

4. ジェスチャー認識を利用した移動ロボットとのコミュニケーション手段の提案 — 誤認識リスクを低減するための設計手法 — *

呂 健**, 姜 偉***

4. Proposal of Communication Method of Mobile Robots by Recognizing Gestures — Design Methods Dedicated to Reduce the Misrecognition Risk —

by Jian LU** and Wei JIANG***

Abstract; As the robot technology advances, in stead of keyboards and control panels, which were previously used as the main communication method with a robot, new communication methods which uses voice recognition and gesture recognition are expected to become common as the robot will work cooperatively with human being and will be widely involved in people's daily life. That is to say, robot will act according to instructions from these kinds of interfaces, therefore the interest of our research is focused on risk assessment and risk reduction related to these new interfaces while they are used for robot control, to assure robot safety. Particularly in our research here, the type of the new interface is restricted to gesture recognition, and the research focus is on reducing the gesture misrecognition. As the first step for assessing and evaluating the different design selections according to the misrecognition risk for gestures, two models were established in the previous stage of our research. The first model was a three dimensional (3D) model for quantitatively expressing gestures, and second model was a quantity analysis model, called gesture distance model (GDM), to measure the difference between different gestures. The examples of applying these models with the average body-size data in the Japanese Body-size Database were given, and the basic data about standard gesture features of average body-size for Japanese people were also obtained. These results were reported previously in the reference¹⁾. In addition, the applications of these results to reduce misrecognition were briefly discussed there.

Here in this report, design considerations are discussed about how to reduce the gesture misrecognition during the entire design work for a gesture recognition system. The design work is divided into three steps: basic design, function design and parameter design. In additional to the two models previously established, two principles are introduced, i.e., simplicity principle and distance principle. Based on these model and principles, the design method and policy are discussed individually for the three steps during the design work. In the basic design step, simplicity principle is mainly used. On the other hand, the distance principle, the 3D model and GDM model are mainly used in the function design steps and parameter design step. The results and the importance of our research are that the quantitative models and methods are provided for evaluating and comparing different design selections for implementing a gesture recognition system. Based on these results, we hope to provide a prototype to establish risk assessment and safety standards for the gesture recognition system to be used in industrial human-robot interfaces.

Keywords; Human-machine Communication, Pattern Recognition, Gesture Recognition, Robot Safety, Risk Assessment

* 本研究の一部は、SIAS2007 (5th International Conference on Safety of Industrial Automated Systems, Tokyo, November 12-13,2007) に発表済。

** 人間工学・リスク管理研究グループ Human Engineering and Risk Management Research Group

*** 科学技術振興機構 元重点研究支援協力員 Japan Science and Technology Corporation, Research Fellow

1. はじめに

従来、産業用ロボットは主に固定された場所で稼働してきたが、これからは人間とコミュニケーションを行い、自律的移動機能も持つタイプが主流になるとと思われる。このような新型産業用ロボットと人間のコミュニケーション手段として、音声認識やジェスチャー認識など新技術の応用が期待されている。そこでジェスチャー認識技術の応用に伴う安全制御について研究を行った結果について報告する。

ジェスチャー認識技術はパターン認識手法の一種の応用であるため、パターン認識手法の他の応用と同様、誤認識の可能性が存在する。そのため、産業用ロボットへ適用するとき、「誤認識」による誤動作を引き起こすリスクがある (Fig. 1)。従って、ジェスチャー認識技術を産業用ロボット制御へ適用する前に、1) 誤認識リスクに基づくジェスチャー認識手法の評価、2) リスク低減のためのジェスチャーシステム的设计・実現手法についての議論が必要である。現状では、誤認識リスクを評価する汎用的、標準的な手段がない。そこで、本研究は誤認識リスクを減少させるための設計手法の開発を目的とした。

前号の報告¹⁾では、開発されたジェスチャーの数値化モデル、ジェスチャー距離モデルの説明を行った。そして、これらモデル及びそのリスク評価への応用の具体例として、日本人身体特徴データベースへの適用、さらに標準ジェスチャーを異なる身長・性別グループのジェスチャー認識へ適用する際のリスクを検討した。

