

# 親子共学環境に向けたマルチモーダル知識オーサリングシステム

## Multimodal Knowledge Authoring System for Empowering Parent-Child Co-Education Environments

川口 紗季<sup>\*1</sup> 大竹 康太<sup>\*1</sup> 山本 剛<sup>\*1</sup> 石川 翔吾<sup>\*1</sup>  
Saki Kawaguchi Kota Otake Goh Yamamoto Shogo Ishikawa

桐山 伸也<sup>\*1</sup> 坂根 信一<sup>\*2</sup> 坂根 裕<sup>\*1</sup> 竹林 洋一<sup>\*1</sup>  
Shinya Kiriyama Shinichi Sakane Yutaka Sakane Yoichi Takebayashi

<sup>\*1</sup> 静岡大学  
Shizuoka University

<sup>\*2</sup> 七田チャイルドアカデミー北大阪教室  
Shichida Child Academy KitaOsaka Classroom

We have developed a system for holistic evaluation of parent-child education, based on collaborative participation of the parents, teachers and experts involved. First, we installed multiple cameras for observing and recording the children's behavior in small-group classroom setting. The audiovisual data recorded is then annotated, using natural language, by the members involved. The gathered data are shared over the Internet in form of multimodal knowledge that consists of movie clips and the annotations. So far, we have gathered over 600 annotated clips, which show that our method is suitable for acquiring practically usable multimodal knowledge that can be used for evaluating the quality of both the education, and the educational environment.

### 1. はじめに

幼児期は人間形成の基礎が培われる重要な時期である。近年、核家族化の進行などを背景に幼児期の子供への関わり方に不安を持つ親も増えており、その中で"良い"教育や教材に対する親の興味が高まっている[1]。教育の"良さ"の指標は親により異なり、子供の育つ環境に大きく影響する。筆者らは親への教育を重視した「親子共学環境」を実現するため、子供が教育を受ける場で、親も教育や子育てについて学べる教室をデザインしている[2]。

教育環境の改善や新しい教材のデザインには、教育現場での子供たちの活動の様子を観察し、実情の問題点を議論する必要がある。教師や教材開発者のみでなく、複数の立場からの意見を総合的に収集することが望ましい。しかし、親を含む様々な立場の人々が教育現場を主観的・客観的に観察し、現場での事例に基づいた議論が可能な環境は例がない。筆者らは、教育の場に関わる様々な人々の間での問題意識の共有や、異なる視点からの意見、各々が持つ知識の連携が重要と考える。

本稿では、筆者らの実践する親子共学環境の中で、教室に参加する親や教師、研究者などが、教室の様子を収録した映像を観察し、解説を加えるためのツールを開発した。入力データは、映像にテキストによる解説を加えたマルチモーダル知識として Web を介して利用者間で共有し、教材の評価や教育環境を改善するための議論の題材として利用する。

### 2. マルチモーダル知識の収集

本研究では、教育現場で子供と活動する親や教師、教育に関する専門的な知識を持つ研究者など、様々な視点からの観察結果を収集し、互いの知識を連携させると共に、実映像と連携させたマルチモーダル知識として提供し、三者間で意識共有を行うことを目的とする。

観察は、子供の行動や心の動き、教材と子供のインタラクション等に注目して行い、それぞれの着眼点での観察結果を記録する。実際に収集する知識には以下の4つの例がある。

- **親の視点** 家での経験や習慣を含めた幼児それぞれの特徴を考慮した解説。初めての行動や反応、いつもと違う行動などの発見

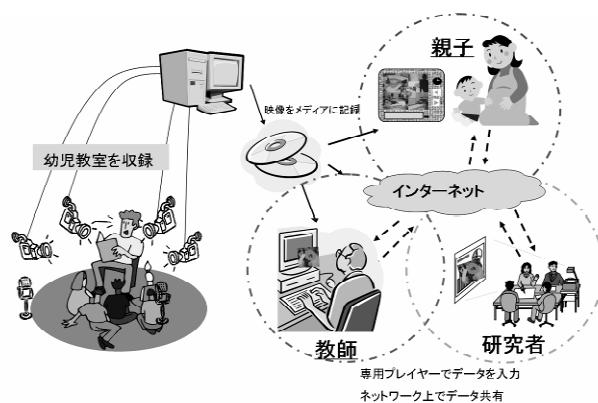


図1 観察から得られる知識収集モデル

- **教師の視点** 教室に来ている家族同士の関係や、それぞれの親子の関係を客観的に観察し、子供の発話や行動、表情や様子から、変化や成長した点を解説
- **教材開発者の視点** 教師がどの様に教材を利用しているか、幼児や親の反応はどうかなど経年的に観測、評価
- **研究者の視点** 言語や教育の専門家が幼児の行動を観察し、気になる行動や発話について、MinskyのEmotion Machine[3]の心のモデルと行動を関連付けて記述

図1に、現場の観察から得られる知識の収集モデルを示す。事例に関する知識を効率よく共有するため、解説は、その内容に関わる映像と関連付けられたマルチモーダル知識として提供する。筆者らのグループでは、映像や音声などのマルチメディアと、テキストなどで表現された知識メディアから構成される「マルチモーダル知識コンテンツ」を、サイクリングや空手、登山などの複数のドメインで収集する取り組みを行っている[4]。コンテンツ作成技術を応用し、幼児教室を収録した映像と、それに対する解説からなるマルチモーダル知識として蓄積・利用することで、知識の共有を円滑化すると共に、子供の成長の過程や教材への反応などを経年的に記録したマルチモーダル知識ベースを構築することができる。

