

ボールジャグリング練習映像に基づいた 言語教示推薦システムの構築と応用

Building the Instruction Recommendation System Based on Ball Juggling Practice Movie and Application

岡田健嗣^{1*} 脇坂崇平¹ 高原慧一¹ 荒川陸^{1,2} 稲見昌彦¹

Taketsugu Okada¹ Sohei Wakisaka¹ Keiichi Takahara¹ Riku Arakawa^{1,2} Masahiko Inami¹

¹ 東京大学

¹ The University of Tokyo

² Carnegie Mellon University

Abstract: One of the methods to acquire motor skills is to receive verbal instructions from a skilled person. While they have been used in various situations, they are based on the subjective judgment of a skilled person who observes the movement of a novice. Therefore, it has not been clearly expressed how the instructions act on the indexes of the motor movement and the structural relationship between the instructions. In this study, we focused on ball juggling learning. We constructed a system “Skill Sapuri” that recommends instruction according to the skill state expressed calculated from the ball juggling video. We also discussed the possibility of using this system to aid in the understanding of embodied knowledge.

1 序論

1.1 はじめに

ヒトが獲得した技や身体運動のコツのことを、身体知と呼ぶ。[1, 2]。これはスポーツや楽器演奏などに限らず、歩く、座る、などの日常的な動作、ひいては音楽やアート鑑賞などの感性、お笑い芸人のボケなども身体知に含まれる。初心者が新しい身体知を獲得する際には、インストラクターから受ける指導がしばしば重要な役割を果たす。例えば接触の伴う指導（いわゆる『手取り足取り』）や、模範的動作の呈示、言語教示などがそれに当たる

（ここで教示による指導とは、“言語によって問題の解決について一定の方向を示唆すること [3], ”を指す）。

最近では、ICT 技術を援用して身体スキル指導を自動化する試みが広く行われているが、その際パーソナリゼーションを考慮する必要がある。すなわち、インストラクターによる対人指導では普通に行われるような、練習者の習熟状態、癖、特性を考慮した上での教示をどのよ

うに実現するか、という課題である。

例えば、ある運動スキルを同程度に実現していたとしても、適性や状況によって身体の使い方に個人差が存在する（スキルの”多様性”）[4]。それは、練習者の癖や特性、あるいは獲得してきた身体知によるが、そのような個人差を考慮せずに一律的に教示を与えると、効果的な教示にならないだけでなく、場合によっては矯正すべき癖の固定が促進されるなど、学習をむしろ妨げることもありえる。異なる練習理論、アプローチに基づく教示を混ぜる場合も、その効果に注意を払う必要がある。対人指導における指導では、その性質上こういった課題がある程度解決されているが、自動教示推薦システムにおいては、まだこれからの課題であるといえる。

個人の癖や特性を考慮せず、理想的状態に近づけることを目的とした教示手法については、実用化されたアプリケーションが存在する（[5, 6] 等）。そのようなアプリケーションは、身体スキルの種類および学習局面においては、十分に効果的であると期待できる。しかし、理想的状態（モデル）を前提として、そこに向かう教示を自動生成するという手法と、パーソナリゼーションに対応した教示生成の間には、現時点では相当隔たりがあるといえる。

* 東京大学情報理工学系研究科システム情報学専攻

〒153-8904 東京都目黒区駒場 4-6-1

E-mail: taketsugu.okada@star.rcast.u-tokyo.ac.jp

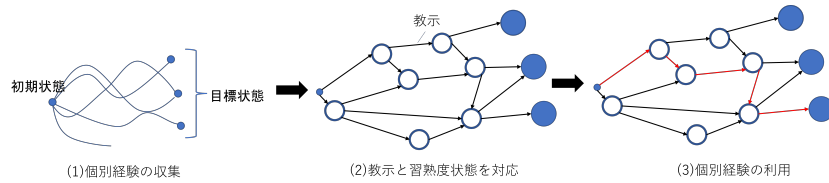


図1 経験サプリメント概略. (1) スキル習熟状態には個人差が存在する. (2) それぞれの習熟状態と対応する解決策 (本研究の場合は教示) を, 経験バンクとして蓄積する. (3) 新規ユーザーが身体スキルを習得する際, 習熟状態に応じた教示を利用する.