本論文では、画像のパターン認識システムとして実現されるジェスチャー認識システムについて、設計手順を整理し、誤認識リスク低減のための一般的な手法を設計の3段階に分けて述べる。

2. ジェスチャー及び画像のパターン認識

2.1 画像のパターン認識

パターン認識に使う入力情報のタイプが画像のみの場合、パターン認識は画像認識という^{3, 4)}。一般的パターン認識課題と同じように、画像認識のプロセスはFig. 2に示すように、入力画像の特徴抽出を行ってから、標準パターンとのマッチングを行う。画像に関する特徴には、カラー、濃度及び抽出された点(ポイント)、ライン、円などの特徴を良く使う。

画像認識手法として、主にテンプレートマッチング法、シンタクティック方法、及び統計方法の3種類がある。テンプレートマッチング法は、二進値テンプレート化した形状を目標物の特徴として、パターン学習照合を行う方法である。生産ラインの製品検査システムなど対象物の形状が固定している場合によく使われる。シンタクティック方法は目標物の構造的な解析を行い、その構成規則を対象物の特徴として、パターン学習照合を行う方法である。この方法は自動翻訳や文字認識などの場合に良く使われる。

統計方法は数値で表現した特徴を用いて、対象物の特徴値空間で分類・認識を行う方法である。対象物の形状が複雑、構造の記述が困難な場合には、通常、統計方法を選択し、画素の濃淡・カラーなど最も基本的な特徴を利用する。

本研究で、ジェスチャーの検出を行う場合において、目標物とする人体画像の形状が複雑な不定形物であり、テンプレートマッチング法やシンタクティック法での記述が困難であるため、統計方法でジェスチャーの検出を行うことにした。

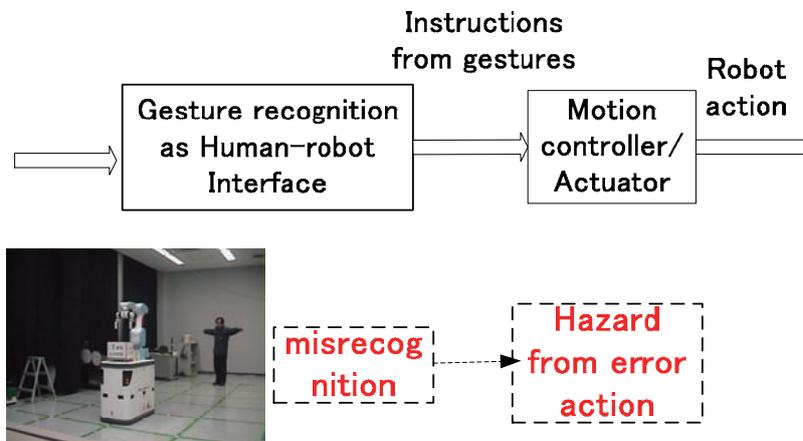


Fig. 1 The misrecognition for gestures and the hazard.
ジェスチャーの誤認識と災害

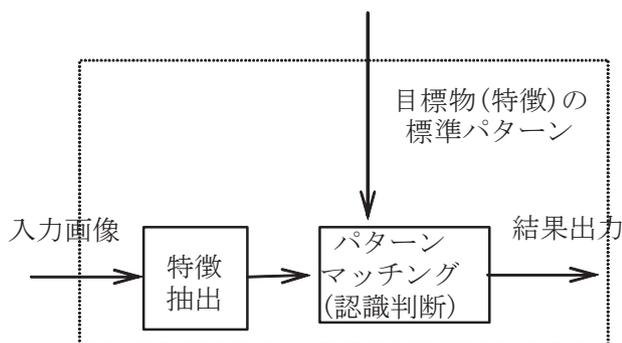


Fig. 2 Image pattern recognition.
画像のパターン認識

2.2 ジェスチャー認識システムの設計選択

Fig. 2で示したように、ジェスチャー認識システムが実現するには、画像認識の手法により人間から示されたジェスチャーをシステム的设计仕様にある標準ジェスチャーの特徴と照らし、どのジェスチャーに属するかを認識・判断する。そのため、次に示すことについて設計選択を行う必要がある。

