

Kinect を用いた ジャグリングの技判定システムの構築とその改良

長岡 俊男^{1,a)} 伊藤 毅志¹

概要: 本研究では, Kinect で取得した身体データから特異値分解で技ごとの特徴量を抽出してデータベース化し, それを用いて技判定を行うシステムを提案した. 扱う関節数を増やす, 類似度の計算法を変えるなどの実験を行い, 判定率の変化について考察した. また特異値分解の有無による判定率の違いや, 判定時間に関する実験も行った.

Juggling-Skill Detecting System by using Kinect and Improvement

NAGAOKA TOSHIO^{1,a)} ITO TAKESHI¹

Abstract: In this research, we propose a system that extracts the feature value of each juggling-skill by singular value decomposition from the physical data by using Kinect and generates a database. It detects juggling-skill by using the database. We examined the changes of the detecting rate by changing the number of joints and the calculation method of the similarity rate. In addition, we compared the difference in the detecting rate by existence of singular value decomposition and examined the detection time.

1. はじめに

身体の動きや道具を用いたパフォーマンスにおいて, 熟達者の技や技術は難易度の高い高度なものほど一般の人にはわかりにくくなることもある. ふだん我々がテレビの中継などでこれらの競技を観る場合には, たいていの場合は実況や解説が付いており, その競技に精通した熟達者の説明を合わせて聞くことでその高度な技を知り, 観戦の助けとすることができる. しかし, そのような実況や解説がそもそも無い映像だけのコンテンツだったり, まだ競技人口が少なく, 十分に解説者を用意できなかつたりするようなマイナーな競技やパフォーマンスも存在する.

そこで, 本研究ではリアルタイムで技の判定ができるシステムを提案する. 熟達者のみができるような技の難易度や完成度を自動で判定できるシステムが存在すれば, 一般の人がよりそのパフォーマンスを楽しむことができ, また

あるいは非熟達者である初心者がその技を習得する際の手助けになることが期待される.

本研究では, ジャグリングにおけるディアボロという道具を用いたパフォーマンスを題材とし, 技の判定を正確に行うシステムの構築を目的とした. 特定の技を複数回撮影し, それらをシステムに判定させて判定率を求め, またその判定率を向上させる方法や判定速度について考察を行った.

2. 関連研究

人間のジェスチャーや歩行動作などの身体動作の解析には多くの手法が提案されており, スポーツやダンスなどの映像に適用した試みも行われている. 2008年に幸らは, 選手が演技をして審判が採点することによって順位を決定するようなスポーツにおいて, 動画像処理による選手の動作解析に基づいて, 自動採点することを目的としたシステムの構築をおこなった [1]. 具体的には体操競技の1つである鉄棒競技の技の判定を, シルエットマッチングという手法によって行った.

近年ではジャグリングを題材とした研究も行われており,

¹ 電気通信大学 情報理工学研究所 情報・通信工学専攻
Department of Computer Science, The University of Electro-Communications

^{a)} nagaoka@minerva.cs.uec.ac.jp

その多くがボールという道具を用いたパフォーマンスに関する研究である。ジャグリングを用いて熟達への様々な影響を研究したものや [2][3][4], ジャグリングそのものを研究対象としたもの [5][6], ジャグリングを用いて「身体知」というものを明らかにしようとしたもの [7], さらに、ロボットとジャグリングを組み合わせた研究 [8] なども行われている。ボール以外の研究としては、デビルスティックという道具の制御に関する研究なども行われている [9].

