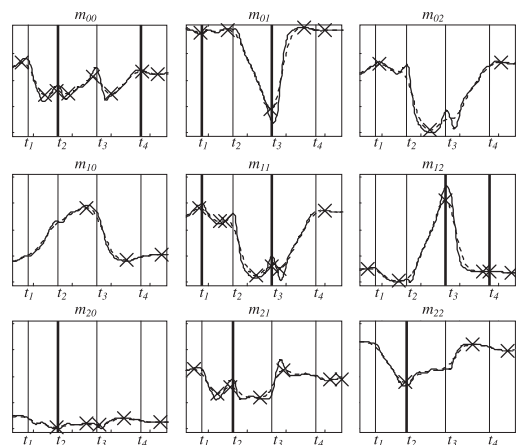


図5 手作業による基本動作分割時刻の特徴
Fig. 5 The feature of the time which divides basic motion manually.

作成していた。本手法により、自動的に分割した基本動作についても同様に、動作語をクエリとした検索が可能になるため、基本動作データベース構築の大幅な省力化が達成できる。

4.2.1 基本動作候補の作成

動作データが、手作業によりどのような時刻で基本動作に分割されているのかを調査した。図5に手作業で作成した右腕の基本動作の分割時刻（図中縦線）と、右手首の回転成分を回転行列に変換した9つの要素（図中実線）を示す。図中の破線は平滑化した回転行列の成分を、×印はその成分の極値をそれぞれ表している。太い縦線は、手作業により分割された時刻のうち、極値をとる時刻と一致する時刻を表している。この例には全部で4つの分割時刻（ $t_1 \sim t_4$ ）があるが、 t_1 は m_{01} 、 m_{11} で、 t_2 は m_{00} 、 m_{20} 、 m_{21} 、 m_{22} で、 t_3 は m_{01} 、 m_{11} 、 m_{12} で、 t_4 は m_{00} 、 m_{12} でそれぞれ極値をとる時刻の1つと一致していることが分かる。

この結果から、着目する部位の回転成分を 3×3 の回転行列に変換し、その9つの要素を平滑化した後に極値を求めると、その極値をとる時刻のいずれかによって、手作業で分割した時刻のすべてをカバーできることが分かった⁶⁾。

この特徴を用いて基本動作の候補を作成する手順を次に示す。

- (1) 動作データの回転成分から変換した 3×3 の回転行列の各要素を平滑化した後に、極

値をとる時刻を求め、集合 B とする。

- (2) 集合 B 中の各要素から重複する値を取り除く。また、互いの時刻が閾値 t_{min} より近接している2つの要素はそれらの平均値と置き換える。
- (3) 集合 B 中から取り出した任意の2つの要素をそれぞれ b_i 、 b_j とする。ただし、事前に定めた基本動作の最小幅と最大幅をそれぞれ f_{min} 、 f_{max} とするとき、 b_i 、 b_j は $b_i < b_j$ かつ $f_{min} \leq b_j - b_i \leq f_{max}$ を満たすものとする。このとき、 b_i 、 b_j がそれぞれ開始時刻と終了時刻になるようなすべての区間の集合を作成し、基本動作の候補 C とする。

作成されたそれぞれの候補に対して、学習データの動作名に含まれる動作語との間で式(9)の類似度を計算し、識別辞書 S を作成する。

4.2.2 基本動作の候補からの目的動作の類似検索

上記の手順で作成した基本動作の候補 c_i ($\in C$)と識別辞書 S から、動作語を検索クエリとして基本動作の類似検索を行う手順について述べる。

はじめに、検索クエリとして与えられた動作語中の2語を組み合わせ、動作語対 W を作成する。次に、動作語対を w_j ($\in W$)、学習データ中での w_j の出現頻度を $P(w_j)$ 、4.2.1項で求めた識別辞書 S 中の基本動作の候補 c_i と動作語対 w_j との類似度を S_{ij} として、すべての基本動作の候補について式(11)により w_j との間の合致度を計算する。この結果の中で、大きな Q_i の値を持つ基本動作の候補が検索目的の基本動作に合致した動作になるため、合致度 Q_i の降順に並べ替えた基本動作の候補を検索結果とする。