1.2 本研究の目的

本稿では, (対人指導で見られるような) パーソナリゼーションを組み込んだ自動教示推薦システム『スキルサプリー』の提案とプロトタイプを紹介を行う. 一言でいうと「対人指導データを大量に収集し, 再利用する」というシンプルな発想であり, このアプローチは黄瀬らが提案した『経験サプリメント』に基づく ([7, 8]. 概要図を図1に示す). より具体的には, 以下の手続きをとる.

教示収集

- 特定の身体スキルの練習動画を撮影し, サーバにアップロード
- 教示者が動画を見て, 教示を行う
- 同時に, 練習動画から練習者の習熟状態を計算
- 以上のループを回して, 習熟状態, 教示の対をデータベースに蓄積

教示利用

- 新規ユーザーが練習動画をアップロードする
- ユーザーの習熟状態を計算
- 教示データベースとユーザーの習熟状態に基づき, 推薦システムが教示を出力

本アプローチを実現するためには, (1) 対象となる身体スキルを決定, (2) 習熟状態を計算可能な指標として数値化, (3) 推薦アルゴリズムの構築 (4) アプリケーション開発, の四点が必要となる. 本論文では, 以上について説明するが, 今後データ収集に伴い, 構築したシステムの妥当性検証, そして収集された習熟状態および教示に対する構造分析をすすめる予定である.

2 スキルサプリー

2.1 対象とする身体スキル: ジャグリング

本研究では, 学習対象となる身体スキルとして, ボールジャグリングの基本技である3ボールカスケード

(3BC) を採用した. 市川らが指摘するように [9], ジャグリングスキル分析には,

- 多くの身体運動には周期的運動が含まれ, ジャグリングのボールの投げる/受けるという運動も同様に周期的である
- 自己組織化の特徴が現れる運動である
- 評価がキャッチ回数という容易な手法で行える
- ジャグリングを習得している人が少ないため被験者実験が容易

という利点がある. これに加え, ジャグリング自体の運動負荷があまり大きくないこと, 広い場所を要しないため気軽に練習を行えること, 習熟状態を表現する指標に関する研究も行われているという利点もある. 更に, 重要な要件である, 「短い練習映像を見ることにより, 教示者が教示を行うことができること」を満たしている. 以上より, スキルサプリーの対象として適切だと判断した.

2.2 ジャグリングの習熟状態を表す指標

本節では, BC の習熟状態を表す指標について記述する. Beek ら [10] は, 3ボールカスケード (3BC) の習熟状態を示す指標として, ボール保持率 k を導入している ($k = \text{ボールが手にある時間} / (\text{ボールが手にある時間} + \text{ボールが保持していない時間})$). Hashizume, 市川によって [11, 9], 実際に熟練者は, 初心者に比べて統計的に k が小さいことが報告されている一方, 学習の初期過程ではこの値が大きい群と小さい群に二分されるという指摘もある [11, 12]. そのため, ボールキャッチ位置安定性に関する研究 [11] や, ボールの軌道に関する研究 [13] がある. また, 高原ら [14] は, k のみでは個人ごとの性質を十分にとらえた習熟状態を表現できないと考え, 肘やボール軌道頂点の位置の安定性をもとに, 習熟状態の表現を試みた.

本論文では, Beek による指標 k , 関節 (肘等) やボール軌道の安定性, 対称性, キャッチ回数をスキルサプリーの習熟状態指標として暫定的に採用している. 後述する推

薦アルゴリズムは、指標の追加が可能なものとなっている。また、習熟状態指標をベクトルとしてまとめたものを習熟状態ベクトル (Skill State Vector:SSV) とする。

2.3 アプリケーション実装及び推薦アルゴリズム

2.3.1 ユーザーの利用フロー

スキルサプリアユーザーは、3BC を数キャッチ以上成功できる初心者想定している。ユーザーは、映像を撮影し、ウェブ経由でアクセスするサーバーに動画をアップロードすることで、教示を得る。教示は (1) 熟練者によって直接付与される方法と、(2)(1) のプロセスで収集した教示データを利用した推薦システムによって教示が付与される方法、の 2 通りが存在する。データベース収集初期は、(1) のみを行う。ユーザー側の操作の流れを図 2 に示す。動画をアップロードし教示が付与されると、メールが届くようになっており、スキルサプリア上で教示内容を確認することができる。