本研究では、最終的にジャグリングというジャンルに属するすべてのパフォーマンスの判定を行うことを目標とし、まずはこれまであまり研究が行われていないディアボロを題材として取り上げた。

2012年に姜らは、身体動作のデータから特異値分解という手法で動作の特徴を抽出し、手招きジェスチャーの認識を行った [10]. 特異値分解は近年注目されている特徴量抽出の手法であり、データ長に依存せず時系列データの個数に対する制約が低いという特徴がある。この研究では、右手に装着した5つのマーカーによって得られた装着部位の時系列データから、特異値分解で特徴量を抽出して従来手法と比較し、結果としてより認識率が高かったことから特異値分解による手法の有用性が示されたと述べられている。

本研究では Kinect にて身体データを取得し、特異値分解による特徴量マッチングにて技の判定を行う。

3. 提案システム

3.1 ジャグリング

これまでジャグリングとは、複数の物を空中に投げ、それを取るという動作を繰り返し、常に1つ以上の物が浮いている状態を維持し続ける技術を指してきた。しかし、近年のジャグリング人口や道具のバリエーションの増加に伴いその意味は拡大され、現在では「道具を使った、修練の必要な特殊な技能、または芸」や「様々な物を巧みに操る、一般の人には出来ないような技術」とされる。今回実験に用いたディアボロ (図1) は、お椀を2個つなげたような道具を2本のハンドスティックに通した紐で回すことにより安定させ、操る道具である。

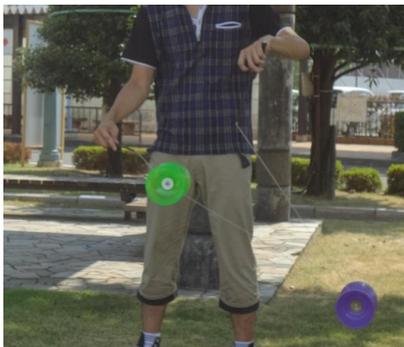


図1 ディアボロ
Fig. 1 Diabolo

3.2 Kinect

Kinect とは RGB カメラ、距離カメラなどを内蔵した Microsoft 製のセンサーである。距離カメラを使用してプレイヤーの認識、プレイヤーの骨格認識が可能である。Kinect の大きな特徴は、身体にマーカー等のセンサーを付けることなく身体データを取得できる点である [11]. Kinect によって得られる関節位置の座標データの値は、 x, y 座標は RGB 画面の中心を点 O として $-1.0 \leq x \leq 1.0$, $-0.9 \leq y \leq 0.9$, z 座標は Kinect からの距離 (m) である。

3.3 特異値分解

時系列データから m 個の離散データを n 個の重複を許して抽出し、 $n \times m$ のハンケル行列 M (逆の対角成分が等しい行列) を構成する。例えば $\{a, b, c, d, e, f, g, \dots\}$ という時系列データが得られたとし、 $m = 4$, $n = 4$ としたハンケル行列 M は、

$$M = \begin{bmatrix} a & b & c & d \\ b & c & d & e \\ c & d & e & f \\ d & e & f & g \end{bmatrix} \quad (1)$$

となる。この行列 M を特異値分解によって次のように分解する。

$$M = U \Sigma V^T \quad (2)$$

U は $n \times n$ のユニタリ行列、 Σ は $n \times m$ で対角成分以外は零で対角成分は非負数の行列、 V^T は $m \times m$ のユニタリ行列である。この Σ の要素が特異値を示し、 U を左特異ベクトル、 V を右特異ベクトルと呼ぶ。左特異ベクトル U は行列 M の特徴を示し、特異値 Σ は M に対する左特異ベクトルの影響の大きさを示している。

本システムではこの左特異ベクトル U の第1列成分を技の特徴量ベクトルとし、このベクトルの類似性で技の判定を行う。

3.4 システム概要

本システムの概要を図2に示す。

本研究では特異値分解を用いて、Kinect で取得した技を行っている間の身体関節位置データから左特異ベクトルを抽出して、それぞれの技ごとの特徴量をデータベース化しておく。そして判定したい技からも左特異ベクトルを抽出し、このデータベースと比較して技の判定を行った。データプロットは0.1秒ごとに3秒間行うので、計30個の時系列データが得られる。この3秒の間に指定された技を行うこととした。ハンケル行列のパラメータは暫定的に $m = 20$, $n = 11$ で行った。よって、得られる特徴量は左特異ベクトル U が 20×20 なので20要素を持つベクトルである。