5. 実験と評価

5.1 基本動作のテストデータ

動作語対と基本動作の特徴ベクトルとの類似度の計算や、評価実験において検索結果が正解かどうかを判定するためには、動作名が付けられた基本動作データが必要である。現在、動作検索の評価の目的に利用可能な身体動作データが存在しないため、検索手法の評価に利用可能な基本動作のテストデータを作成した。

提案手法が対象とする身体動作は、短かいまとまりのある動作（基本動作）に分割可能であること、および基本動作に名前を付けられることが必要である。これらの条件に適合する動作として空手道の形（かた）を選び、形に含まれる空手道の技を基本動作に、その技の名前を動作名として基本動作のテストデータを作成した。このテストデータは、3人（A、B、C）が演じたそれぞれ4種類の形（ $K_1 \sim K_4$ ）の動作データからなる。各データは形の動作

表 2 テストデータ諸元
Table 2 Specifications of test data.

形	部位	基本動作種類	基本動作数	平均基本動作長 (フレーム)
K_1	右腕	19	96	41.4
	左腕	21	102	42.0
	脚	10	42	32.4
K_2	右腕	25	78	21.8
	左腕	23	78	23.5
	脚	23	69	20.7
K_3	右腕	29	90	23.0
	左腕	29	90	23.3
	脚	24	75	22.0
K_4	右腕	26	114	28.2
	左腕	27	117	32.7
	脚	8	30	27.5

表 3 基本動作名および動作語の例
Table 3 Examples of basic motion name and words.

(a) Examples of basic motion name

部位	基本動作名
右腕・左腕	引き手 → 正拳中段突き
	引き手 → 中段横受け
	正拳中段突き → 中段横受け
	中段横受け → 引き手
脚	下段払い受け → 引き手
	内八文字立ち → 平行立ち
	左三戦立ち → 前進右三戦立ち
	右三戦立ち → 後退左三戦立ち

(b) Examples of basic motion words

部位	動作語
右腕・左腕	引き手, 正拳, 中段, 下段, 突き, 受け, 払い, 横
脚	左, 右, 内, 八文字立ち, 平行立ち, 三戦立ち, 前進, 後退

についての知識を持つ者により部位別に基本動作への分割および名前付けが行われた。表 2 にテストデータの諸元を示す。また、表 3 (a) にテストデータに含まれる基本動作名の例を、(b) にその基本動作名に含まれる動作語を示す。

5.2 実験方法

提案手法の有効性を検証するために、作成した基本動作のテストデータを使用した検索実験を行った。実験に使用した機器は Apple MacPro (CPU: Intel Xeon X5482 3.2 GHz, メモリ: 2 Gbyte, OS: Windows XP Service Pack 2) である。検索実験プログラムは MATLAB R2008a を用いて作成した。

実験は、提案手法との比較のためのベースライン手法、手作業で分割した基本動作を対象とした検索 (以下、提案基本手法)、自動的に作成した基本動作の候補を対象とした検索 (以下、基本動作自動分割手法) の 3 通りについて行う。各実験は、テストデータを動作の演技者により 3 グループに分け、3 分割の交差検証法により行った。すなわち、2 グループを学習データとして識別辞書の作成に、残りの 1 グループを実験データとして基本動作の検索に使用する実験を分割ごとに繰り返し行い、各分割の評価値の平均を結果とした。

提案手法との比較のために文献 7) で提案されている手法をベースラインとした。この手法ではまず、学習基本動作データの動作名に含まれるすべての動作語について、その動作語を使用する基本動作の特徴ベクトルを集め、3.3 節で述べた特徴ベクトルの各要素の分散と平均を計算する (ただし 3.2 節で述べた正規化は本論文で新たに提案する処理であり、ベースライン法では正規化は行っておらず、標準化のみしか行っていない)。次に、分散が閾値以下の要素の平均値を、その動作語の特徴を表す特徴ベクトルとする。動作語とその特徴ベクトルとの対応は識別辞書に保存する。