- 1) 特徴抽出部がジェスチャーのどの特徴を抽出するかについて決定・選択する。つまり下記ジェスチャーの基本仕様について決定・選択する。
 - a) ジェスチャー指示部の決定：腕，手，指など人体のどの部位でジェスチャーを示すのか？又は全ての人体部位を用いてジェスチャーを示すのか？
 - b) ジェスチャーは指示部の動きを持った動作で示されるのか？それとも静止的動作で示されるのか？
 - c) 指示部の画像からどんな特徴抽出を行うのか？例えば、ポーズの幾何学形状を抽出するのか？又は指示部の特別な部分（ラインまたはポイント）を抽出するのか？
- 2) ジェスチャーの意味（ロボットのどのような制御を行うのか？）。
- 3) 認識部はどんな手法により判断を行うのか？などについて設計選択を行なう必要がある。

2.3 三つの設計ステップ

ジェスチャー認識システムの設計段階において、ジェスチャーの仕様設計、特徴選択、ジェスチャーの認識（分類）の方法選択などについて決定する必要がある、これらについての決定・選択は、ジェスチャーの誤認識リスクに影響する。そこで、第3節において、どんなに設計・選択をすれば、誤認識リスクの減少に繋がるかについて議論する。まず、設計段階を3つのステップに分ける。

1) 基本設計ステップ

このステップでは、ジェスチャー認識システムの構造

やジェスチャーの基本仕様について設定する。すなわち前記1)のa) - c)について選択する。

2) 機能設計ステップ

このステップは2.2節で示した2)に相当し、どのようなロボット制御内容をどのようなジェスチャーに対応させるのかについて設計する。

3) パラメーター設計ステップ

このステップは3)に相当し、ジェスチャー認識の判断に使う閾値パラメーターを設計選択する。

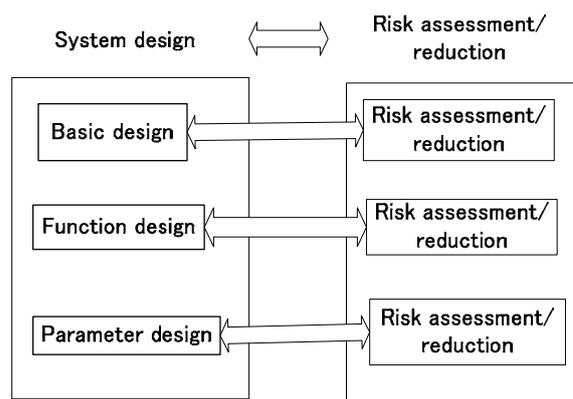


Fig. 3 Three steps in design stage and risk reduction.
設計段階の3つのステップとリスク減少

3. 誤認識を減少するための設計

3.1 基本的な考え方

ジェスチャー認識は統計分類を用いて、ジェスチャーのパターン認識を行う。特徴空間を複数の重ならない領域に分割し、認識しようとするジェスチャーが特定の領域へ属するようにする。そのため、認識するジェスチャーの数が多くなるほど、領域が重なりやすくなり、誤認識リスクが高くなる。従って、システム的设计上では、多数のジェスチャーを定義しないこと（以下、簡単化原則と呼ぶ）が必要である。

また、異なるジェスチャーの特徴空間に対応する領域が大きく離れるほど、誤認識リスクが少ない（以下、距離原則と呼ぶ）。そこで、前研究¹⁾では、異なるジェスチャーの距離原則に基く離れ程度を測るため、3次元空間におけるジェスチャーの数値表示モデルと、それに基づくジェスチャー間の差の数値化（ジェスチャー間の距離モデル）を提案した。