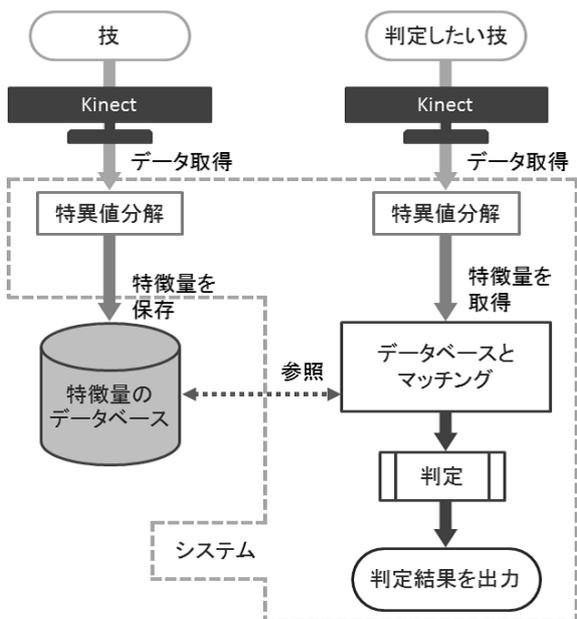


図 2 本システムの構成図
Fig. 2 Outline of this system

4. 実験

4.1 判定結果と判定精度向上への実験

4.1.1 手順

実験はジャグリングサークル所属のディアボロを専門とする男性 2 名で行った。ディアボロを用いた技を 9 種類 15 回ずつ行わせ、データを採取し判定を行った。15 回のうち、初めの 5 回はデータベース用のデータであり、残りの 10 回がマッチング用のデータである。判定は、各関節の各座標ごとに類似度が最小となる技を判別し、すべての類似度のカウント個数が最大の技を判定結果とする「楽観判定」、そのあと技ごとにデータベースに登録されている 5 回分のデータを合計したカウント個数が最大の技を判定結果とする「多数決判定」の 2 種類を行った。

まず両手、両肘の 4 関節のデータで判定を行った。次に判定率の向上に向けて、それらに両肩、両股関節、両ひざ、両足のデータも加えた 12 関節データでの判定を行い、結果を比較した。さらに、上 2 つの判定では関節ごとの各座標要素で類似度の計算を行ったが、各座標要素の類似度を合計することで関節ごとの類似度を計算し、関節ごとにマッチングする判定も行い、結果を比較した。この判定は 12 関節データで行った。

4.1.2 結果

4 関節のデータでの判定、12 関節データでの判定の結果と関節単位でマッチングを行った判定法による判定の結果を表 1 に示した。各技の判定率は 10 回判定を行った際の平均値である。

判定率の平均を見ると、4 関節データより 12 関節デー

表 1 判定法の違いによる結果の比較 (%)
Table 1 Comparison of The Results
by Difference in Detecting Method(%)

技名	4 関節判定		12 関節判定		関節単位判定	
	楽観	多数	楽観	多数	楽観	多数
腕回し	85	75	85	90	95	90
足回し	90	100	95	100	100	100
チャイニーズ	60	45	73.3	70	73.3	90
エレベーター	100	100	90	88.3	100	100
インテグラル	100	100	100	100	100	100
ジェノサイド	80	90	100	100	100	100
スピードループ	85	90	72.5	80	71.7	75
ふりこ	80	90	100	95	100	100
アンブレラ	82	90	90	95	85	85
平均	84.7	86.7	89.5	90.9	91.7	93.3

タを用いた判定の方が良い結果となり、関節単位での判定がそれよりさらに良い結果となった。しかし技ごとに見ると、判定率が上がった技と判定率が下がる技があった。

4.1.3 考察

関節データを増やすと、判定率の平均は上がっているが、個別に見ると判定率が下がっている技も多く確認できた。これは増やしたパラメータが誤判定を招いている可能性があり、技ごとに判定に必要な関節情報が異なることが原因だと考えられる。

また、座標要素ごとより関節単位でのマッチングの方が全体の判定精度は上がるということがわかった。これは判定の際に「手」や「肘」といった関節自体の動きに意味があるということだと考えられる。

