

修士論文

教材指示情報を付与した
講義コンテンツの作成

指導教官 美濃 導彦 教授

京都大学大学院情報学研究科
修士課程知能情報学専攻

丸谷 宜史

平成 16 年 2 月 13 日

教材指示情報を付与した講義コンテンツの作成

丸谷 宜史

内容梗概

本研究では、講義中に提示された表示教材への講師の指示に関する情報を講義コンテンツに付与し、提示する手法について提案する。講義コンテンツとは、一人の講師が複数の受講者に教示内容を講述する一斉講義型の対面講義の様子をアーカイブ化したものである。

対面講義では受講者は講師や表示教材と空間を共有しており、講師の指示と表示教材間の関係を把握しながら教示内容を理解できる。一方、国内外の大学で既に配信されている講義コンテンツは、講義の様子を講師のみを撮影した映像と表示教材の映像とを同期をとって受講者に同時に提示しているが、講師の教示内容と表示教材上の指示対象との関係の直感的な把握が困難であり、利用者に教示内容の理解だけでなく関係把握にも気を使わせねばならない問題がある。

そこで本研究では、講師が講義のどの時刻で表示教材のどの領域を指示しているかという情報(教材指示情報)を抽出して講義コンテンツに付与し、受講者に提示することで教示内容と表示教材との関係を伝達することを目標とする。

教材指示情報は、指示している領域の特定と、指示している時間区間の特定により抽出できる。領域の特定には表示教材内で講師の意図する指示の領域が講師の教材への4種類の指示動作に対応するという教育学分野の研究成果から、本研究では指示動作の分類を領域の特定に用いる。

指示動作の分類は、ある時間区間に検出される指示動作を構成している指示点等の特徴と、予め定義した時間区間毎の各指示動作モデルとを比較して、最も類似している指示動作であると判断することで達成する。また、このときの時間区間の終了時刻は指示領域が切り替わる時刻でもあるため、上記の方法で教材指示情報の特定が可能になる。

教材指示情報を付与した講義コンテンツの提示は、従来の表示教材の映像において表示教材中の指示領域を指示動作の種類に応じた表現で強調表示することにより実現した。実験の結果、分類は平均82.6%の精度で分類ができ、また受講者による教材指示の把握もアンケート結果から従来のものよりもわかりやすくなったことを確認した。

Making a lecture contents added deictic informations about lecture materials

Takafumi MARUTANI

Abstract

In this paper we propose a method to assign deictic information for computer presentations to lecture contents in electronic archives.

Existing lecture contents carry a lecture situation by displaying a movie of a lecturer and lecture materials synchronically. However, students cannot grasp the relationship between lecture content and objects on a slide through existing archives. Therefore they have difficulty in understanding the lecture content compared with the situation when they attend the class in person.

This research extracts information when and where the lecturer points (deictic information about lecture materials) and overlays it to lecture content. The contents with the deictic information enhance the students' understanding of a lecture.

As a starting point of our research we studied the Lilians' group research results on deictic information. According to their research results, there are four deictic gestures which point different area. Therefore, the information of points the lecturer indicates can be determined by means of classifying time series data. Students can grasp the relationship between lecture content and lecture item on slides because the lecture content is displayed with the classification.

We verified the effectiveness of our method by the success rate of classification and the questionnaire given to students in order to estimate the understanding of the relationship between lecture content and objects on a slide.

教材指示情報を付与した講義コンテンツの作成

目次

第1章	緒論	1
第2章	講義コンテンツのための教材指示情報	3
2.1	対面講義と従来の講義コンテンツで受け取る情報の違い	3
2.1.1	対面講義における教示内容の伝達	3
2.1.2	従来の講義コンテンツによる教示内容の伝達	3
2.1.3	対面講義と従来の講義コンテンツで伝達される情報の 違い	5
2.2	講義における教材指示情報	6
2.2.1	教材指示情報とは	6
2.2.2	教材指示情報の抽出	7
2.2.3	教材指示情報の提示	11
第3章	教材指示情報の付与と提示	14
3.1	教材指示情報の付与	14
3.2	スライドの領域分割	16
3.3	指示意図領域と指示領域の変更時刻の推定	17
3.3.1	指示領域の変更時刻の検出のための評価関数	18
3.3.2	指示動作の分類	18
3.3.3	指示点の軌跡の評価	20
3.3.4	指示点の存在する領域の種類の評価関数	25
3.3.5	講師の体の向きに関する評価関数	25
3.3.6	講師の体の位置に関する評価関数	27
3.3.7	指示意図領域と指示領域の変更時刻の決定	27
3.4	指示領域の決定	29
3.5	アーカイブに付与された教材指示情報の提示	30
第4章	実験と評価	33
4.1	実験の目的	33
4.2	実験環境	33
4.3	予備実験	33

4.3.1	手動抽出した教材指示情報の評価実験	33
4.3.2	講義中の講師の指示に関する実験と検討	37
4.4	教材指示情報抽出に関する実験と評価	37
4.5	教材指示情報の提示の実験と評価	39
第5章	結論	43
	謝辞	45
	参考文献	46

第1章 緒論

計算機の記憶容量や計算能力の向上, 通信ネットワークの広帯域化により, ネットワークを通じて様々な種類の映像コンテンツが利用されるようになってきた. 大学教育の場においては, 映像コンテンツは Web 教材の一部として [1], また遠隔講義 [2] や講義アーカイブ [3][4] で利用されている.

遠隔講義や講義アーカイブでは, 一人の講師が複数の受講者に教示内容を講述する一斉講義型の対面講義を配信や記録の対象としている. 対面講義では講師が受講者に対して, 発話や動作, 講義室前面のスクリーン等に表示した教材 (表示教材) を通じて講義内容を伝達する. 本研究では講師が受講者に語りかけたり, 表示教材を指示しながら解説する対面講義の様子の記録と伝達を目的とし, 講義室の様子を記録したものを講義コンテンツと呼ぶ.

講義コンテンツを提示する際に複数の映像間で時間同期を取ることの重要性は従来の講義アーカイブに関する研究において指摘されている [3]. その点に関しては SOI (School Of Internet) [7] に代表される従来の講義コンテンツでも反映されており, 講師のみを撮影した映像 (講師映像) と OHP やスライドのような表示教材の映像 (表示教材映像) を同時に提示することにより視聴者の講義の様子を伝達している. しかしながら, このような2つの映像の同時提示のみの枠組みでは, 講義コンテンツの視聴者は講師と表示教材の関係を把握することが困難であるため, 視聴者は講師が表示教材のどの部分に関して説明しているかという情報 (教材指示情報) を推測して理解する必要があり, 視聴者にとって負担が大きい.

対面講義において講師と表示教材との関係が把握しやすい状況とは, 講義中に講師の行う4種類の動作 (beats・iconics・metaphorics・deictics) の中の特に deictic な動作 (指示動作) が講師により表示教材に対して行われ, 表示教材内の特定対象への注目が誘導される状況である [6]. しかしながら, 上で述べたような従来の講義アーカイブでは, 講師が講義のどの時点で, 表示教材内のどこを指示しているかがわからないため, 講師と表示資料との関係がわからない. そこで本研究では, 対面講義から教材指示情報を抽出して講義コンテンツに付与することを目指す.

講義コンテンツに教材指示情報を付与するには, 講師がどの時間区間に, 表示教材内のどの領域を指示しているかを特定すればよい.

講義中の講師の教材への指示動作は Lilian らの研究 [8] によると 4 種類 (Emphasizing・Highlighting・Pointing・Outlining) に分けられることが知られている。これらの動作の違いはそれぞれ講師が受講者に伝えたい表示教材内の指示対象の抽象度が異なるために生じるものであり、それに応じて表示教材内で指示する領域は異なる。したがって、講師が表示教材内のどの領域を指示しているかを特定するためには、講師の動作を上述べた 4 種類の指示動作に分類すればよい。

指示動作の分類に関しては Lilian らの研究 [8] で紹介されている指示動作を規定する特徴を利用し、任意時間区間に検出されるそれらの特徴と、予め定義した 4 種類の任意時間区間毎の指示動作モデルをそれぞれ比較し、ある時間区間における指示動作モデルと最も類似していると判断された指示動作に分類する。また、このときの時間区間の終了時刻が、指示動作が切り替わるタイミングであり、指示領域が切り替わるタイミングでもある。

講義コンテンツに付与した教材指示情報は、講義アーカイブの表示教材映像を強調表示することによって視聴者に伝達する。表示教材映像への教材指示情報の強調表示は、従来の表示教材映像の中で、表示教材中の指示領域を指示動作の種類に応じた表現で書き換えたものを置き換えることにより実現する。

本論文の構成は以下の通りである。2 章では講義コンテンツのための教材指示情報についてその分類と提示について述べる。3 章では教材指示情報の付与と提示の手法について述べ、4 章では本研究で行う実験の環境と実験の概要について述べる。5 章では本研究についてのまとめと将来的な課題について述べる。

第2章 講義コンテンツのための教材指示情報

講義室において対面講義を受講する場合と講義コンテンツを視聴する場合では、対面講義の受講者と講義コンテンツの視聴者が受け取ることが可能な情報量は異なる。本章ではまず対面講義の受講者と講義コンテンツの視聴者が受け取ることが可能な情報量について検討し、次に講義コンテンツにその違いを補うために付与する必要がある情報について議論する。