本論文では、ジェスチャー認識システムの全設計過程をFig. 3のように、基本設計、機能設計、パラメーター設計の3段階に分けて、前記簡単化原則と距離原則を応

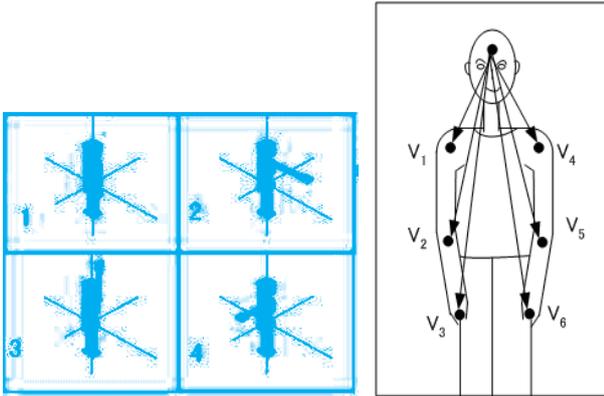


Fig. 4 Basic gestures and digital model.
基本的ジェスチャーとその特徴点による表示

用したリスク減少策について述べる。すなわち基本設計段階では、簡単化原則の応用、ジェスチャーの3D数値化モデル及びジェスチャー間距離の3Dモデルを説明する。次にこれらモデルに基く距離原則を機能設計段階とパラメーター設計段階へ適用した事例について述べる。

3.2 基本設計

基本設計段階では、システムの基本要素を決める設計選択を行う。例えば、認識できるジェスチャーは腕によるジェスチャーか、手によるジェスチャーか、又は指によるものかについて、選択する必要がある。具体的に、本研究では簡単化原則に従い、腕による静止ジェスチャーを選択した (Fig. 4左)。選択した理由は静止ジェスチャーは動的ジェスチャーより不確定性が小さく、腕ジェスチャーは手あるいは指のジェスチャーより識別しやすいことによる。また、ジェスチャーの識別特徴をポイント、ライン、又は2次元図形のどちらを選択するかについて、ポイント特徴の汎用性と簡単性により、Fig. 4右のようにポイントに基く特徴、すなわち特徴点を選択した (Fig. 4右図参照)。片腕は上、下、前及び側面へ腕を伸ばす4種の基本ジェスチャーがあり、両腕における組み合わせは16種のジェスチャーとなる (Fig. 5参照)。

3.3 ジェスチャーの数値モデルとジェスチャーの距離

ジェスチャー認識については関連する研究やその成果としての応用も出始めているが、現状ではジェスチャーの設計手法と認識手法が系統的に確立されていない⁶⁾。ジェスチャーの誤認識リスクを定量的に解析する第1段階として、ここでは、静止したものを対象にして、ジェスチャーを腕による右肩、右肘、右手首、左肩、左肘、

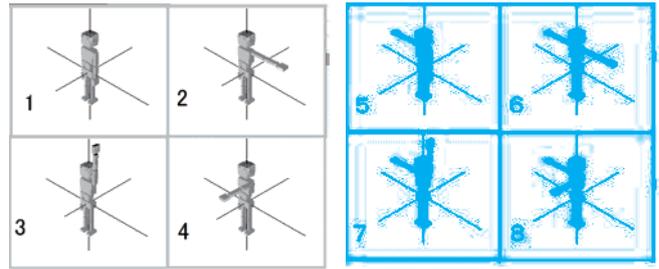


Fig. 5 (a) Basic gestures 1-8.
基本ジェスチャー1-8

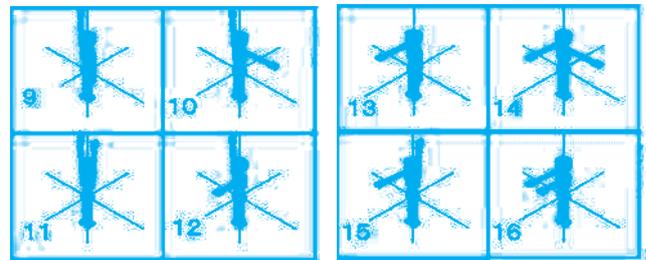


Fig. 5 (b) Basic gestures 9-16.
基本ジェスチャー9-16

左手首の6つの特徴点を用いて、3次元空間におけるジェスチャーの数値表示モデルと、それに基づくジェスチャー間の差を数値化した指標 (ジェスチャー間の距離モデル) を提案する。