動作語を検索クエリとして基本動作を検索するためには、検索クエリから作成した各動作語の特徴ベクトルを識別辞書から取り出し、各要素の値の平均を計算する。その結果が検索対象の基本動作の特徴ベクトルを表すものであるとして、データベース中のすべての基本動作の特徴ベクトルとの間のユークリッド距離を求め、距離が近い順に整列したものを検索結果とする。

本論文では、検索結果の向上を目的とした要素技術として、動作データの正規化、動作語の 2 語の組合せ、合致度の計算における重み付けを提案した。それぞれの要素技術の効果を評価するため、ベースライン手法と提案基本手法では、各要素技術 (重み付けは提案基本手法のみ) の使用の有無について検索実験を行った。基本動作自動分割手法では、すべての要素技術を使用して実験を行った。

検索実験の対象とする基本動作は次の条件により選択した。

- 基本動作長が 10 フレーム以上であること。
- 学習データの基本動作名に 5 回以上出現する動作語の組合せにより、動作名を作ること

表 4 実験データ諸元
Table 4 Specifications of experimental examples.

部位	基本動作の総数	検索に使用した基本動作数	検索すべき基本動作数の平均
右腕	67	9	3.8
左腕	88	8	4.1
脚	68	4	4.5

ができること。

- 腕の動作については、本研究で使用したモーションキャプチャ装置では区別が不可能な手指の形のみが異なるような基本動作を除く。

実験データの諸元を表 4 に示す。表中の基本動作の総数は各部位の基本動作の数、検索に使用した基本動作数は各部位で検索実験に使用する基本動作の数、そして検索すべき基本動作の平均数は 1 つの検索基本動作あたりの正解として検索すべき基本動作数の平均をそれぞれ表している。

検索実験は、動作語に分割した基本動作名を検索クエリとして行った。検索結果は、検索クエリが表す基本動作に類似している順に整列した基本動作が出力されるが、ベースライン手法と提案基本手法では、検索クエリが表す基本動作名を持つ基本動作を正解とした。基本動作自動分割手法では、手作業により分割した実験データの基本動作から、検索結果の基本動作の候補と時間軸で重複する部分が最も多い基本動作を求め、その基本動作名が検索クエリの基本動作名と等しい場合を正解とした。

なお、識別辞書の作成に使用する基本動作データの各成分のリサンプリング点数 n 、および類似度の計算に使用するパラメータ κ は、予備実験で最も良い検索結果を示した $n = 9$ 、 $\kappa = 0.9$ に定めた。基本動作の候補を作成する際に使用するパラメータ t_{min} 、 f_{min} 、 f_{max} はそれぞれ $t_{min} = 5$ 、 $f_{min} = 18$ 、 $f_{max} = 180$ に定めた。

5.3 実験結果

提案手法による基本動作の類似検索の性能を評価するため、ランキング検索の評価に広く用いられている次の 2 つの指標を用いる²³⁾。

平均精度 (Average Precision) なるべく多くの正解を検索したい、という検索システムの利用を前提とした評価指標である²³⁾。正解基本動作数を D 、検索結果の第 i 位の基本動作が正解のときに 1、不正解のときに 0 である値を r_i とする。また第 k 位までの検索結果に含まれる正解の割合を

$$P_k = \frac{1}{k} \sum_{i=1}^k r_i \quad (12)$$

とする。このとき、 P_k と最も下位に見つかった正解の順位 N から、平均精度 AP は式

$$AP = \frac{1}{|D|} \sum_{k=1}^N r_k \times P_k \quad (13)$$

により定義される。

逆数順位 (Reciprocal Rank) 正解が 1 件見つければよい、という検索システムの利用を前提とした評価指標である²³⁾。検索結果が 1 つも正解を含まない場合は 0、そうでない場合、検索結果中の最も上位にある正解の順位を r' としたとき、逆数順位 RR は式