2.1 対面講義と従来の講義コンテンツで受け取る情報の違い

本節では対面講義の受講者と従来の講義コンテンツの視聴者が受け取る情報の違いを明らかにする。

2.1.1 対面講義における教示内容の伝達

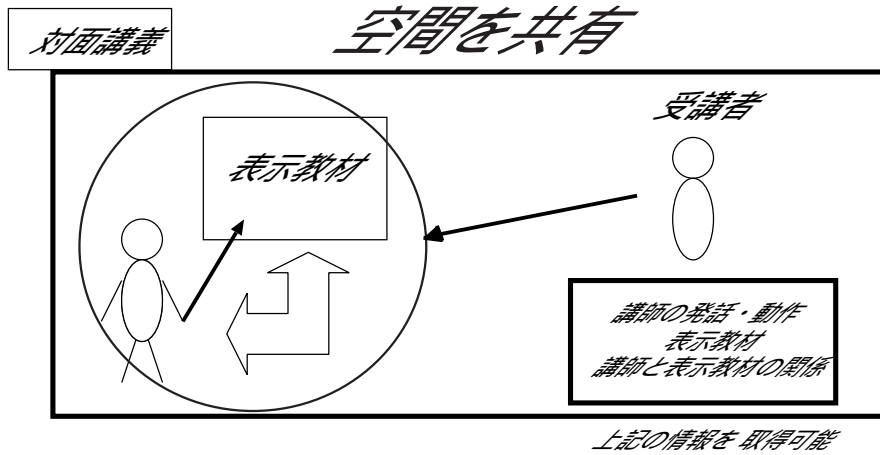
講義室における対面講義では、講師は受講者に語りかけたり、表示教材を指示しながら解説することにより教示内容を伝達する。一般に、コミュニケーションの場において、聞き手は話し手の発話と同時に行われる動作をリアルタイムに解釈することで、発話単体や動作単体のときよりも相手の意図を理解することが知られている [5]。

対面講義の受講者は、講師と空間を共有しているため、一般のコミュニケーションの場合と同様に講師の発話と動作を同時に確認することができ、動作に応じたタイミングで表示教材を確認し、講師と表示教材との関係を把握することができる。そのため、受講者が教示内容の理解のために用いることができる情報は、図 1(a) で示すように、講師発話・動作、表示教材、講師と表示教材の関係の3つの情報である。

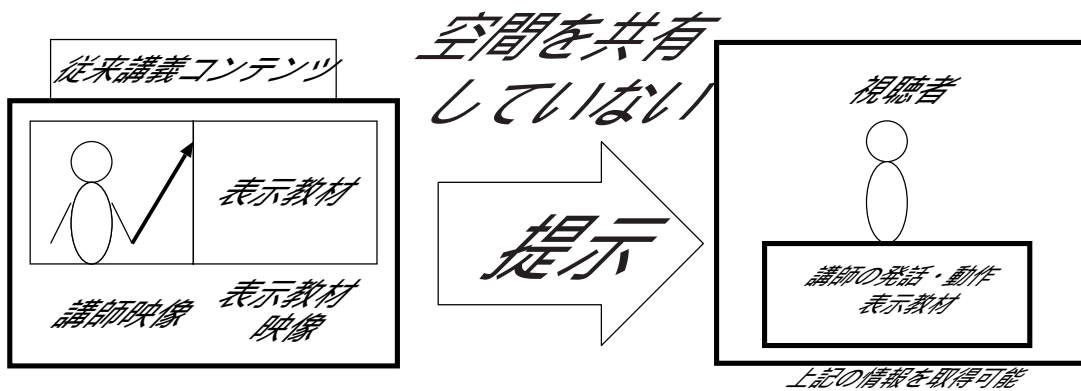
2.1.2 従来の講義コンテンツによる教示内容の伝達

従来の講義コンテンツでは、講師映像、表示教材映像を同時に視聴者に提示することで教示内容の伝達を行う。ここでいう講師映像とは講師の動作や表情がわかりやすいように講師のみを撮影した映像である。また、表示教材映像とは表示教材をビデオカメラなどを利用して電子的に記録した映像である。

講義コンテンツの視聴者には、講師の発話・動作、及び表示教材の更新が同期して提示されるようになっており、図 1(b) で示すように一般のコミュニケーションの場合と同様に教示内容を理解するために講師の発話・動作を同時に確認することができる。ただし、講師映像と表示教材映像は講義中の講師の動作、



(a) 対面講義の場合



(b) 従来講義コンテンツの場合

図1: 対面講義と従来講義コンテンツで受講者が受け取ることができる情報

表示教材を別々に表現メディア化したものであり、かつ講師と視聴者は空間を共有していないため、対面講義の場合と異なり、講師の動作と表示教材間の関係に関して確認することができない。そのため、従来の講義コンテンツの視聴者が教示内容を理解するために用いることができる情報は講師の発話・動作、表示教材の2つの情報である。従来の講義コンテンツの提示例を図2に示す。

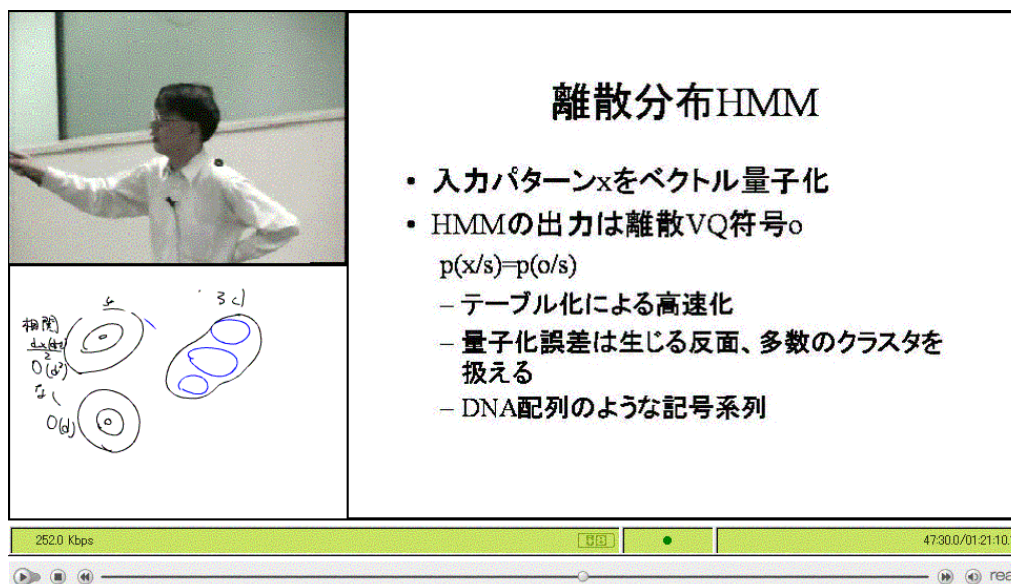


図2: 講義コンテンツの提示例

2.1.3 対面講義と従来の講義コンテンツで伝達される情報の違い

2.1.1 節で考えた対面講義で受講者が把握することができる情報と、2.1.2 節で考えた従来の講義コンテンツで視聴者が把握できる情報を比較すると、従来の講義コンテンツでは講師と表示教材の関係の情報が得られないことがわかる。講師と表示教材の関係の情報は、講義中の教示内容と表示教材を結びつけ、教示内容の理解を促すために重要な情報であるので、講義コンテンツにおいてその情報が伝達されるようにする必要がある。そこで次節では、講師と表示教材の関係の情報の伝達について考察する。

2.2 講義における教材指示情報

2.2.1 教材指示情報とは

対面講義にて講師と表示教材との関係が伝達される状況は、発話と共に講師自身が明示的に表示教材との関わりを示す動作を行う状況である。発話と共に行われる動作は McNeil[6] によると、以下の4つの性質に分類される。

- beats 現在の発話内容を強調するための動作。
- iconics 腕や手の動きにより実在のものや出来事を表現する動作。
- metaphorics iconics と類似しているが、抽象的な概念を視覚的に表現する動作。
- deictics 相手の注目を向けるため、実在もしくは抽象的な対象を指し示す動作。

上記の分類の中で、講師自身が明示的に表示教材との関わりを示す動作は deictics な動作（指示動作）であり、指示動作は教材中の特定の対象に受講者の注目を向けさせることを目的としている。

つまり、講義の各時点で、講師が表示教材中のどの対象に注目させたいかが視聴者にわかる形で明示すれば、講師と表示教材との関係情報が伝達できると考えられる。そこで本研究では、講義の各時点で表示教材のどこを指しているかという情報（教材指示情報）を抽出して講義コンテンツに付与することにより、従来の講義コンテンツでは得ることができなかった講師と表示教材との関係の情報を視聴者に提示し、教示内容の理解を助けることを目標とする。

教材指示情報とは、前述のように、講師が表示教材中のどの領域を、どの時間区間指示しているかという情報のことである。指示している表示教材の領域は、現在の話題が、表示教材中の図の形状のような具体的な対象から、表示教材中の大体このあたりというような抽象的な対象までのように、講師の指示が示す表示教材中の対象の抽象度と具体的な位置を示すものである。また、指示している時間区間は、講師の指示動作の開始時刻終了時刻の間の区間であり、講師の動作と表示教材が関わっているタイミングを示す情報である。