ジェスチャーGは6つのベクトルで表すことができる。

$$G = \{V_k | k=1, \dots, 6\} \quad (1)$$

ここではV₁, V₂, V₃, V₄, V₅及びV₆は頭部からそれぞれ右肩、右肘、右手首、左肩、左肘、左手首へのベクトルである。あるジェスチャー認識システムに定義されたジェスチャーがm個あると、i番目のジェスチャーG_iは式(2)で示される。

$$G_i = \{V_{ik} | k=1, \dots, 6\} \quad (2)$$

{G_i | i=1, ..., m} は全てのジェスチャーからなる集合である。

頭部を原点とした3次元空間において、式(2)のV_{ik}は式(3)で示すことができる。

$$V_{ik} = (x_{ik}, y_{ik}, z_{ik}) \quad (3)$$

ジェスチャーG_iとG_j間の距離は式(4)に示すように各

ベクトル成分の距離和として示される。

$$\begin{aligned} dist(G_i, G_j) &= \sum_{k=1}^6 \| V_{ik} - V_{jk} \| \\ &= \sum_{k=1}^6 \sqrt{(x_{ik} - x_{jk})^2 + (y_{ik} - y_{jk})^2 + (z_{ik} - z_{jk})^2} \end{aligned} \quad (4)$$

3.4 機能設計

ジェスチャー機能設計段階ではロボット制御用のジェスチャーとして、ロボットから誤認識の少ないジェスチャーを選択する必要がある。つまり、選択可能なジェスチャーからロボットの制御に必要な数のジェスチャーを選択し、ロボットへの伝達を行う。

例えば、選択可能な16種類のジェスチャーに対応する特徴空間領域がFig. 6のように示されており、かつ5種類の制御用ジェスチャーが必要とされる場合、{1, 3, 5, 7, 10}の方は{9, 10, 11, 12, 13}より誤認識が発生する可能性が小さい。その理由は、後者の場合は各ジェスチャーの特徴空間領域が近接し、距離原則に反することによる。

機能設計段階のリスク減少策は、グループ内ジェスチャー間の距離和が最大であるグループを選択しなければならない。実際に、Table 1によるFig. 5の16のジェスチャーから5つのジェスチャーを選択する場合、距離和が最大であるジェスチャーグループ{1, 2, 5, 12, 15}を選択する。

次に、選択されたこれら5つのジェスチャーをロボットの制御用、すなわち準備、前進、停止、左折、右折という5つの動作に対応させる設計を行う時の方法について述べる。Fig. 5を参照すれば、G1が準備、G2が左折、G5が右折という対応付けは容易に考えられる。しかし、G12、G15はロボットの停止と前進にそれぞれ対応つける場合、人間の日常的な行動習慣からみると適切ではない。さらにG2のような左へ伸ばすジェスチャーをロボットに対する「右折」命令に対応つけると極端にリスクの高い設計となる。このような人間の習慣とまったく反す

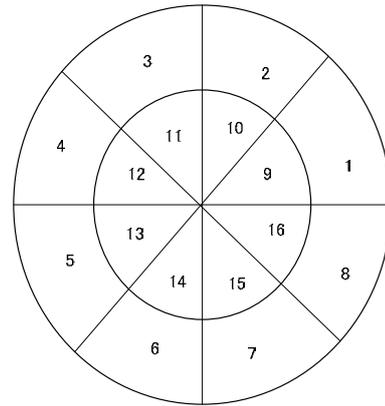


Fig. 6 Sub-regions in feature space of gesture. ジェスチャーの特徴空間領域の例

Table 1 Standard distances between 16 basic gestures by applying average body-size. 人体平均寸法を用いた16種類の基本ジェスチャー間の標準距離(文献1)より引用