4.2 他人のデータベースを用いた判定

4.2.1 手順

4.1 節の実験ではそれぞれの被験者は自分の技のデータベースを用いて判定を行った。そこで、それぞれの被験者のデータベースを入れ替えて判定を行うと判定率がどう変化するかを検証した。判定は 4 関節 (両腕、両肘)、6 関節 (4 関節+両肩)、12 関節 (6 関節+両股関節、両ひざ、両足) のデータを用いてそれぞれ行い、すべて 4.1 節で結果の良かった座標要素をまとめた関節ごとの判定法を用いた。

4.2.2 結果

12 関節データを用いた判定において被験者 B が、被験者 B(自分) のデータベースを用いて判定を行った結果と、被験者 A(他人) のデータベースを用いて判定を行った結果を表 2 に示した。

結果として他人のデータベースを用いて判定を行うと、判定率が 60%前後まで大きく下がることがわかった。なお、被験者 A が被験者 B のデータベースを用いて判定を行った結果も、同様に大きく判定率が下がる結果となった。

また他人のデータベースを用いた判定において、4 関節

表 2 自他人のデータベースを用いた判定結果の比較 (%)

Table 2 Comparison of The Results with One's Database and Others(%)

技名	自分の DB		他人の DB	
	楽観	多数	楽観	多数
腕回し	95	90	60	77.5
足回し	100	100	48.3	52
チャイニーズ	73.3	90	55	70
エレベーター	100	100	75	95
インテグラル	100	100	38.3	70
ジェノサイド	100	100	100	100
スピードループ	71.7	75	65	75
ふりこ	100	100	36.7	13.3
アンブレラ	85	85	10	10
平均	91.7	93.3	54.3	62.5

表 3 他人のデータベースを用いた場合の関節データ数の違いによる結果の比較 (%)

Table 3 Comparison of The Result by The Difference in The Number of Joint Data with Others' Database(%)

技名	4 関節判定		6 関節判定		12 関節判定	
	楽観	多数	楽観	多数	楽観	多数
腕回し	100	100	95	100	73.3	58.3
足回し	95	95	95	88.3	100	90
チャイニーズ	55	50	36.7	30	20	30
エレベーター	100	100	100	100	80	100
インテグラル	55	55	78.3	83.3	90	90
ジェノサイド	30	30	0	11.7	0	3.3
スピードループ	32.5	32.5	25	25	60	60
ふりこ	95	95	100	100	100	100
アンブレラ	100	100	90	100	65	60
平均	73.6	73.1	68.9	70.9	65.4	65.7

データで判定した結果, 6 関節データで判定した結果, 12 関節で判定した結果を表 3 に示した. データベースは被験者 A のものを使い, 判定は被験者 B の技を判定した.

結果として他人のデータベースを用いた判定においては, 関節データ数を減らして両肘と両手のデータだけを用いて判定を行った結果が最も判定率が高かった. これは自分のデータベースを用いて判定を行った際の, 関節数を増やせば判定率が上がるという結果とは逆の結果である.

4.2.3 考察

データベースを変えた実験に関して, 本人のデータベースを用いて判定を行った場合より大きく判定率が下がった. このことから, 動きが同じであっても技を行う人によって微妙な差異があり, 本実験の手法ではその差異が障害となり判定率の低下に繋がっていると考えられる.

また, 自分のデータベースを用いた際の判定とは異なり関節データ数を減らすことで判定率が上がった. このことから, 共通の動きを示している関節とそうでない関節の差が大きいのではないかと考えられる. 今回は暫定的に 4 関

節, 6 関節, 12 関節で実験を行ったが, さらに効率の良い組み合わせが存在する可能性も十分ありえる.

4.3 特異値分解を介さない判定

4.3.1 手順

4.1 - 4.2 節の実験では姜らの研究に倣い, 時系列データから特異値分解によって左特異ベクトルを求め, その第 1 列成分を特徴量として用いて判定を行った. そこで, 本節では特異値分解を介さずに時系列データのみによるマッチングを行い, 特異値分解の有無による結果の比較を行った.