そこで、講義コンテンツに教材指示情報を付与し、講師と表示資料との関係を伝達させるためには、指示動作で表示教材のどこを指示しているかを特定し、いつその指示動作が行われているかという情報を明示する必要がある。実際に、後述する予備実験の結果からも講義の各時点で指示している領域を明確にした

ほうがわかりやすいという結果が得られているため、本研究では講義中に講師が指示している領域及び指示時間区間が教材指示情報の要素であるとし、それらを抽出して従来の講義コンテンツに付与することを考える。

2.2.2 教材指示情報の抽出

2.2.1 節で述べたように、教材指示情報は以下の2つの要素で表現される。

- 指示動作により指示される領域（指示領域）
- 指示動作が行われている時間区間

まず、指示領域の特定について説明する。指示領域とは、ある時点で表示されているスライドで文や図といった表示教材の構成要素の存在する領域と、講師がある対象への指示を意図して行った動作から推定される領域（指示意図領域）との積領域によって決定できる領域のことであり、本研究では指示領域の粒度は指示意図領域に応じて表1のように定義する。

指示領域の特定は図3に示すようにして行う。まず、ある時点での講師の表示教材への指示動作を取得する。次に、その時点で表示されているスライドを文や図といった構成要素ごとへの領域分割と、講師の指示動作から推定される教師がどの表示教材のどのあたりを指示しようとしていたか、という指示意図領域の推定を行う。最後に、分割されたスライド内の領域と推定された指示意図領域の積領域を求め、それを指示領域とする。

そこで指示領域の特定のために、表示教材の構成要素ごとへの領域分割と講師の指示意図領域の推定について以下で説明する。なお、図3、表1中の文領域、図領域についてはスライドの構成要素ごとへの領域分割の説明のときに定義する。

表1: 指示意図領域と指示領域の粒度の対応

指示意図領域		指示領域の粒度
文領域	図領域	粗
文領域の一部	図領域の一部	
-	形状	
語句	点	細

最初にスライドの構成要素ごとへの領域分割について説明する。

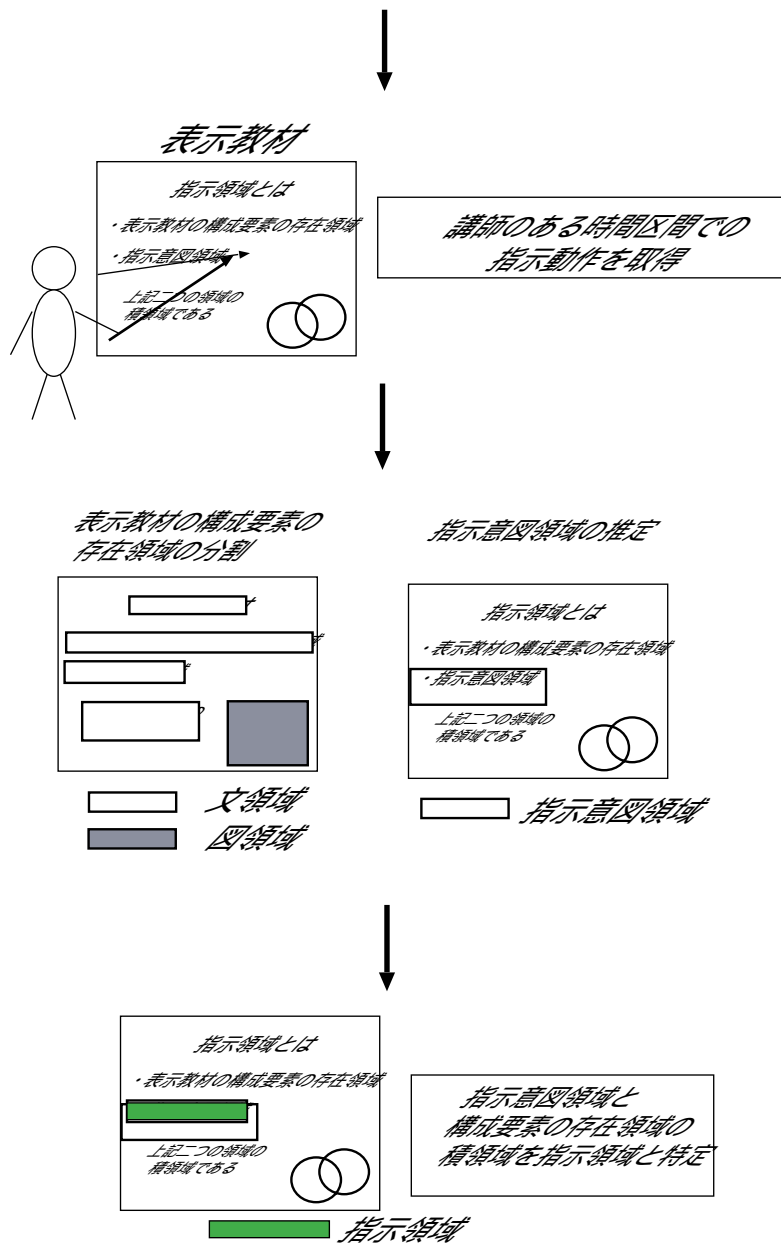


図 3: 指示領域の特定

対面講義の各時点で提示されるスライドは、複数の文や図が配置されることで構成されている。本研究で想定している表示教材への講師の指示において一番粒度が粗いものは、一文や一つの図といったおおまかな対象であり、指示の意図により、指示する対象が文の一部であったり、図の一部であったりというように粒度が異なってくる。本研究では、講師の指示で一番粒度が粗いものが表示教材中に存在する領域を文や図かという性質に応じて、領域分割し、それぞれを文領域・図領域と呼ぶ。

この文領域や図領域は指示領域の粒度として本研究で想定している指示領域の中で一番粗いものなので、実際に指示される領域は文領域・図領域の中に含まれる。そして、文領域や図領域と講師の動作から推定される指示意図領域との積領域が指示領域であるとする。

次に、指示意図領域の推定について述べる。本研究では、指示意図領域の特定のために、Lilian らの研究 [8] による指示動作分類を利用する。Lilian らの研究の研究では指示動作は以下に述べる 4 種類に分類される。

- emphasizing(図 4(a))
スライド内の文・図の構成の強調・明確化の働きがあり、その動作で意図される指示領域は、スライド内の文領域（一つの項目、一文）や図領域（一つの図）である。
- highlighting(図 4(b))
表示教材内から特定の語句や領域を切り出し注目対象を限定化する働きがあり、その指示で意図される指示領域は、動作で切り出される文の一部や図の一部である。
- pointing(図 4(c))
注目対象を固定、継続させる働きがあり、その指示で意図される指示領域は語句や点、以前の動作で示していた領域になる。
- outlining(図 4(d))
この動作はとくに表示教材の図の形状特徴を記述する働きがあり、この動作で意図される指示領域はその指示を行った軌跡が示す形状になる。

上記の 4 種類の指示動作は、それぞれが意味することが異なっており、またその指示動作により意図される指示の領域は表 1 で示した指示領域の構成要素に対応している。そのため、指示意図領域の推定は指示動作を上述の 4 種類に分類することで特定できる。