$$RR = \frac{1}{r'} \quad (14)$$

により定義される。

基本動作自動分割手法の評価では、これらの指標のほかに、手動で分割した基本動作と検索結果の基本動作が一致する度合いを評価するために次の指標を用いる。

一致度 2 つの区間が完全に一致しているときに 1、ある区間の終点とその次の区間の始点だけが重なっているときに 0、2 つの区間に重なり部分がないときに負の値をとる評価値である。手動で分割した基本動作の区間を $[a_1, a_2]$ 、自動分割された基本動作の区間を $[b_1, b_2]$ とするとき、式

$$C = \frac{2(\min(a_2, b_2) - \max(a_1, b_1))}{a_2 - a_1 + b_2 - b_1} \quad (15)$$

により定義される。

手作業により分割した基本動作データを用いたベースライン手法と提案基本手法による検索実験結果を表 5 と表 6 に、基本動作自動分割手法による検索実験結果を表 7 にそれぞれ示す。

ベースライン手法では、2 語の動作語を組み合わせることによる曖昧性の除去が、提案基本手法ではデータの正規化が、それぞれ検索性能の向上に最も貢献していることが分かった。提案基本手法では、重み付けが 2.0%、2 語の動作語の組合せが 8.3%、データの正規化が 14.2%、それぞれ AP の値を向上させている。また、すべての要素技術の組み合わせることにより 21.8% の向上が見られた。脚部の基本動作は腕部よりも悪い結果を示している

表 5 ベースライン手法による検索実験結果

Table 5 Experimental results of a baseline method using manually divided basic motions.

手法	部位	AP(%)	RR
ベースライン手法	右腕	23.2	0.54
	左腕	28.1	0.64
	脚	13.7	0.33
	平均	23.2	0.54
ベースライン手法 +2 語の組合せ	右腕	38.0	0.87
	左腕	34.6	0.78
	脚	23.7	0.62
	平均	32.1	0.76
ベースライン手法 +データの正規化	右腕	23.9	0.52
	左腕	21.3	0.50
	脚	12.8	0.27
	平均	19.3	0.43
ベースライン手法 +データの正規化 +2 語の組合せ	右腕	35.5	0.81
	左腕	32.3	0.74
	脚	22.2	0.52
	平均	30.0	0.69

が、この原因として平面的な動きの基本動作が実験に使われたため、特徴ベクトルに動作の特徴が現れにくいことが考えられる。

すべての要素技術を組み合わせた提案基本手法では、ベースライン手法と比較して AP で 50.8% 増の性能向上を達成しており、RR の値も 0.91 と正解が検索結果の上位に検索されていることを示している。なお、検索に要する時間は、1 つの基本動作あたり平均 0.002 秒であった。

基本動作への分割を自動で行っている基本動作自動分割手法では、AP の値がすべての要素技術を組み合わせた提案基本手法の検索結果より 3 割程度劣るものの、検索結果の上位に正解が集まっていることを示している。RR の値は平均では 0.72 であり、正解が上位に検索されていることを示しているが、特に左腕では 1.00 と目標の基本動作がすべて最初に検索されていることが分かった。一致度は平均 0.64 であり、検索結果である自動分割された基本動作の分割範囲が手動で分割した基本動作の範囲とおおむね一致していることを示している。腕部に比べ脚部の結果が悪い理由として、提案基本手法の場合と同様、特徴を表しにくい基本動作が使われたことが考えられる。

なお、基本動作自動分割手法では、生成される基本動作の候補数が平均約 6,100 個と多く、1 つの基本動作の検索時間が平均 1.2 秒であった。検索時間の短縮のために基本動作の

表 6 提案基本手法による検索実験結果

Table 6 Experimental results of a proposed method using manually divided basic motions.