Gesture No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	0	138	195	138	138	276	333	276	195	333	390	333	138	276	333	276
2	138	0	138	138	276	138	276	276	333	195	333	333	276	138	276	276
3	195	138	0	138	333	276	138	276	390	333	195	333	333	276	138	276
4	138	138	138	0	276	276	276	138	333	333	333	195	276	276	276	138
5	138	276	333	276	0	138	195	138	138	276	333	276	138	276	333	276
6	276	138	276	276	138	0	138	138	276	138	276	276	276	138	276	276
7	333	276	138	276	195	138	0	138	333	276	138	276	333	276	138	276
8	276	276	276	138	138	138	138	0	276	276	276	138	276	276	276	138
9	195	333	390	333	138	276	333	276	0	138	195	138	138	276	333	276
10	333	195	333	333	276	138	276	276	138	0	138	138	276	138	276	276
11	390	333	195	333	333	276	138	276	195	138	0	138	333	276	138	276
12	333	333	333	195	276	276	276	138	138	138	138	0	276	276	276	138
13	138	1.276	333	276	138	276	333	276	138	276	333	276	0	138	195	138
14	276	2.138	276	276	276	138	276	276	276	138	276	276	138	0	138	138
15	333	3.276	138	276	333	276	138	276	333	276	138	276	195	138	0	138
16	276	276	276	138	276	276	276	138	276	276	276	138	138	138	138	0

る対応付けは誤ったジェスチャー指示、つまりヒューマンエラーを招くリスクが高い。

従って、リスク評価の計算のみによるジェスチャー設計ではなく、人間の習慣をも考慮すべきである。そこで、前述の準備、前進、停止、左折、右折という5つの命令に対し、最大距離グループ{1, 2, 5, 12, 15}のG12とG15の代わりに、G6とG11を入れ、G6を停止、G11を前進に対応させた。

3.5 パラメーター設計

1) サンプル分散の考慮

パラメーター設計段階では、ジェスチャーのサンプル分散に対応する調整を行う。ジェスチャー G_i のサンプル分散とは、そのジェスチャーを行う複数の人間の個体差、すなわち腕の長さや身長の間隔などから生じる標準ジェスチャーの特徴値との距離である。未知の認識対象サンプル G_x をジェスチャー G_i に同定するため、 G_x と G_i の標準ジェスチャーとの距離 $dist(G_x, G_i)$ がジェスチャー G_i のサンプル分散の範囲内であるかを判断するための閾値 T_i を決める。この T_i が、 $dist(G_x, G_i) < T_i$ の条件を満たせば、 G_x は G_i と同定することができ、ちなみに、 $x=i$ となる。

次に、閾値 T_i 設定の考え方をFig. 7を用いて、ジェスチャーG1、G2とその対応する特徴空間領域により説明する。ここで、G1とG2は異なるジェスチャー G_i の内、標準ジェスチャー距離が最小となる組み合わせとする。それぞれの特徴空間領域の円心は標準ジェスチャーであり、それと最大分散を有するサンプルの距離は r_1 (ま

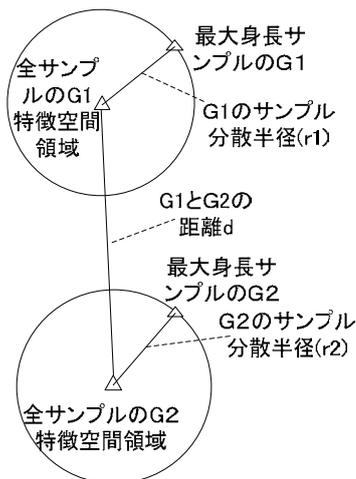


Fig.7 Distribution of the same gesture VS. distance between different gestures.
同じジェスチャーのサンプル分散と異なるジェスチャー間の距離

たは r_2)で表す(以下、分散半径と呼ぶ)。ジェスチャーG1において、閾値 T_1 は、 $T_1 > r_1$ を満足しなければならない。しかし、実際には、あらゆる実サンプルに対応できる適切な r_1 は事前に得られないため、 T_1 を余裕を持たせて大きめにする必要はある。ただし、過大な r_1 の選択は、G1とG2に対応する二つの円を接近させることになり、その結果G2のサンプル分散の一部をG1と誤認識するリスクが生じる。

2) 人体寸法データベースを用いた設計例

実サンプルに対応できる分散半径を適切に選択するため、人体寸法データベース⁷⁾から、最大身長サンプルグループを用いることにした。すなわち、このサンプルグループは標準ジェスチャーと最大分散を有するサンプルと見なすことができる。