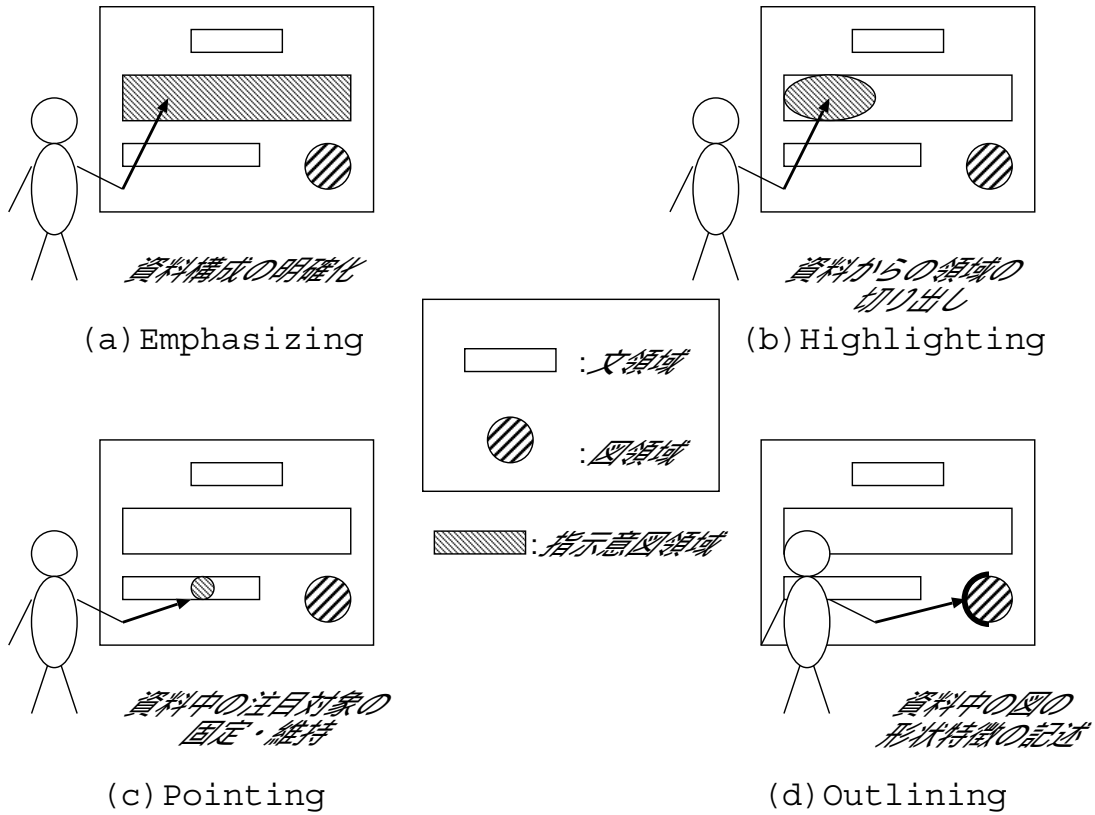


図 4: 4 種類の指示動作による指示意図領域

指示動作の分類に関しては、Lilian らの研究で、以下の特徴に基づき分類されることがわかっているのでそれを利用する。分類の詳細については次章で述べる。

- 指示棒先端の軌跡がどうなっているか
- どの種類の領域を指しているか
- 講師がどの方向をむいているか
- 講師がどこにいるか

次は、指示動作の行われている時間区間の特定について考察する。指示動作の行われている時間区間は、各時点で指示されている領域を明確にするための情報である。図 5 は横軸を講義の時間軸とし、講義中の講師の動作と表示教材の時間軸における関わりを示したものである。図 5(a) で示すように、対面講義では任意時点の講師がどのような指示を行っていてどこを指示しているかはリアルタイムに解釈が可能である。それに対し、図 5(b) で示すように、本研究における講義コンテンツでは教材指示情報が提示されている時間区間でしか講師の指示に関しては判断できない。そのため、講義の任意の時刻を取り出したときに対面講義で解釈できる講師の指示情報と講義コンテンツで得ることができる指示情報が同じになるためには、対面講義で指示動作が行われている区間と教材指示情報を提示する時間帯が正しく対応していなければならない。そこで本研究では、講義のビデオを観察した結果から指示が行われている時間区間は指示棒の先端が教示教材にほぼ接触させているときとし、指示棒を接触させている時間区間において、指示領域の変更時刻を検出することで指示動作の開始時間と終了時間を検出し、指示動作の行われている時間区間を特定する。なお指示動作がなされていないときは、何の処理も行わないものとする。

2.2.3 教材指示情報の提示

2.2.2 節で得られる教材指示情報を付与した講義コンテンツにより情報が伝達される様子を、図 6 に示す。

従来の講義コンテンツの視聴者が教示内容の理解のために受け取ることができる情報は、2.1.2 節で説明したように講師の発話・動作及び表示教材情報である。

それに対し、本研究の講義コンテンツは、得られた教材指示情報を図 6 のように表示教材映像上で強調表示することにより提示する。これにより、視聴者は教材指示情報を受け取ることができ、講師が表示教材のどこを指して説明し

