

## A study of presentation system for e-learning

有安香子

山田一郎

住吉英樹

柴田正啓

八木伸行

Kyoko Ariyasu

Ichiro Yamada

Hideki Sumiyoshi

Masahiro Shibata

Nobuyuki Yagi

## 1. はじめに

高速大容量ネットワークの整備・普及に伴い、教育の分野においてもインターネットの高速性・双方向性を活用した遠隔教育が活発になってきている。そこで我々は、放送局の持つ豊富な映像資料、情報から得られる知識、経験を元にした映像表現能力を活かし、統合的な教育を支援するシステムの研究を行っている[1]。その1つとして、今回、遠隔グループ学習システムを試作したので報告する。

本システムでは、ネットワークに接続した端末から、チャット形式で入力した各学習者の発言を、仮想空間内での討論番組として表示しグループ学習を行う。この仮想教室でのグループ学習を中心とし、必要に応じて授業内容に関連した番組を視聴し、わからない事はグループ学習に参加しているエージェントアバターに質問するといった形で、知識の向上・学習意欲の活性化などの学習効果の促進をはかる学習システムである。

## 2. システムの概要

今回試作したシステムの概要を図1に示す。

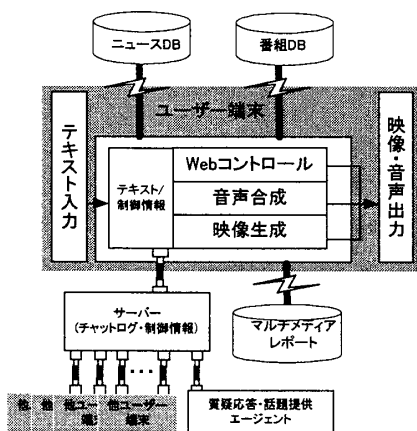


図1: 遠隔グループ学習プレゼンテーションシステム概要

本システムでは、学習者の発言内容がテキストでリアルタイムに送受信され、CGアバターを使った仮想空間上で討論番組形式の映像音声として出力される。このテキストデータと共に、必要に応じてURLなどがユーザー端末の表示をコントロールする制御情報と共に送受信され、自然な流れの中で番組視聴やWebによる調べ学習を行うことができる。

また、サーバーはチャット内容を随時エージェントに送信しており、エージェントはこの内容から話題推定を行い、会話が途切れた際の話題提供[2]や質疑応答などを行っている。

更に、マルチメディアレポート作成ツール[3]をシステムに組み込むことで、学習した内容をまとめ、ネットワーク上に公開することも可能となり、「見る・調べる・話し合う・まとめる」といった統合的な学習を支援するシステムを構築している。

以下に、チャットによる話し合いとエージェントの話題提供・質疑応答を、仮想空間上に討論番組形式で映像音声化する手法について詳しく述べる。

## 3. 仮想空間内での討論番組映像自動生成 (仮想教室)

ネットワークを通じたグループ学習では、グループの構成要因が動的であるため、テキストの送受信だけでは発言者を特定しにくいという問題点がある。そこで、学習参加者のIDと発言内容を元に、発言者の特定がしやすく、見ていて飽きの来ない映像を自動付加するため、実際に放送された討論番組(42対談、30時間、9000カット)を様々な角度から分析し、映像演出の統計的算出を行いこれを実装した。



図2: 仮想教室画面例

システムの出力映像は図2に示すように、エージェントと各参加者がCGアバターとして仮想教室上で討論番組を行う形式のものである。この様な対談番組形式で、発言者が特定しやすく、飽きの来ない演出を行うために特に重要な、ショットの決定とスイッチングタイミングについて、発言者に着目して分析を行った結果を以下に述べる。

## 3-1. ショットの決定

討論番組など、会話を主体とする番組の構成の基本は発話者のワンショットなので、当然発言開始時には発話者のワンショットが映される確率が高く、その後挿入されていくショットの決定は、様々な要素と密接に絡み合っているが、一般に直前のショットとの関係が高い。

そこで、ショットを話者ワンショット・話者周辺ショット・

ドリーショット・参話者(発言者以外の討論参加者)周辺ショット・参話者ワンショットの5つにわけ、発言開始時と、その後の各ショット間の遷移の確率を調査し、結果を表1に記す。

前	後	発言開始	話者1S	話者周辺	ドリー	参話周辺	参話1S
話者1S		70%	11%	85%	73%	77%	72%
話者周辺		19%	32%	4%	11%	14%	6%
ドリー		5%	8%	1%	1%	2%	1%
参話周辺		1%	15%	2%	4%	3%	2%
参話1S		5%	35%	8%	11%	4%	19%

表1: ショット種類の直前ショットとの関係

本システムでは、各ショットの出現頻度がこれらの確率に近くなるよう、ショットを切り替えるタイミングごとに乱数で決定した。

### 3-2. スイッチングタイミングの決定

次に、3-1で決定した各ショットの持続時間(スイッチングタイミング)の算出方法について検討を行った。実際に放送された番組データのショット切り替えタイミングとショット種類の関係を見ると、相関係数は0.42であり、あまり相関が高いとは言えない。そこで、スイッチングタイミングを決定するために必要な、ショット種類以外の主要な要因を洗い出すため、様々な要因を用いて数量化一類により分析を行った。その結果、表2に示す発話の長さ・画面効果・ショット種別・ジェスチャや名前引用などの要因を用いてスイッチングタイミングを決定すると重相関係数0.83(寄与率69%)でスイッチングタイミングを予測出来ることがわかった(表2には、上記各要因のカテゴリ分けとカテゴリ数量を示す。)