教育現場で知識を効率よく収集、共有するため次の3つの問題点を解決する必要がある。

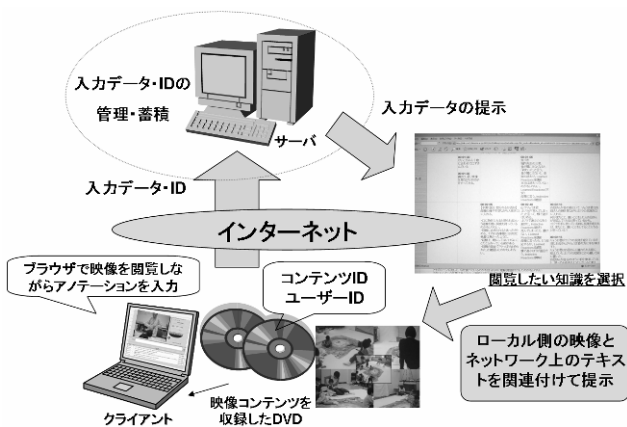


図2 システム構成



図3 実装したシステム

### (1) 収集する知識データ

記述する内容は非常に複雑であり、知識表現言語や形式化された文法を用いて表現することは難しい。また入力者の専門知識の有無も様々であるため、誰でも手軽に複雑な記述が可能な自然言語を用いる。

### (2) 作業負担の軽減

子供が実際にした行動や発話に対し、幅広い人々の視点からの観察を行うため、各々の観察者の視点に合った映像を提供する必要がある。このため映像はマルチカメラで収録する。複数の映像ソースから、見たいシーン・アングルを探し、その映像に対してデータを入力するなどの作業は本来非常に時間がかかる。幅広い人からの知識の収集を実現するためには、誰もが利用できる環境で作業可能で、環境作業負担の軽減を考慮したインタフェースを備える必要がある。

### (3) 映像とテキストを含むマルチモーダル知識の共有

入力の際に他者が入力した情報を参照し合うことで、互いに情報を補うなど、連携したアノテーションが入力されることを狙う。このため、入力した知識はリアルタイムに共有できる必要がある。時間のかかる編集などをせずに、入力と同時にマルチモーダル知識として他者に提供できる仕組みが必要となる。

## 3. システムの試作と試用

### 3.1 システムの実装

映像の提示と知識の入力・閲覧を行う専用ブラウザを開発した。システムの構成図を図2、実装したシステムを図3に示す。

映像は、多視点映像や複数の音声を手軽に切り替えて視聴することが可能なDVDメディアに記録して配布する。専用のブラウザをDVDに同梱しており、DVDを視聴可能なPC環境で利用できる。DVDとPC、インターネット環境があれば、どこからでも知識オーサリングを行うことができる。

入力者はブラウザを利用して映像を視聴しつつ、視点を切りかえる、気になったところで停止、巻き戻し、早送り等の作業しながらテキストを入力することができる。テキスト情報を入力する際、映像コンテンツの種類と映像中のタイムコード、入力者情報が一緒に記録されサーバ上で管理される。入力された知識はネットワークを介して共有するが、映像データのネットワーク上での共有は大きな帯域を必要とするため実用的でない。収集した知識は、該当する映像コンテンツが記録されたメディアをPCに挿入すれば、インターネット上で自然言語による解説として一覧できる。解説を選択すると、関連する映像をメディアから読み出し

して再生する。ローカルの映像データとネットワーク上のテキストデータと連携させて提示することで実用的な知識として配信する。

### 3.2 システムを用いた知識収集と評価

実際の幼児教室に多視点カメラを置いて約10時間の映像を収録した。システムを利用し、幼児教室に通う母親、教室で教える教師、筆者らを含む研究者が、映像に対するアノテーションを行った結果、約600個が収集された。

本システムを用いることにより、子供の行動や表情の変化、新しい発話やなど、一連の映像の中でそれぞれの立場から見た重要箇所を、容易に蓄積できた。またインターネットを利用して、様々な場所にいる利用者が、同じ観察環境において連携的に作業することを可能とした。

## 4. まとめ

本稿では、幼児の発話や行動に着目した観察に基づく知識を、様々な立場の人から総合的に収集し、実映像と自然言語からなるマルチモーダル知識として利用するためのシステムを開発した。実際にシステムを試用し、実用的なマルチモーダル知識を収集できる見通しを得た。

今後は入力された知識を解析し、有効な知識が収集されているか検証すると共に、収集されたマルチモーダル知識に対し検索などが行える知識ベースとして利用できるよう改良する。

## 謝辞

本研究の遂行にご協力いただいたMITのMarvin Minsky教授とPush Singh氏、また静岡大学情報学部の堀内裕晃教授に謝意の意を示す。

## 参考文献

[1] 文部科学省中央教育審議会: "子供を取り巻く環境の変化を踏まえた今後の幼児教育の在り方について", [http://www.mext.go.jp/benu\\_shingi/chukyo/chukyo0/toushin/04102701.html](http://www.mext.go.jp/benu_shingi/chukyo/chukyo0/toushin/04102701.html).

[2] 川口, 坂根, 坂根, 竹林: "親子のコミュニケーションを促進する幼児教育の一検討", 第67情報処理学会全国大会, 5A-2, 2005.

[3] M Minsky: "The Emotion Machine (draft)", <http://web.media.mit.edu/~minsky/>.

[4] 吉滝, 田森, 坂根, 竹林: "サイクリスト支援のためのマルチモーダルナレッジとコミュニケーション環境のデザイン", 第18回人工知能学会全国大会, 1B3-05, 2004.