実験は自分のデータベースを用いて判定を行った場合と他人のデータベースを用いて判定を行った場合の, 12 関節データを用いて座標要素をまとめた関節ごとの判定法で, 特異値分解の有無それぞれで判定を行い, 結果を比較した.

特異値分解を行わずに判定を行う際, 関節ごとに取得した 30 個の時系列データをそれぞれの要素順に比較してその差分の合計を類似度とした.

4.3.2 結果

自分のデータベースを用いた場合の, 特異値分解の有無による結果の比較を表 4 に, 他人のデータベースを用いた場合の, 特異値分解の有無による結果の比較を表 5 に示した.

どちらのデータベースを用いた場合も, 特異値分解を行わない方が判定率が高い結果となった. 特に自分のデータベースを用いた場合の結果は 99% を超える判定率となり, また他人のデータベースを用いた場合の結果は楽観の判定率が大きく上がった.

4.3.3 考察

比較しているデータ数が 20 個と 30 個で違うとは言え, 特異値分解を用いない方が判定率が明らかに高い結果となった. そもそも特異値分解は次元圧縮の際に用いられることが多く, 今回のように 30 個のデータを 20 個に圧縮す

表 4 自分のデータベースを用いた場合の特異値分解の有無による結果の比較 (%)

Table 4 Comparison of The Results by Existence of SVD with One's Database(%)

技名	特異値分解あり		特異値分解なし	
	楽観	多数	楽観	多数
腕回し	95	90	100	100
足回し	100	100	100	100
チャイニーズ	73.3	90	95	95
エレベーター	100	100	100	100
インテグラル	100	100	100	100
ジェノサイド	100	100	100	100
スピードループ	71.7	75	100	85
ふりこ	100	100	100	100
アンブレラ	85	85	100	100
平均	91.7	93.3	99.4	97.77

表 5 他人のデータベースを用いた場合の特異値分解の有無による結果の比較 (%)

Table 5 Comparison of The Results by Existence of SVD with Others' Database(%)

技名	特異値分解あり		特異値分解なし	
	楽観	多数	楽観	多数
腕回し	60	77.5	100	100
足回し	48.3	52	100	100
チャイニーズ	55	70	23.3	30
エレベーター	75	95	100	100
インテグラル	38.3	70	36.7	30
ジェノサイド	100	100	85	85
スピードループ	65	75	77.5	77.5
ふりこ	36.7	13.3	43.3	33.3
アンブレラ	10	10	80.8	70
平均	54.3	62.5	71.9	69.5

るという程度の場合では用いるべきではなかったのかもしれない。

4.4 判定時間

4.4.1 手順

特異値分解の有無、関節データ数、時系列データの長さの違いなどによる判定を行い、判定にかかった時間を比較した。計測したのは時系列データからハンケル行列を作成するのにかかった時間、ハンケル行列を特異値分解するのにかかった時間、特異値分解で得られた特徴量をデータベースとマッチングして判定結果を出すまでにかかった時間である。用いる関節データ数の違いによる判定時間の差を調べるため 6 関節データでの判定と 12 関節データでの判定を行った。また、用いる時系列データの長さの違いによる判定時間の差を調べるため、これまでの 30 個の時系列データを用いた判定と、0.05 秒毎に 3 秒間データプロットを行い 60 個の時系列データを用いた判定を行った。60 個の時系列データを用いる際のハンケル行列の大きさは 20×41 とし、得られる特徴量の大きさはこれまでと変わらないようにした。

4.4.2 結果

手順で述べたそれぞれの判定条件に関する番号付けを表 6 に示した。以降、この番号付けを参考に議論を進める。そして、それらの判定時間に関する結果を表 7 に示した。結果は 10 回測定を行った平均値である。また、この結果をグラフにしたものを図 3 に示した。

4.4.3 考察

全体的に見て、ハンケル行列の作成にはあまり時間はかかっておらず、ほとんどがマッチングにかかっている時間である。

(1) と (2) を比較すると、このパラメータにおいては特異値分解の有無ではそれほど大きく時間は変わらないことが

表 6 判定条件の番号付け

Table 6 Numbering of Detecting Condition

番号	SVD	関節数	データ長
(1)	あり	12	30
(2)	なし	12	30
(3)	あり	6	30
(4)	なし	6	30
(5)	あり	12	60
(6)	なし	12	60