2.3.2 スキルサプリア構成

この章ではユーザーに教示を提示する、スキルサプリアの構成について概観する (図 3)。入力された映像をもとに、骨格検出、ボール座標検出を行い、データの後処理を行ったのち SSV を計算する。(1) の、熟練者によって直接言語教示が付与される場合は、SSV は言語教示とセットにしてデータベースに保存される。(2) の、データベースに基づいて言語教示が付与される場合は、SSV とデータベースを用いて言語教示推薦が行われる。

SSV を計算するために、ユーザーの骨格、ボール座標の検出を行う必要がある。ユーザーの骨格検出には、OpenPose[15] を用いた。また、ボール座標の検出には、Åkerlund による畳み込みニューラルネットワークを用いたアルゴリズム [16] を使用した。これらの値は、動画の解像度、ユーザーの身長によって補正されたのち、正規化される。

ユーザーの骨格や、ボール座標に対し、欠損値計算、平滑化などの前処理をおこなったのち、習熟状態指標計算を行う。本論文執筆時点では、習熟状態を表現する指標として、Beek の指標、高原らが検討した、肘の平均、標準偏差、ボール軌道頂点位置座標の平均、標準偏差、ボールキャッチ回数をを用い、SSV とした。このベクトルと教示を対応させるが、後からでも習熟状態を追加できるようにスキルサプリアの設計を行った。

熟練者による言語教示は、自由入力する方法と、過去に用いた教示の中から選択する方法がある。過去に用いた教示と異なる教示をしたいときのみ自由入力で教示を

してもらおう。

スキル習熟状態に基づいた教示を自動で利用可能にするため、ユーザーベース協調フィルタリングの推薦アルゴリズムの手法の一つである、Factorization Machines[17] を用いた推薦を行う。Factorization Machines は、

- 推薦に用いる指標の追加が容易
- ドメインに関する知識がなくてよい
- 習熟状態指標同士の相互作用を用いて推薦可能
- スパースなデータに対応

という利点がある。概要図を図 4 に示す。ユーザーと教示、教示の評価値からなる行列に対して、教示の評価値が高いものを推薦するアルゴリズムである。この行列の値をより正確に表現するために、評価値の行列を、ユーザーの習熟状態を含む、図 4 の右図のような形式に変換することで、ユーザーの習熟状態を利用することができる。また、Factorization Machines には、Ribeiro らによって提案された Local Interpretable Model-agnostic Explanations:LIME[18] を用いることにより、出力 (推薦結果) にどの SSV の要素が効果的に働いたかを確認することができるという利点がある。また教示と習熟状態、習熟状態の相互作用の対応関係を定量的に観察することも可能である。さらに教示同士の構造についても考えることができる。例えば、類似した SSV が複数の異なる教示を推薦した場合、それらの教示は (1) 類似した効果をもつ、もしくは (2) 異なるアプローチの教え方に属する、といった解釈がありえる。習熟状態間の構造、教示間の構造の具体的分析については、今後解析に必要なデータ数が蓄積した後に行う予定である。

推薦スコア予測値 \hat{y} は、次の式で計算される。

$$\hat{y}(\mathbf{x}) := w_0 + \sum_{i=1}^n w_i x_i + \sum_{i=1}^n \sum_{j=i+1}^n \langle \mathbf{v}_i, \mathbf{v}_j \rangle x_i x_j \quad (1)$$

\hat{y} : 教示の推薦スコア予測値

x : 特徴ベクトル (ユーザー, SSV, 教示)

w : 重み

\mathbf{v} : 潜在ベクトル

w, \mathbf{v} が学習パラメータである。

3 考察

3.1 スキルサプリアとメタ認知

メタ認知自体は 1979 年に Flavell によって [19] 提唱された、自分の認知活動を見直したり調整したりするメカニズムに意識的に気づくという概念である。Nelson ら