アイテム	カテゴリ		カテゴリ数量	
発話の長さ	X11	0s~30s	A1	-0.75432
	X12	30s~60s	A2	-0.0639
	X13	60s~90s	A3	0.393439
	X14	90s~120s	A4	0.517815
	X15	120s~180s	A5	0.409238
	X16	180s以上	A6	1.721694
画面効果有無	X21	スーパー	B1	10.86068
	X22	フリップ	B2	6.077252
	X23	なし	B3	-0.68935
ショット種別	X31	話者1S	C1	2.219515
	X32	話者込み周辺1	C2	-1.32513
	X33	話者込み周辺2	C3	-2.06837
	X34	話者込みドリー	C4	2.320856
	X35	参話者周辺	C5	-3.28907
	X36	参話者ドリー	C6	0.209537
	X37	フォロー(割込)	C7	-3.94285
	X38	フォロー(身振)	C8	-2.81013
	X39	フォロー(表情)	C9	-1.8129
	X30	その他	C0	-4.69
その他要因	X41	ジェスチャ	D1	-2.8671
	X42	名前引用	D2	-2.8549
	X43	ドリー→話者	D3	9.143352
	X44	割り込み	D4	-4.59026
	X45	議論	D5	-7.75546
	X46	ドリー→全体	D6	9.245671
	X47	構成上	D7	20.77709
	X48	なし	D8	-0.36744
定数項		E	-0.36744	

表2: スイッチングタイミング決定のカテゴリと数量

この結果をもとに、

- i) 入力テキストの長さを標準的読速(150文字/分)で計算し発話の長さを計算した結果で発話の長さをカテゴリわけする。
- ii) 各学習者の初回の発話時には画面効果としてスーパーを行い、画面効果のスーパーのカテゴリに場合わけする。入力テ

キスト内に参考URLが存在するときにはフリップカテゴリに場合わけする。

iii) 3-1の遷移確率にしたがってショット種別のカテゴリわけを行う。

iv) 入力GUIの感情表現ボタンの押下による感情表現入力が行われたときはジェスチャをキャラクタに付与しその他要因をジェスチャカテゴリにする。入力テキストと学習者の名前のパターンマッチングを行い、名前引用があった場合にはその他カテゴリを名前引用に、入力テキスト同士のパターンマッチングを行い他者発言引用があった場合にはその他カテゴリを発言引用にする。

という方法で具体的に入力テキストを場合わけし、各要因の該当カテゴリを1(それ以外は0)とし、スイッチングタイミングを予測するために各カテゴリ数量を係数に用いて予測式

$$\sum_{i=1}^6 A_i X_{1i} + \sum_{j=1}^3 B_j X_{2j} + \sum_{k=1}^{10} C_k X_{3k} + \sum_{l=1}^8 D_l X_{4l} + E \quad (\text{秒})$$

によりスイッチングタイミングを決定した。各ショットに切り替わった時間を0とし、決定したスイッチングタイミングの時間が経過した後、発言が終わっていない場合には、3-1に基づき、新たに挿入ショットをきめ、そのショットのスイッチングタイミングを決定する、という方法を発言が終わるまで繰り返すことで実装した。

### 4. まとめ

以上に述べた映像生成規則を用いて、学習者の発言テキストと、エージェントの質問応答・話題提供テキストを、仮想教室上にマッピングしたCGキャラクタをアバタとしたグループ学習映像としてシステム上で表現した。これにより、発言者の適切なフォロー、話題参加者を含めた適切な周辺ショット、などを気ぜわしくなりすぎない適切なタイミングで表現し、学習者の興味をひきつけるような映像を自動付加することが出来た。システムの実装にはTVML[4]の外部制御機能を用い、リアルタイムでカメラスイッチングやCGキャラクターのジェスチャや感情表現などを生成し、映像音声の出力を行った。

また、多くの知識と検索能力を持ったエージェントをグループ学習に擬人化オブジェクトとして参加させることで、他の学習者との交流による知識の向上・学習意欲の活性化など、グループ学習を行なうことにより期待される効果をよりいっそう高める事ができた。

今後は、会話テキストの映像化手法の入力対象となるテキストの範囲を、解説文など汎用的なものに広げ、テキストを映像化する際に必要な付加情報の洗い出しを行い、コンテンツ記述方式の必要要件の洗い出しへと、発展させていく予定である。

### 参考文献:

- [1]NHK デジタル教材「おこめ」 <http://www.nhk.or.jp/okome>
- [2]山田ほか: “仮想教室上の学習コミュニティにおける発想支援エージェントの検討”, FIT2002 情報科学技術フォーラム, Sep, (2002)
- [3]住吉ほか: “ネットワークによる意見交換を目的とした映像レポート作成システム”, 映像情報メディア学会メディア工学研究会, ME2002-8, July, (2002)
- [4] <http://www.str1.nhk.or.jp/TVML/>