表 7 判定時間の比較

Table 7 Comparison of Detection Time

番号	行列作成	SVD	マッチング	合計
(1)	0.02	0.07	0.64	0.73
(2)	—	—	0.71	0.71
(3)	0.01	0.08	0.32	0.41
(4)	—	—	0.37	0.37
(5)	0.02	0.24	0.66	0.92
(6)	—	—	0.72	0.72

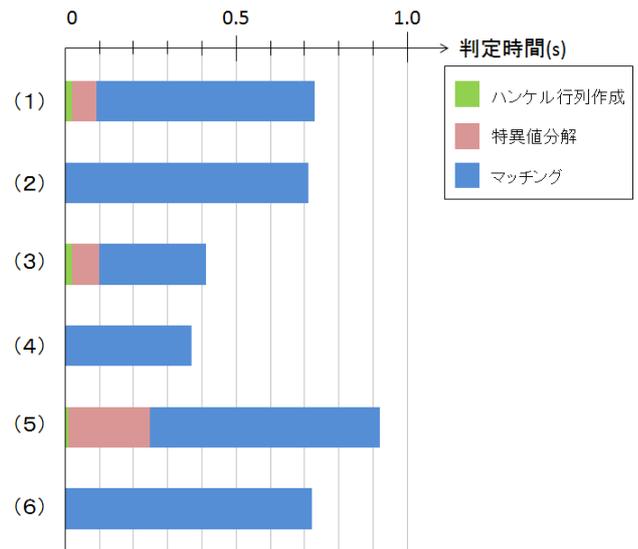


図 3 判定時間の内訳

Fig. 3 Distribution of The Detection Time

わかる。

(1) と (3) を比較すると、扱う関節データ数を減らすことによってマッチングにかかる時間がおよそ半減している。しかし特異値分解にかかる時間には大きく変化がなかった。

(1) と (5) を比較すると特異値分解にかかった時間が大きく増加している。

(2) と (6) を比較すると、比較に用いたデータ長が 30 個から 60 個になっているにも関わらず、マッチングにかかる時間がほぼ変化していないことがわかる。

以上のことから、ハンケル行列が大きくなる（データ長が長くなる）と特異値分解にかかる時間が増えること、扱う関節データ数が増えるとマッチングにかかる時間が増

えること、比較するデータ長が増える程度ではマッチングにかかる時間に影響は無いことがわかった。最後の結果に関しては、マッチングにかかる時間のうち大半を占めているのがデータベースアクセスやデータの open, close にかかっている時間なのではないかと考えられる。

5. おわりに

本研究では、Kinect で取得した体の関節位置データを元に特異値分解によって左特異ベクトルを抽出することで、技ごとの特徴量をデータベース化し、それを用いて技判定を行うシステムを提案した。また全体の技の判定率を上げるために、取得する関節を増やす、類似度の計算法を変えるなどの実験を行い、判定率の変化についても考察した。

結果、データとして取得する関節数を増やし、類似度計算の方法を関節単位での計算法に変えることで、判定率を上げることに成功した。このことより、技の判定は得られたデータを単に比較するだけでなく、類似度の計算法などを工夫することで判定率を上げられることがわかった。また、他の人のデータベースを用いた判定では判定率が下がる結果となり、1人のデータで万人の特徴量をカバーできるわけではないということがわかった。しかし、自分のデータベースを用いて判定を行った際とは逆に、判定に用いる関節データ数を減らすことで判定率を上げることに成功した。このことに関しては、さらに被験者を増やして実験を行い、様々なパフォーマーにも対応できるような判定条件の組み合わせを模索していく必要がある。

また、特異値分解を介さず時系列データをそのままマッチングに用いて判定を行い、特異値分解を用いた結果と比較する実験や、条件を変えて判定時間を比較する実験を行った。結果として、特異値分解を用いない方が判定率は高く、さらに判定時間に関しても短い判定時間で済むという結果となった。こちらも上で述べたものと同じように、さらに多くの被験者で実験を行って本当に特異値分解を用いるべきではないかを調べていく必要がある。