手法	部位	AP(%)	RR
提案基本手法	右腕	62.9	0.72
	左腕	86.1	0.94
	脚	16.3	0.27
	平均	59.0	0.66
提案基本手法 +重み付け	右腕	65.6	0.78
	左腕	89.3	1.00
	脚	28.1	0.32
	平均	61.0	0.70
提案基本手法 +2 語の組合せ	右腕	78.5	0.96
	左腕	89.4	1.00
	脚	34.0	0.41
	平均	67.3	0.79
提案基本手法 +データの正規化	右腕	87.2	0.92
	左腕	89.5	0.98
	脚	42.8	0.50
	平均	73.2	0.80
提案基本手法 +重み付け +2 語の組合せ +データの正規化	右腕	93.2	0.96
	左腕	92.7	1.00
	脚	57.5	0.76
	平均	80.8	0.91

表 7 基本動作自動分割手法による検索実験結果

Table 7 Experimental results of a proposed method using automatically divided basic motions.

部位	候補数	AP(%)	RR	一致度
右腕	7,584.7	73.2	1.00	0.70
左腕	7,750.3	73.9	1.00	0.65
脚	3,057.0	23.7	0.13	0.57
平均	6,130.7	57.0	0.72	0.64

候補数を削減することが、今後の検討課題の 1 つである。

6. おわりに

本研究では、身体動作データベースの効率的な構築と利用の容易化のために、動作データと動作名に含まれる語句との関連を定量化する手法を提案した。提案手法により、対象分野の動作についての学習データを用意することで、基本動作データベースの自動インデクシ

グ、動作語を検索クエリとする基本動作データの検索、および動作データの特徴により作成した基本動作の候補から、動作語により適切な基本動作を検索することが可能になる。

提案手法を評価するために作成したテストデータを使用した実験により、提案手法が基本動作の自動生成、および動作語による基本動作の類似検索のための有効な手法であることが分かった。

今後の課題として、提案手法を舞踊や体操など他の種類の身体動作へ適用すること、検索性能を向上させるための特徴量および特徴抽出手法の改良、および自動的に作成される基本動作の候補数の削減があげられる。

謝辞 本研究の一部は、平成 17～19 年度総務省戦略的情報通信研究開発制度 (SCOPE)、および平成 18, 19 年度科学研究費補助金基盤研究 (C) (18500071) の支援によって実施した。また、動作データの取得に協力していただいた NPO 法人日本伝剛柔流空手道巧志会、および (株) わらび座デジタルアートファクトリの皆様に感謝の意を表す。

参 考 文 献

- 1) 湯川 崇, 海賀孝明, 長瀬一男, 玉本英夫: 舞踊符による身体動作記述システム, 情報処理学会論文誌, Vol.41, No.10, pp.2873–2880 (2000).
- 2) Müller, M., Röder, T. and Clausen, M.: Efficient Content-Based Retrieval of Motion Capture Data, *ACM Trans. Graph.*, Vol.24, No.3, pp.677–685 (2005).
- 3) 玉本英夫, 湯川 崇, 海賀孝明, 水戸部一孝, 三浦 武, 吉村 昇: 産学官連携による民俗芸能伝承のためのデジタルコンテンツ制作技術の開発, 電子情報通信学会誌, Vol.91, No.4, pp.303–308 (2008).
- 4) 湯川 崇, 小原直子, 玉本英夫: モーションキャプチャデータからの基本動作分割点の自動検出, 情報処理学会論文誌, Vol.45, No.4, pp.1198–1201 (2004).
- 5) 柴田 傑, 湯川 崇, 海賀孝明, 門脇さくら, 横山洋之, 玉本英夫: 舞踊符自動生成のための自己組織化マップによる動作の分割化, 秋田大学工学資源学部研究報告, Vol.28, pp.17–24 (2007).
- 6) 湯川 崇, 海賀孝明, 玉本英夫: モーションキャプチャデータの基本動作データへの自動分割法, 情報処理学会東北支部研究会 (山形大学), 07-A3-4 (2008).
- 7) Yukawa, T., Obara, N. and Tamamoto, H.: Automatic Construction of the Motion Database which Allows to Search Contents by a Motion Name, *Proc. IAPR Conference on Machine Vision Applications (MVA 2005)*, pp.534–537 (2005).
- 8) Forbes, K. and Fiume, E.: An Efficient Search Algorithm for Motion Data Using Weighted PCA, *Proc. 2005 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, pp.67–76 (2005).
- 9) 小杉尚子, 小島 明, 片岡良治, 串間和彦: 大規模音楽データベースのハミング検索

システム, 情報処理学会論文誌, Vol.43, No.2, pp.287–298 (2002).