図2 スキルサプリーユーザー側の操作動画をホームページ上にアップロードし、技の種類、ボール数、練習している時間を指定する。教示が付与されるとメールが届く。ユーザーからは教示がシステムによって付与されたものか、熟練者が付与したものかはわからない。

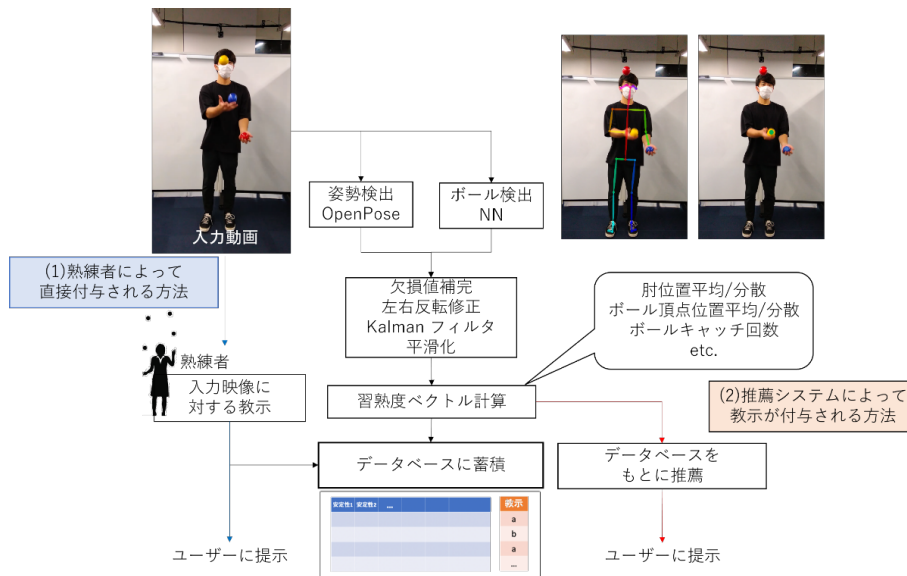


図3 スキルサプリー構成図。入力された動画は、教示付与の方法にかかわらず、習熟状態指標が計算される。熟練者によって言語教示が付与される場合、習熟状態指標と教示がデータベースに格納され、システムによって言語教示が付与される場合、習熟状態指標とデータベースに基づいて言語教示が推薦される。

によって [20] メタ認知知識とメタ認知制御によるメタ認知プロセスがモデル化されている。身体スキルに関するメタ認知は、メタ認知言語化として、諏訪ら [1] によって提唱されてきた。これは、スキル習熟が進行するにつれ、環境や道具との関わりや、身体部位の使い方が漸進的に変化するが、それを言語的に表出させることで、身体と環境のインタラクション自体を発展させることを指している。メタ認知言語化によって、それまで意識していなかった変数に対する意識が生じる。メタ認知言語化はあくまで練習者が一人称的、主体的に行うことが重要であり、単純に教示（コーチング）を与えるのみでは”

受け売り”になってしまうという指摘があったが、スキル習熟に応じた教示を与えることができれば、メタ認知言語化を促進できる可能性がある。

メタ認知言語化に基づき、ジャグリング習熟プロセスにおける言語化の内容と、ジャグリングのパフォーマンスを関連させることで、ジャグリング習熟中に意識する内容がどのように変化するのかを調査した研究が行われてきている ([21],[9],[22])。これらの研究は、スキル習熟プロセスにおいて、主観的な視点での言語化に基づいて分析を行っている。ジャグリングにおける言語化内容の質と、スキル習熟状態に対し一定の関連があることが示