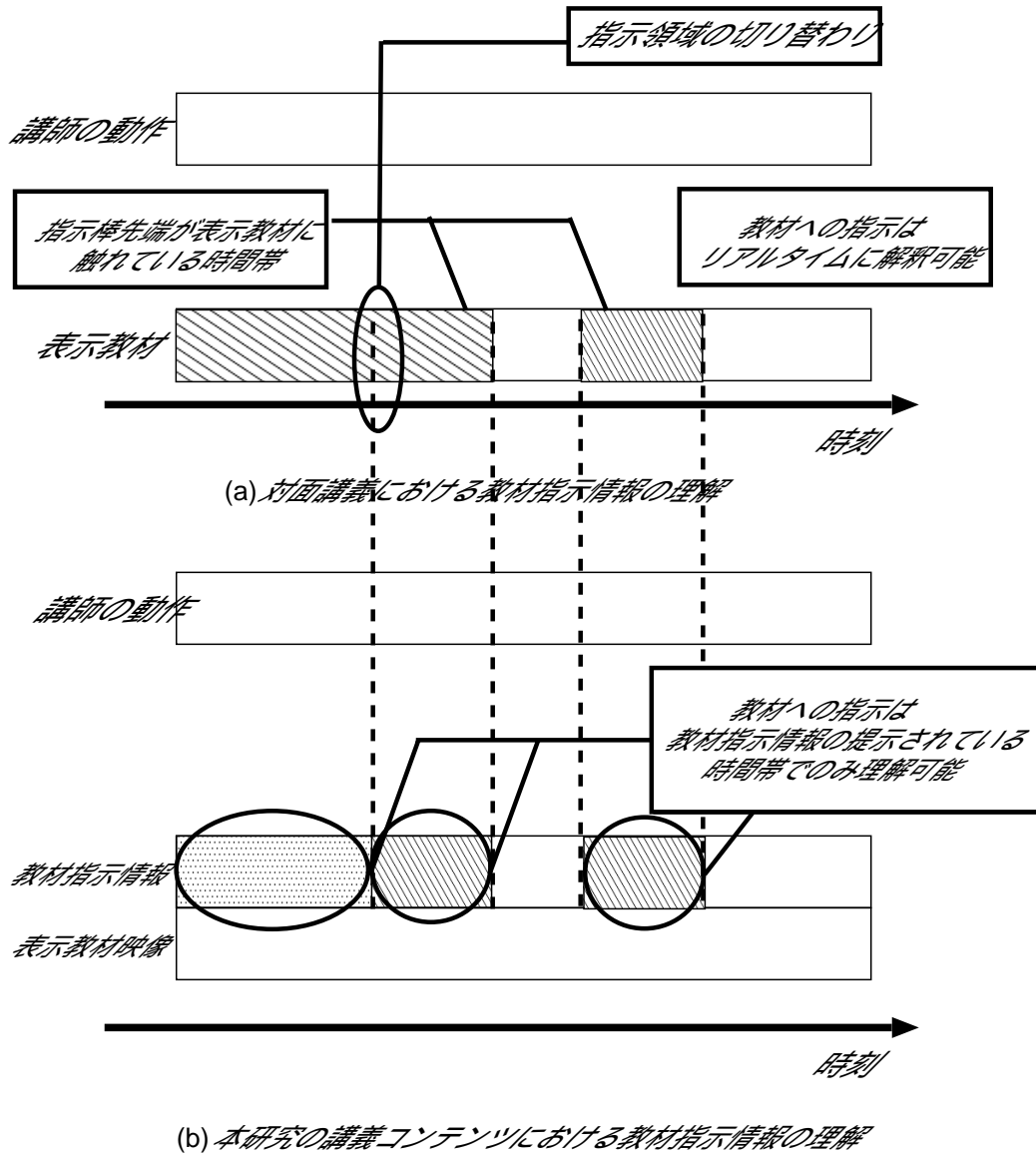


図 5: 対面講義と本講義コンテンツにおける教材への指示の理解

ているか把握できるので、教示内容を深く理解することができる。

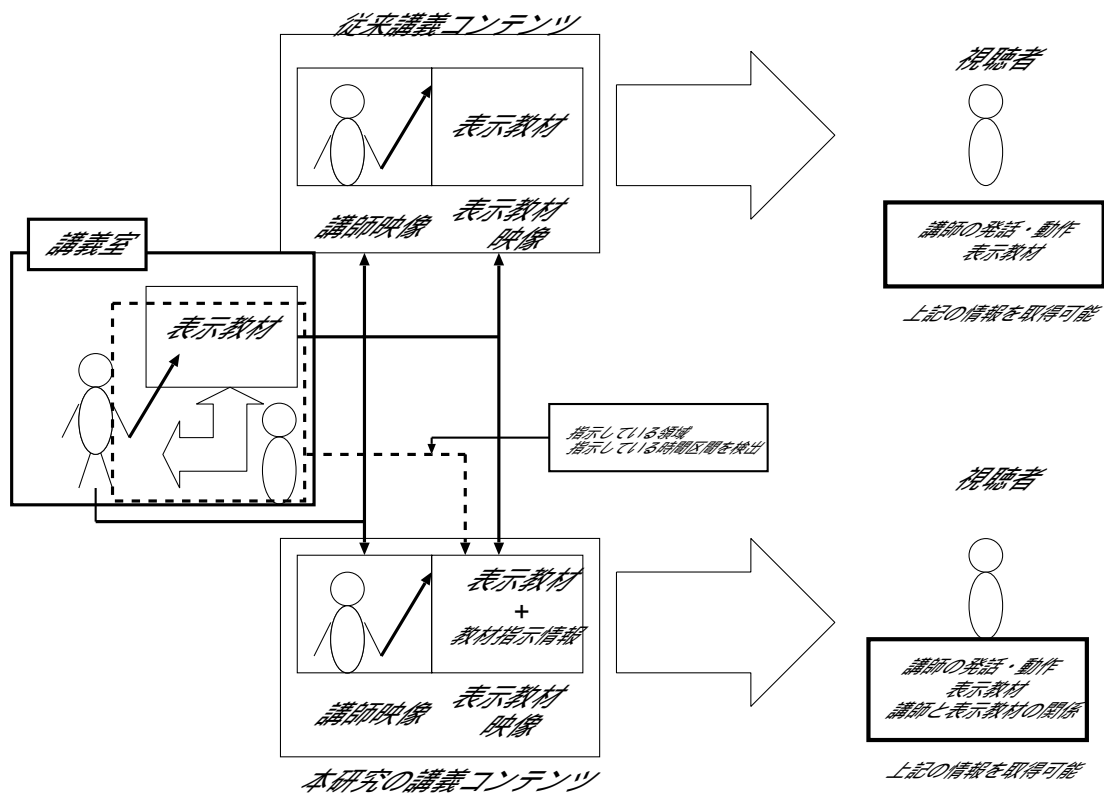


図 6: 従来と本研究での講義コンテンツによる視聴者の情報把握

第3章 教材指示情報の付与と提示

この章では、教材指示情報を講義アーカイブに付与し、視聴者に提示する手法について述べる。

3.1 教材指示情報の付与

教材指示情報は、講義の「どの時間区間」に「どの指示領域」を指示しているかという情報であり、スライドの領域分割、講師の指示意図領域の推定、及び指示領域の変更時刻の検出を行うことで取得できる。

本節では、教材指示情報の付与のための処理の流れについて説明する。講師は講義中に表示教材上の文や図を指示棒を用いて指示するという動作を繰り返す。そこで本研究では、講義が終了した後に処理を行う際に、講義のどの時間区間にどのスライドが提示されていたかという情報と、講義中に得られるセンサ情報を利用する。また、講義を観察して得られた知見より、講師の指示は指示棒の先端が教材にほぼ触れている状態で行われているものとする。

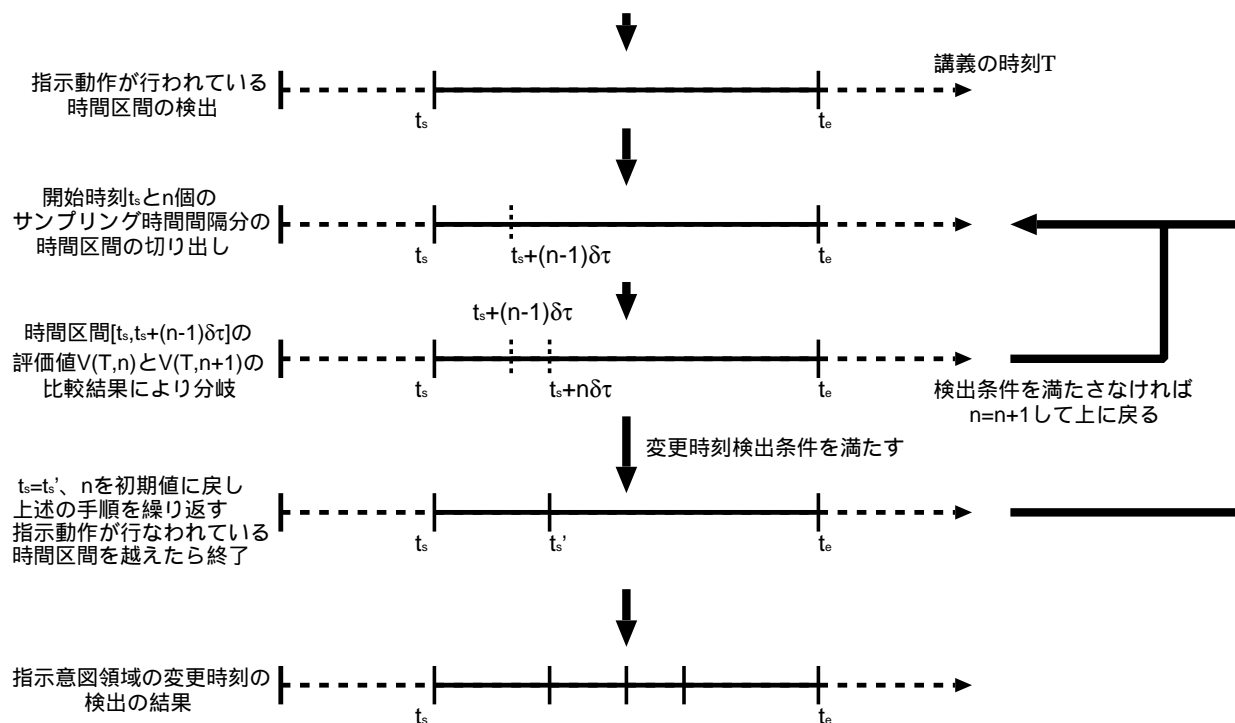


図7: 指示領域の推定とその変更時刻の検出

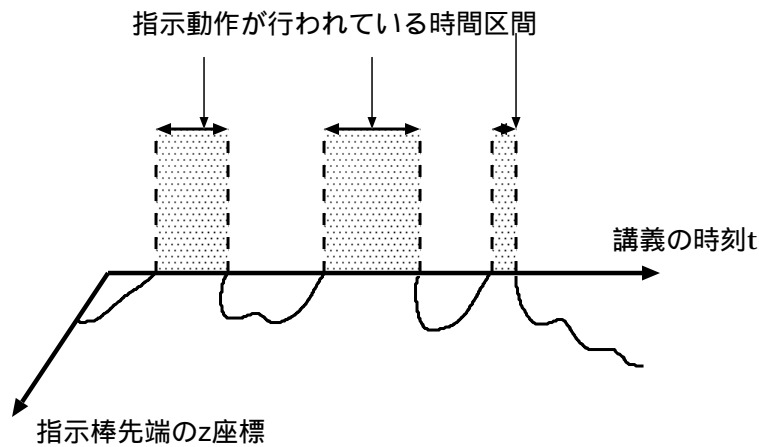


図 8: 指示動作が行われている時間区間の検出

以上のような仮定の下，本研究では図 7 に示すような流れで指示意図領域の推定と指示領域の変更時刻の推定を行う．まず，講義中に指示棒の先端が表示教材に触れている時間区間を図 8 のように指示棒先端の座標を利用して検出する．次に各時間区間で開始時刻 t_s ，終了時刻 $t_s + (n - 1)\delta\tau$ とする部分時間区間で各指示動作の確からしさを表す評価関数 $V_E(t_s, n)$ ， $V_H(t_s, n)$ ， $V_P(t_s, n)$ ， $V_O(t_s, n)$ を計算する．ただし， $\delta\tau$ はサンプリング時間間隔であり， n は 2 を最小値とし，一回評価するごとに 1 ずつ増加していく整数である．次に $V_E(t_s, n + 1)$ ， $V_H(t_s, n + 1)$ ， $V_P(t_s, n + 1)$ ， $V_O(t_s, n + 1)$ を計算し，結果を比較する．その結果，指示領域の変更が行われていないと判断されれば，次は n を 1 だけ増加してまた同様の操作を繰り返す．もし，そこが指示領域の変更時刻であると判断されれば，その時刻 $t_s + (n - 1)\delta\tau$ を指示領域の変更時刻とみなし， $[t_s, t_s + (n - 1)\delta\tau]$ の時間区間で評価関数の値が最も高い指示動作が行われたと推定する．そして $t_s = t_s + n\delta\tau$ を新たな t_s とし，上記処理を繰り返す．上記の操作は指示棒の先端が表示教材に触れている時間区間の間繰り返す．これをすべての指示動作が行われている時間区間に対して行うことにより，講師の指示領域と変更時刻の推定を行う．

以降の節ではスライドの領域分割，講師の指示意図領域の推定，及び指示領域の変更時刻の検出の詳細について述べる．

3.2 スライドの領域分割

まず本研究における座標系について説明する．本研究で扱う座標系は図9にあげている講義室の座標系・表示教材の座標系の二つである．講義室の座標系は，講師の動作を捉える際に用いる座標系で，表示教材が $X_L Y_L$ 平面上に存在するように設定してある．表示教材の座標系は指示領域の特定のために用い，両座標系は変換行列 R により相互変換される．