Fig. 8にこのデータベースを用いたジェスチャーの適切な分散半径の設定例を示す。ここでは、ジェスチャーの分散半径(各ジェスチャーの対応する第1列)はすべて50以下である。具体的に、G1の第1列は39.76を示す。また、異なるジェスチャー間の最小距離(各ジェスチャーの対応する第2列)はすべて160以上である。閾値 T_1 はできる限り小さい方が他のジェスチャーとの混同を避けるという点で望ましいが、実測のばらつきを考慮すると、 r_1 の1割マージンを含んだ44が適切な閾値と考えられる。すなわち、 $T_1 = 44$ となる。

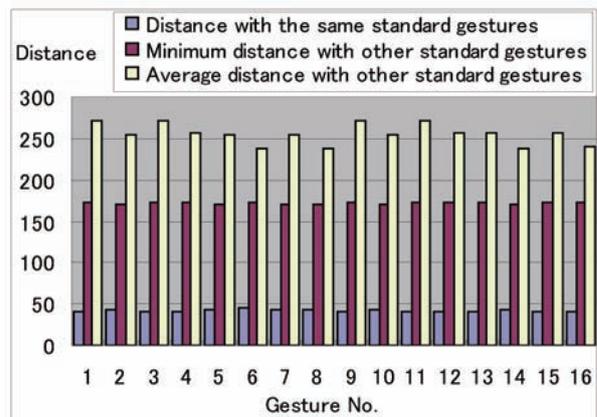


Fig. 8 Analysis of distances between gestures of the max height sample and the standard.
最大身長サンプルと標準ジェスチャーの距離

4. まとめ

ジェスチャー認識技術をロボットの制御に応用する場合を対象にして、誤認識から起因する災害を減少するためのジェスチャー認識システムの設計手法を提案した。

ジェスチャー認識システムの設計段階において、ジェスチャーの仕様定義、特徴選定、ジェスチャーの認識(分類)方法などについて設計選択が必要である。これら設計選択について、ジェスチャーの誤認識リスクの減少に繋がる手法を設計段階の3ステップに分けて説明した。具体的には、基本設計ステップでは、簡単化原理について説明し、機能設計ステップ及びパラメーター設計ステップでは、ジェスチャーの数値化モデル及びジェスチャー間距離モデルについて説明した。このような設計手法はジェスチャー認識を用いた高機能ロボットに適用でき、ジェスチャー認識手法の安全評価の標準化に繋がると考える。

ター, AIST人体寸法データベース1997-98, <http://www.dh.aist.go.jp/AIST91DB/>

(平成20年11月13日受理)

参考文献

- 1) 呂健, 姜偉, 濱島京子:オペレータのジェスチャー認識を利用した移動ロボットとのコミュニケーション手段—制御用ジェスチャーの数値化モデルと誤認識リスクの減少—, 産業安全研究所特別研究報告, NIIS-SRR-NO.33 (2005) pp43-52
- 2) J.Lu : Assessing and reducing the misrecognition risk for robot control gestures using pattern recognition method, Proceedings of SIAS2007 (5th International Conference on Safety of Industrial Automated Systems), Tokyo, November 12-13 (2007) pp.391-396
- 3) IEC 61508, Functional safety of electrical / electronic /programmable electronic safety-related systems Part1-7, IEC, Geneva, Dec. 1998-Feb. 2000.
- 4) 佐藤吉信, 井上紘一, 熊本博光:人間—ロボット系の安全性評価について(第3報, 順序依存型故障論理の定量化について), 日本機械学会論文集(C), vol52, no.475 (1986) pp.1110-1117
- 5) H. Kumamoto, E. J. Henley : Probabilistic Risk Assessment and Management for Engineers and Scientists, IEEE PRESS, pp.166-172, New York, 1996.
- 6) V. Parameswaran and R. Chellappa : View independent human body pose estimation from a single perspective image, CVPR 04, Vol.2, pp.16-22, 27 June-2 July 2004.
- 7) 産業技術総合研究所デジタルヒューマン研究セン