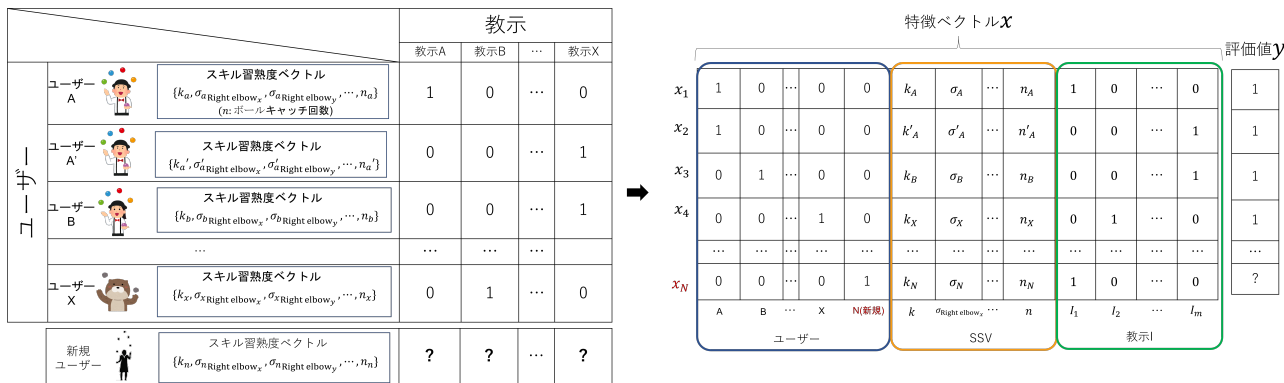


図4 Factorization Machines の概要図。左表のようなユーザーと教示の組のみを用いて教示推薦を行うのではなく、右表のような、ユーザー、SSV、教示を組み合わせた特徴ベクトル x を用いて教示推薦を行う。同一のユーザーに対して、SSV が変化していれば、新しく特徴ベクトルとして加える。SSV を構成する各指標の説明は省略。

唆されている。スキルサプリア自体は熟練者からの教示を与えるだけであるが、これらの研究をもとに、自分でメタ認知言語化した情報を入力することができるようになったり、教示を受けた際に連鎖的に発見したことを記入することができるようになれば、他者と言語化した情報の共有を行えるようになる可能性がある。また、スキルの習熟過程において、習熟状態がどのような変遷をするのかわかれば、それを初心者に提示することで、メタ認知を促進しうる。

3.2 教示者自身への効用

教示者はスキルサプリアの教示収集モードを介してユーザーに対し教示を与えるが、これは教示者自身のスキル習熟も促進すると期待している。Duran らによると [23], 「教えることで学ぶ」ことは以下の段階に分けることができる。すなわち、(1) 保有する知識を他者に伝搬する教授 (telling knowledge), (2) 他者により積極的に理解を促すために知識構築 (building knowledge) を伴う解説, (3) 質疑応答などのあるやりとり (interaction), (4) 教える側と教えられる側による共同作業 (collaboration) の4段階である。この分類に従えば、スキルサプリアでは、通常に教えるだけで得られる (1) だけの効果だけでなく、(2) のプロセスも含めることができ、教示を通して大きな学びの効果を期待することができる。

3.3 流派、方法論の整理

コミュニティや流派、方法論によって、教え方が異なる教示と習熟状態の対応を調査することで、方法論の差異を明確にし、初心者が方法論を選択して学習できる。

3.4 ライブストリーミング映像への対応

スキルサプリアでは、撮影した動画を手動でアップロードするという手法を用いている。アップロード作業に最短でも1分ほどかかることに加え、教示出力計算に2分程度かかる。すなわち、動画を撮影してから教示を得られるまでに、ある程度の時間が経過してしまうという問題が存在する。習熟状態や教示内容によっては、できる限り早く教示しないと効果が出ない場合も考えられる。そのため、手動アップロードする方式に加えて、ストリーミング方式（練習中撮影していると、適宜教示が与えられる）を追加実装することが望ましい。

3.5 汎用性

本稿では3BCに特化したアプリケーションを開発した。ただし、映像をみて教示が可能、習熟状態を数値指標化できること、などの要件を満たすならば、他のスキルを扱うこともできる。例えば5ボールジャグリングは、メタ認知言語化や個人の状況にあわせた教示が特に難しい。そういったスキルではスキルサプリアの恩恵がさらに明確に現れると期待している。さらにジャグリングにかぎらず、応用先についても検討する。

4 結論

スキル習熟プロセスにおける、教示と身体動作の特徴量を対応付けることで、教示構造を明らかにするためのツールとして、スキルサプリアを作成した。また、スキルサプリアを用いることで分析できることについて議論を行った。習熟状態に応じ教示を利用することを目的として