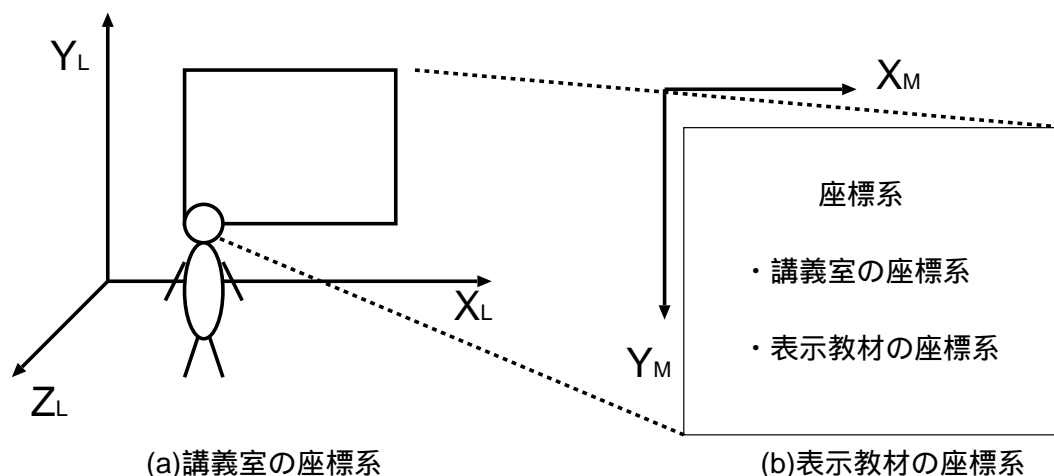


図9: 本研究で扱う座標系

本節では，スライドの領域分割について説明する．スライドは教示内容の主旨を表す簡潔な文や図から構成されており，講義における指示は，その簡潔な文や図に対して行われる．これらの簡潔な文や図という単位は先山らの研究[10]で用いられていた図字領域と同じものであることから，領域分割は先山らの研究で行われていた方法を利用する．上述の研究では，図字領域は文字や図を囲む最小の矩形領域により記述する．そして各矩形領域の代表する座標値を求めることで分割を行っている．本研究でもその手法と同様にして図10のように分割を行う．まず，ある時間区間に教室前面に提示されていた表示教材に関して，一つの文，一つの図の単位で，それぞれを含む最小の矩形領域を決定する．そして各矩形領域の左上と右下の座標値を代表座標値1，2と呼ぶこととする．なお，本研究では，各領域が図であるか，文であるかを利用するため，スライドごとに矩形領域のID，代表座標値1，代表座標値2，領域の種類という4つ組のデータの集合で記述する．以上の処理を講義中に提示されていたスライドす

べてに関して行うことで講義中の全スライドの領域分割を行う。

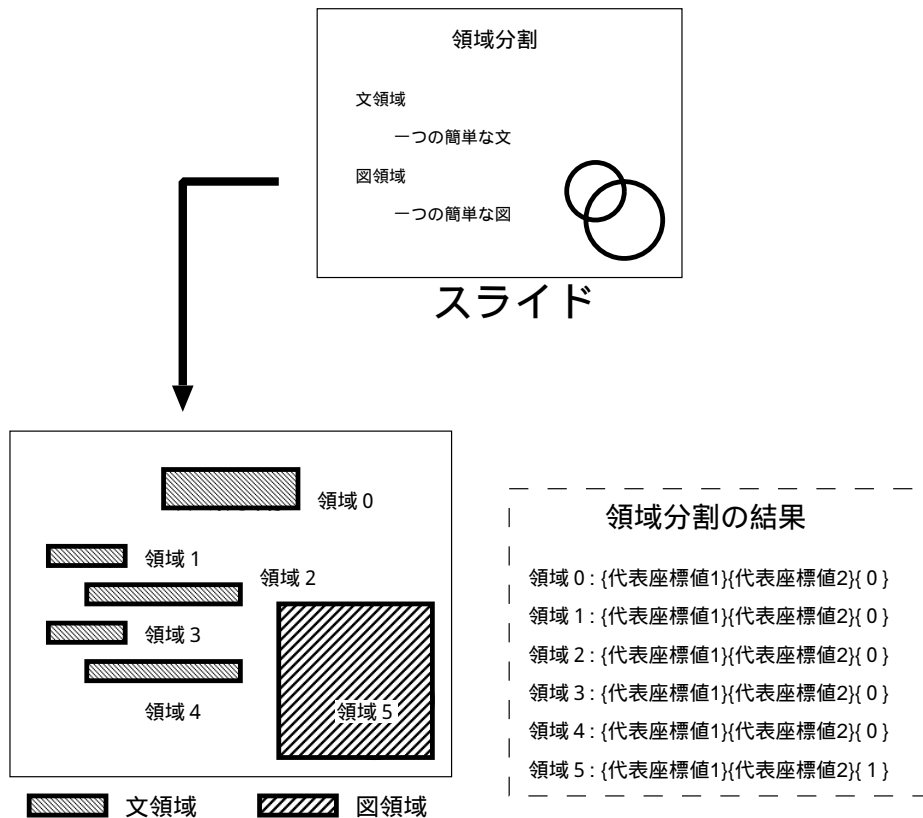


図 10: スライドの文領域・図領域への分割

3.3 指示意図領域と指示領域の変更時刻の推定

3.1 節で説明したように、本研究では講師の指示動作と指示意図領域の変更時刻を同時に把握するための評価関数を用いる。講師は大抵、指示をしようとするスライド上の領域に指示点を載せる。また、2.2.2 節で述べたように、指示意図領域は指示動作の種類に応じて決まる。指示動作の特定のためには、指示点や指示している領域の種類、講師の位置、講師の体の向きを検出する必要がある。以降では、指示意図領域の変更時刻を検出するための評価関数の定義、指示動作を規定する特徴及び評価関数の定義を行い、最後にそれらを総合した評価値計算による指示意図領域とその変更時刻の推定について述べる。

3.3.1 指示領域の変更時刻の検出のための評価関数

まず指示領域の変更時刻の検出のための評価関数について述べる．講師が指示棒を用いて指示を行うときは主に指示したい領域の上に指示点が多く、表示教材を投影しているスクリーンは大抵目線より高く、指示棒の先端が意図しない領域にくることもある．そのため、指示棒の先端が講師の意図する領域に来る確率が高いという仮定をおく．この仮定のもと、その時間区間内 $[t_s, t_s + (n - 1)\delta\tau]$ に一つの領域に指示点の分布が偏ったかを評価する評価関数 $C_O(t_s, n)$ を定義する． $C_O(T, n)$ は、時間区間 $[t_s, t_s + (n - 1)\delta\tau]$ における指示点が存在する領域の ID が i である回数を返す関数 $c_{(t_s, n)}(i)$ を用い以下のようにして定義する．

$$C_O(t_s, n) = \frac{1}{n} \max_i c_{(t_s, n)}(i)$$

この関数により、時刻 $t_s + (n - 1)\delta\tau$ と $t_s + n\delta\tau$ の間が指示領域の変更時刻であるとき $C_O(t_s, n)$ と $C_O(t_s, n + 1)$ の間で大きな値の変化を確認することができる．

3.3.2 指示動作の分類

まず、指示動作の分類に必要な特徴と各分類がその特徴によりどのように規定されるかについて述べる．Lilian らの研究 [8] によると指示動作を規定する特徴は指示棒の先端により示される表示教材上の点 (指示点) が示す軌跡、指示点が存在する表示教材上の領域の種類、講師の体の向き、講師の位置の 4 つである．

指示点の軌跡

指示点の軌跡は、指示領域の変更時刻を示す指標であるとともに講師の意図している指示領域を表す特徴であり、図 11 のように点軌跡・直線軌跡・楕円軌跡の 3 つに分けられる．

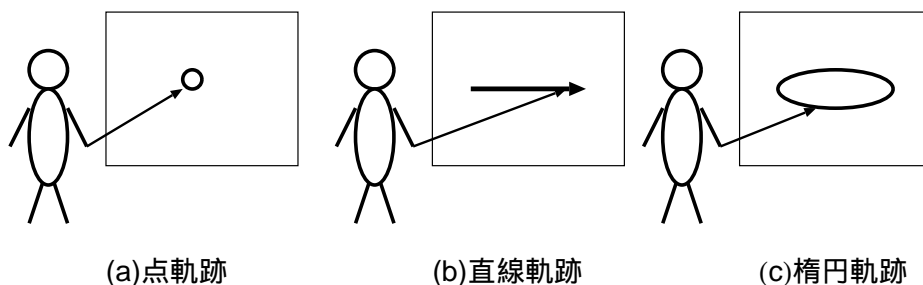


図 11: 指示点の軌跡

指示点が存在する表示教材上の領域の種類

指示点が存在する表示教材上の領域の種類とは、指示点が文領域、図領域、何もない領域、のどの領域上にあるかということであり、これは何を指しているかを表す特徴である。

講師の体の向き

講師の体の向きは、表示教材への講師自身の注目の度合いを表し [9]、図 12(a) のように講師が体を表示教材側に向けているときはよりスライド中の具体的な対象についての話題に触れ、逆に図 12(b) のように体を受講者側へに向けているときは講師は表示教材中の抽象的な対象に関する話題についてに触れるというように、指示領域の粒度を示す特徴である。