- 10) 片山紀生, 佐藤真一: マルチメディア時代の類似検索のための索引技術, 情報処理, Vol.42, No.10, pp.958–964 (2001).
- 11) Lv, Q., Josephson, W., Wang, Z., Charikar, M. and Li, K.: Ferret: A toolkit for content-based similarity search of feature-rich data, *EuroSys '06: Proc. 1st ACM SIGOPS/EuroSys European Conference on Computer Systems 2006*, New York, NY, USA, ACM, pp.317–330 (2006).
- 12) Rafei, D. and Mendelzon, A.: Efficient retrieval of similar time sequences using DFT, *Proc. 5th FODO 1998*, pp.249–257 (1998).
- 13) Chan, K.-P. and Fu, W.-C.: Efficient time series matching by wavelets, *Proc. 15th International Conference on Data Engineering*, pp.126–133 (1999).
- 14) Keogh, E.J., Palpanas, T., Zordan, V.B., Gunopulos, D. and Cardle, M.: Indexing Large Human-Motion Databases, *VLDB*, pp.780–791 (2004).
- 15) 石井健一郎, 上田修功, 前田英作, 村瀬 洋: わかりやすいパターン認識, オーム社 (1998).
- 16) 福井和広: 部分空間表現に基づく顔画像認識, 電子情報学会技術研究報告 PMRU, パターン認識・メディア理解, Vol.102, No.651, pp.33–40 (2003).
- 17) 小坂谷達夫, 山口 修, 福井和広: 制約相互部分空間法を用いた顔認識システムの開発と評価, 情報処理学会論文誌, Vol.45, No.3, pp.951–959 (2004).
- 18) 福井和広: 部分空間法の今昔 (下): 最近の技術動向: 相互部分空間法への拡張とその応用事例, 情報処理, Vol.49, No.6, pp.680–685 (2008).
- 19) 西山正志, 山口 修, 福井和広: 制約相互部分空間法を用いたジェスチャー認識, 第 10 回画像センシングシンポジウム講演論文集 SSI104, pp.439–444 (2004).
- 20) O'Brien, J.F., Bodenheimer, B., Brostow, G.J. and Hodgins, J.K.: Automatic joint parameter estimation from magnetic motion capture data, *Graphics Interface*, pp.53–60 (2000).
- 21) 島田静雄: CAD・CG のための基礎数学, 共立出版 (2000).
- 22) ISO/IEC 19774:200x: Information technology – Computer graphics and image processing – Humanoid Animation (H-Anim).
- 23) 酒井哲也: よりよい検索システム実現のために, 情報処理, Vol.47, No.2, pp.147–158 (2006).

(平成 20 年 6 月 6 日受付)

(平成 20 年 12 月 5 日採録)



湯川 崇 (正会員)

昭和 43 年生。平成 8 年秋田大学大学院鉱山学研究科博士後期課程退学。同年秋田経済法科大学短期大学部講師。平成 12 年同助教授。平成 19 年ノースアジア大学教養部准教授。平成 20 年同法学部准教授。コンピュータグラフィクス、身体動作解析の研究に従事。秋田大学大学院工学資源学研究科博士後期課程在学中。



玉本 英夫 (正会員)

昭和 24 年生。昭和 51 年東京大学大学院博士課程修了。同年秋田大学鉱山学部電子工学科講師。昭和 55 年同学科助教授。平成 3 年同情報工学科助教授。平成 9 年同工学資源学部情報工学科教授。現在に至る。この間、論理回路の故障診断、画像計測、マルチメディア等の研究に従事。工学博士。電子情報通信学会、人工知能学会、計測自動制御学会、IEEE 各会員。