スキルサブリを構成したが、教示に関する様々な知見を得るのに利用できる可能性がある。今後は、実際にスキルサブリを用い、データベースを構築し、教示に関する分析や、ユーザー体験の分析を中心に行っていく。

参考文献

- [1] 諏訪正樹. 身体知獲得のツールとしてのメタ認知的言語化 (< 特集 > スキルサイエンス). 人工知能, Vol. 20, No. 5, pp. 525–532, 2005.
- [2] 諏訪正樹. 身体が生み出すクリエイティブ. 2018.
- [3] 小玉耕平, 中山正吉. 体育における言語的教示に関する研究. 島根大学教育学部紀要 (教育科学), Vol. 23, No. 1, pp. 17–24, 1989.
- [4] 古川康一. スキルサイエンス入門—身体知の解明へのアプローチ. 人工知能学会, 2009.
- [5] NTT docomo. Golf ai. <https://golfai.jp/>.
- [6] 小坂田光, 多田昌裕. Android 端末を用いたリアルタイム安全運転アドバイス提供システムの試作. 映像情報メディア学会技術報告メディア工学, pp. 11–14. 一般社団法人映像情報メディア学会, 2016.
- [7] 国立研究開発法人科学技術振興機構. 経験サブリメントによる行動変容と創造的協働. https://www.jst.go.jp/kisoken/crest/project/1111083/1111083_09.html.
- [8] Sohei Wakisaka, Atsushi Hiyama, and Masahiko Inami. Transmission of experiences with augmented human techniques. In *UbiComp*, 2017.
- [9] 市川淳, 三輪和久, 寺井仁. 運動計測と言語報告に基づく身体スキル獲得に関する実験的検討. 人工知能学会論文誌, Vol. 30, No. 3, pp. 585–594, 2015.
- [10] PJ Beek and AAM van Santvoord. Learning the cascade juggle: A dynamical systems analysis. *Journal of Motor Behavior*, Vol. 24, No. 1, pp. 85–94, 1992.
- [11] Ken Hashizume and Tomoyuki Matsuo. Temporal and spatial factors reflecting performance improvement during learning three-ball cascade juggling. *Human Movement Science*, Vol. 23, No. 2, pp. 207–233, 2004.
- [12] 山本耕太. 3 ボールジャグリング学習過程における時空間的変数の変化. 3.
- [13] Pamela S Haibach, Gregory L Daniels, and Karl M Newell. Coordination changes in the early stages of learning to cascade juggle. *Human Movement Science*, Vol. 23, No. 2, pp. 185–206, 2004.
- [14] 高原慧一, 脇坂崇平, 荒川陸, 檜山敦, 稲見昌彦. 経験バンクに基づくジャグリング学習支援システムの提案. エンタテインメントコンピューティングシンポジウム 2020 論文集, Vol. 2020, pp. 199–203, 2020.
- [15] Zhe Cao, Tomas Simon, Shih-En Wei, and Yaser Sheikh. Realtime multi-person 2d pose estimation using part affinity fields. In *CVPR*, 2017.
- [16] Rasmus Åkerlund. Real-time localization of balls and hands in videos of juggling using a convolutional neural network, 2019.
- [17] Steffen Rendle. Factorization machines. In *2010 IEEE International conference on data mining*, pp. 995–1000. IEEE, 2010.
- [18] Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. ” why should i trust you?” explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, pp. 1135–1144, 2016.
- [19] John H Flavell. Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American psychologist*, Vol. 34, No. 10, p. 906, 1979.
- [20] Thomas O Nelson. Metamemory: A theoretical framework and new findings. In *Psychology of learning and motivation*, Vol. 26, pp. 125–173. Elsevier, 1990.
- [21] 田中彰吾, 小河原慶太. 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, Vol. 19, No. 1, pp. 69–82, 2010.
- [22] 内山光太, 伊藤毅志ほか. ジャグリングの熟達における思考過程の分析-3 ボールカスケードの事例より-. *SIG-SKL*, Vol. 22, No. 07, pp. 34–41, 2016.
- [23] David Duran. Learning-by-teaching. evidence and implications as a pedagogical mechanism. *Innovations in Education and Teaching International*, Vol. 54, No. 5, pp. 476–484, 2017.