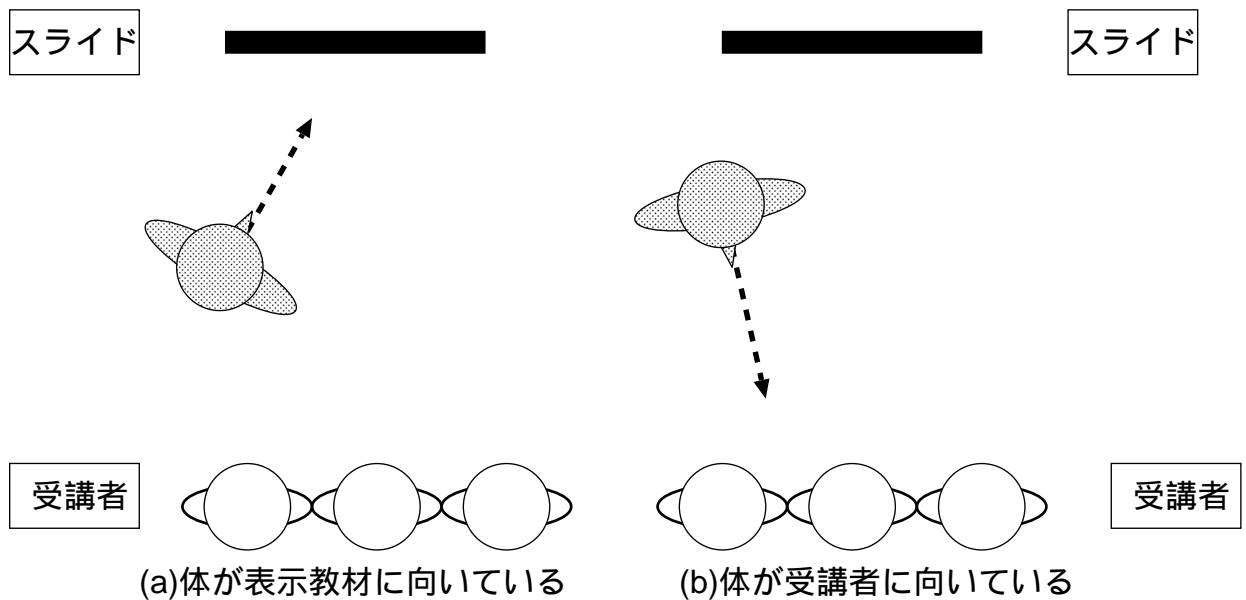


図 12: 講師の体の向き

講師の位置

講師の位置とは、講師と表示教材と受講者との位置関係のことであり、この特徴は、講師による学生の表示教材への注目の調整の指標である。図 13(a) のように表示教材がよく見えるように移動する場合は表示教材への注目の誘導であり、逆に図 13(b) のように表示教材が講師の体に隠れてもあまり講師自身が動かない場合には講師自身が表示教材中の具体的対象への注目を誘導しておらず、受講者にとっても抽象的な対象としか意味を持たないと考えられる。

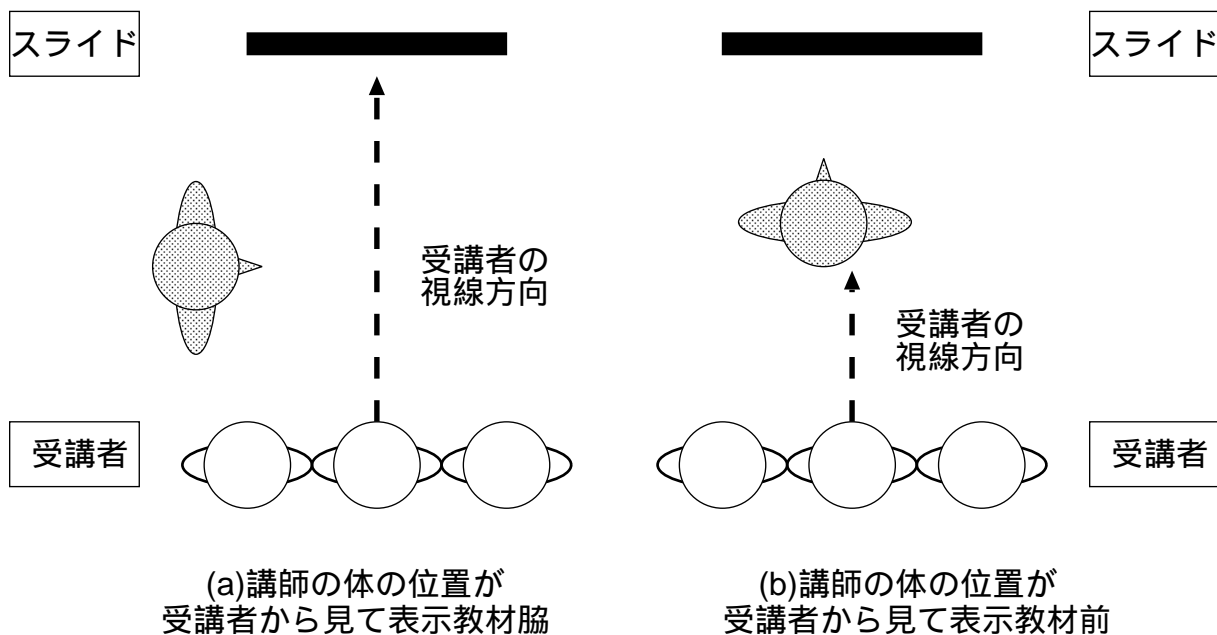


図 13: 講師の体の位置と受講者の視線方向

指示動作は，Lilian らの研究 [8] により表 2 のように，上述の各特徴の組み合わせにより規定できることがわかっている．そこで，各時間区間における各特徴に関して評価し，それらの評価を統合して評価した結果がもっとも良いものをその時間区間の指示動作であるとし，以降では各特徴の評価関数について説明する．

表 2: 指示動作を規定する特徴の組み合わせ

指示動作	指示点軌跡	指示領域の種類	講師の体の向き	講師の位置
Emphasizing	一点・直線・楕円	文・図領域	不定	不定
Highlighting	楕円	図領域	表示教材側	表示教材脇
	直線・楕円	文領域	表示教材側	表示教材脇
Pointing	一点	文・図領域	不定	表示教材脇
Outlining	不定 (対象依存)	図領域	表示教材側	表示教材脇

3.3.3 指示点の軌跡の評価

各時刻の指示点の位置の取得

実際の講義で行われた指示の動作を観察した結果，講義における指示は指示棒の先端を対象に接触させるようにして行うことがほとんどであった．そこで，各時刻の指示点の位置の取得は以下のように行う．その様子を図 14 に示す．なお，指示点とは各時刻で指示棒の付け根から先端方向へ向かうベクトルの延長線とスライドとが交差する点であるとする．接触しているかの判断は指示点とマーカ間の距離と指示棒の長さを比較することで行う．

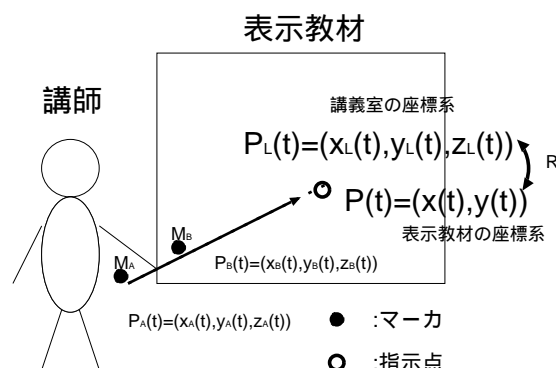


図 14: 指示点の検出

1. 指示棒に2つのマーカを M_A, M_B を設置し，それらの時刻 t における講義室の座標系での座標値 $P_A(t) = (x_A(t), y_A(t), z_A(t)), P_B = (x_B(t), y_B(t), z_B(t))$ を測定する．
2. マーカの座標値から指示棒の付け根から先端方向へ向かうベクトル $\overrightarrow{M_A M_B}$ を求める．
3. 手順2で求めたベクトル $\overrightarrow{M_A M_B}$ がスライドと交差する点 $P_L(t)$ を求める．ただし，指示は指示棒の先端をほぼスライドにくっつけて行うので，移動等である程度 (α) 離れることを考慮して，交点 P_L とマーカ M_A との距離 $|\overrightarrow{M_A P_L}|$ が指示棒の長さ l よりも α 以上長いような場合 ($|\overrightarrow{M_A P_L}| > l + \alpha$)，指示をおこなっていないものとして次のステップには進まないものとする．
4. $P_L(t)$ を変換行列 R により表示教材で定義する座標系での2次元座標値 $P(t)$ に変換する．

以上により各時刻での指示点の位置が検出される．

指示点の軌跡の種類を判別するための評価関数

上述の指示点位置の取得により，講義中の任意の時刻 t における 2 次元指示点位置 $\mathbf{P}(t)$ が得られる．以下では指示点のサンプリング間隔を $\delta\tau$ とし，講義中の任意時間区間 $[t_s, t_s + (n-1)\delta\tau]$ における軌跡を点・直線・楕円の観点から評価する評価関数について説明する．なお，これらの評価については，時間区間内で得られる全指示点がそれぞれの軌跡をとると仮定した場合，点軌跡なら代表となる点が直線軌跡なら代表となる直線が，楕円軌跡なら代表となる楕円が存在するとし，全指示点がそれら代表となる軌跡とどれだけ近いかで評価する．

- 点軌跡

まず，軌跡が点であることの評価関数 $T_P(t_s, n)$ について述べる．本研究では，その時間区間を代表する点を時間区間内の指示点群の重心とし，各点との距離から評価値を計算し，その評価値の平均をこの時間区間内の評価関数 $T_P(t_s, n)$ とする．以下に計算手順を示す．