結論として、1人の技を判定するのであればその人のデータベースを用いることで90%を超える確率で判定できる結果となった。一方、どんな人の技でも判定できるようにするためには問題が残る結果となった。しかし、判定しやすい技とそうでない技があったため、これからは技ごとにさらに詳細に判定結果を調べていく必要がある。例えば、判定率の低い技がいったいどのような技と誤認されているかを調べたり、判定率の高い関節を調べるなどである。

この実験を通して、今回特徴量としてきた「左特異ベクトル」がどのような特徴を表しているのかという点が新たな問題であった。4.3.3節でも述べたとおり、特異値分解はもともと次元圧縮でよく用いられる手法であり、姜らの研究においてこれによって得られたデータを身体知と位置づけているが、それが直接何を指すのかまで指摘されてい

ない。これを解明できれば、この特徴量では足りない要素を補うために新たなデータを取得しマッチングすることで、さらに判定精度を上げられる可能性がある。

今後の課題としては、本実験で暫定的に設定したハンゲル行列の行数と列数や比較する左特異ベクトル数、データプロットの間隔、技ごとのデータベース数などのパラメータの調整、ディアボロの位置取得などが挙げられる。また、今回はあまり考慮しなかったリアルタイム性に向けて、連続した動きの中からの判定や、判定の速度を考慮しながら改良を加えていく必要もある。

参考文献

- [1] 辛 貞殷, 小沢 慎治: 『動画像処理によるスポーツ運動解析の研究: 鉄棒競技の自動採点システムに向けて (テーマセッション, センシングのための認識・理解)』電子情報通信学会技術研究報告. PRMU, パターン認識・メディア理解 Vol.108, No.46, pp.13 - 18(2008.05.15)
- [2] 妹尾 江里子: 『Ball Juggling 課題に及ぼす視覚-運動イメージ・リハーサルの効果』日本体育学会大会号, 39A, pp.165(1998.08.20)
- [3] 山近 慎二, 阪口 豊: 『お手玉の運動制御における視覚情報の役割』電子情報通信学会技術研究報告. NC, ニューロコンピューティング Vol.97, pp.223 - 228 (1998.03.19)
- [4] 山近 慎二, 阪口 豊: 『お手玉の運動制御における視覚情報の役割 (2)』電子情報通信学会技術研究報告. NC, ニューロコンピューティング Vol.98, pp.112 - 122(1999.03.19)
- [5] 西野 順二: 『ジャグリングの連続技生成』ゲーム情報学, Vol.27, pp17 - 23(2002.03.15)
- [6] 西野 順二: 『ロボカップサッカーとジャグリング-計算される日常-』情報処理学会研究報告. AL, アルゴリズム研究会報告, No.5, pp31 - 37 (2007.01.23)
- [7] 田中 彰吾, 小笠原 慶太: 『身体知の形成: ボールジャグリング学習過程の分析』人体科学 No.1, pp.69 - 82, (2010.05.30)
- [8] 脇屋 真一, 西島 良雅, 並木 明夫: 『高速ハンドアームと高速ビジョンを用いたボールのダイナミックマニピュレーション』ロボティクス・メカトロニクス講演会講演概要集 2010, “1A1-F30(1)” - “1A1-F30(4)” (2010)
- [9] 川井田 康礼, 中浦 茂樹, 大島 龍介, 三平 満司: 『デビルスティックの持続回転運動のフィードバック制御』「振動と運動の制御」シンポジウム講演論文集, No.8, pp540 - 545(2003.10.30)
- [10] 姜 銀来, 林 勲, 王 碩玉: 『特異値分解による運動動作の特徴獲得』日本知能情報ファジィ学会誌. 知能と情報 Vol.24, No.1, pp.513 - 525(2012)
- [11] 中村 薫, 田中 和希, 宮城 英人: 『KINECT for Windows SDK プログラミング C#編』株式会社ゲッシュ (2012)