1. 時間区間 $[t_s, t_s + (n-1)\delta\tau]$ における指示点群の重心 \mathbf{P}_g を求める．

$$\mathbf{P}_g = \frac{1}{n} \sum_{k=0}^{n-1} \mathbf{P}(t_s + k\delta\tau)$$

2. 時間区間中の時刻 t における指示点 $\mathbf{P}(t)$ と重心 \mathbf{P}_g との距離の評価関数 $E(t)$ を代表点から離れれば離れるほど小さくなり，かつ $0 \leq E(t) \leq 1$ となるように正規化するため以下のように定義し，その平均から評価関数 $T_P(T, n)$ を定義する．

$$E(t) = \frac{1}{1 + |\mathbf{P}(t) - \mathbf{P}_g|}$$

$$T_P(t_s, n) = \frac{1}{n} \sum_{k=0}^{n-1} E(t_s + k\delta\tau)$$

- 直線軌跡

次に指示点の軌跡が直線であることの評価関数 $T_L(t_s, n)$ のについて説明する．その時間区間を代表する直線は時間区間内の各指示点からの誤差が最小となる直線であるとする．そこで，本研究では任意時間区間内の指示点群から最小二乗誤差直線を求め，直線と各指示点との誤差の評価値の平均を評価値 $T_L(t_s, n)$ と定義する．

1. まず，時間区間 $[t_s, t_s + (n-1)\delta\tau]$ 内の指示点の分布から最小二乗誤差直線のパラメータを計算する．ここで時刻 t における指示点 $\mathbf{P}(t)$ の座標を $(x(t), y(t))$ とすると，求める最小二乗誤差直線の式 $y = a + bx$ のパラメー

タは以下のように求められる .

$$\bar{x} = \frac{1}{n} \sum_{k=0}^{n-1} x(t_s + k\delta\tau)$$

$$\bar{y} = \frac{1}{n} \sum_{k=0}^{n-1} y(t_s + k\delta\tau)$$

$$a = \frac{\bar{y} \cdot \sum_{k=0}^{n-1} x^2(t_s + k\delta\tau) - \bar{x} \cdot \sum_{k=0}^{n-1} (x(t_s + k\delta\tau) \cdot y(t_s + k\delta\tau))}{\sum_{k=0}^{n-1} x^2(t_s + k\delta\tau) - n\bar{x}^2}$$

$$b = \frac{\sum_{k=0}^{n-1} (x(t_s + k\delta\tau) \cdot y(t_s + k\delta\tau)) - n\bar{x} \cdot \bar{y}}{\sum_{k=0}^{n-1} x^2(t_s + k\delta\tau) - n\bar{x}^2}$$

2. 1 で求めた直線のパラメータから時刻 t における指示点 $P(t)$ と直線との距離 $D(t)$ を計算する .

$$D(t) = \frac{|bx(t) - y(t) + a|}{\sqrt{b^2 + 1}}$$

3. 時間区間中の時刻 t における指示点 $P(t)$ と直線との距離の評価値 $E(t)$ を代表となる直線から離れば離れるほど小さくなり , かつ $0 \leq E(t) \leq 1$ となるように正規化するため以下のように定義し , その平均から評価関数 $T_L(t_s, n)$ を規定する .

$$E(t) = \frac{1}{1 + D(t)}$$

$$T_L(t_s, n) = \frac{1}{n} \sum_{k=0}^{n-1} E(T + k\delta\tau)$$

- 楕円軌跡

最後に指示点の軌跡が楕円軌道であることの評価関数 $T_C(t_s, n)$ のについて説明する . 任意時間区間内の指示点の軌跡を表示教材の座標系に水平な楕円とすると , 時間区間内の指示点群は楕円に外接する矩形領域内に存在すると考えられる . そこで , この時間区間を代表とする楕円軌跡は時間区間内の指示点群を囲む矩形領域の中心を楕円の中心とし , 矩形領域の長辺を楕円の長径 , 矩形領域の短辺を短径とした軌跡を描くと仮定する . そして評価関数 $T_C(t_s, n)$ は楕円と各指示点との誤差の評価値の平均により定義する .

1. まず , 時間区間 $[t_s, t_s + (n-1)\delta\tau]$ 内の指示点の分布から楕円のパラメータを計算する . ここで時刻 t における指示点 $P(t)$ の座標を $(x(t), y(t))$ とすると , 求める楕円 $\frac{(x-x_c)^2}{a^2} + \frac{(y-y_c)^2}{b^2} = 1$ のパラメータは以下のようにして求め

られる。

$$x_c = \frac{\max_t x(t) + \min_t x(t)}{2}$$

$$y_c = \frac{\max_t y(t) + \min_t y(t)}{2}$$

$$a = \frac{\max_t x(t) - \min_t x(t)}{2}$$

$$b = \frac{\max_t y(t) - \min_t y(t)}{2}$$

2. 1で求めた楕円のパラメータから時刻 t における指示点 $P(t)$ と楕円との距離 $D(t)$ を計算する。なお、ここでは楕円と指示点との距離は図 15 に示すように、楕円の中心と指示点を結んだ直線を考えてとき、この直線の楕円外の部分の長さであると定義し、以下のように計算できる。

$$D(t) = \left| \sqrt{(x_t - x_c)^2 + (y_t - y_c)^2} * \left(1 - \frac{1}{\sqrt{(x_t - x_c)^2/a^2 + (y_t - y_c)^2/b^2}} \right) \right|$$

3. 時間区間中の時刻 t における指示点 $P(t)$ と楕円との距離の評価値 $E(t)$ を代表となる楕円軌跡から離れれば離れるほど小さくなり、かつ $0 \leq E(t) \leq 1$ となるように正規化するため以下のように定義し、その平均から評価関数 $T_C(t_s, n)$ を定義する。

$$E(t) = \frac{1}{1 + D(t)}$$

$$T_C(t_s, n) = \frac{1}{n} \sum_{k=0}^{n-1} E(t_s + k\delta\tau)$$

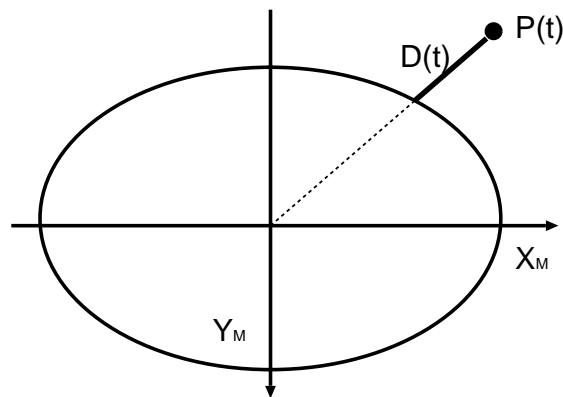


図 15: 楕円軌跡と指示点との距離の計算

3.3.4 指示点の存在する領域の種類の評価関数

次は指示点の存在する領域の種類の評価関数について説明する．これはある時間区間における指示点群の存在する領域が文か図かその他に偏っているかの評価であり，文に偏っているかの評価関数 $R_S(t_s, n)$ ，図に偏っているかの評価関数 $R_P(t_s, n)$ その他に偏っているかの評価関数 $R_E(t_s, n)$ の算出方法について述べる．これらは時刻 t においてそれぞれの領域に含まれていたら 1 を，そうでなければ 0 を返す関数 $r_S(t)$ ， $r_P(t)$ ， $r_E(t)$ を用いて以下のように定義できる．

$$R_S(t_s, n) = \frac{1}{n} \sum_{k=0}^{n-1} r_S(t_s + k\delta\tau)$$

$$R_P(t_s, n) = \frac{1}{n} \sum_{k=0}^{n-1} r_P(t_s + k\delta\tau)$$

$$R_E(t_s, n) = \frac{1}{n} \sum_{k=0}^{n-1} r_E(t_s + k\delta\tau)$$

$$(R_S(t_s, n) + R_P(t_s, n) + R_E(t_s, n) = 1)$$

3.3.5 講師の体の向きに関する評価関数

次に体の向きに関する評価関数について説明する．この関数では，講師が指示していない状態と比べ，受講者の注目を表示教材へと誘導させようとしているかということの評価をしたいので，講師が指示していない状態から，どれだけ表示教材側に身体を向けているかを評価する．そこで講師がどの方向を向いているかの検出の手順について図 16 に従って説明する．

1. 講師の両肩にマーカを設置し，講義室の座標系での両肩の 3 次元座標値を測定する．
2. $y = 0$ で定義できる平面に両肩のマーカ位置を投影する (P_L, P_R) ．
3. 両肩の位置を投影した点を結ぶ線分の中心を頭の位置 (P_C) とし，中心から体の前方方向に垂直に直線をのばす．
4. のばした直線と表示教材との交点 P_F から図 16 のように時刻 t で講師の向きを $o_B(t)$ を $-1 \sim 1$ の値で算出する．具体的には以下の式に従う．

$$o_B(t) = \begin{cases} 1 & , (P_F \text{ が } M_L \text{ と } M_R \text{ の間にある場合}) \\ \frac{1}{(|\vec{M}_C P_F| - |\vec{M}_C M_L|) / |\vec{M}_L M_R|} & , (P_F \text{ が } M_L \text{ と } M_R \text{ の間でない場合}) \end{cases}$$

$$(o_B(t) = -o_B(t) \quad (\text{身体の正面方向が表示教材側ではない場合}))$$

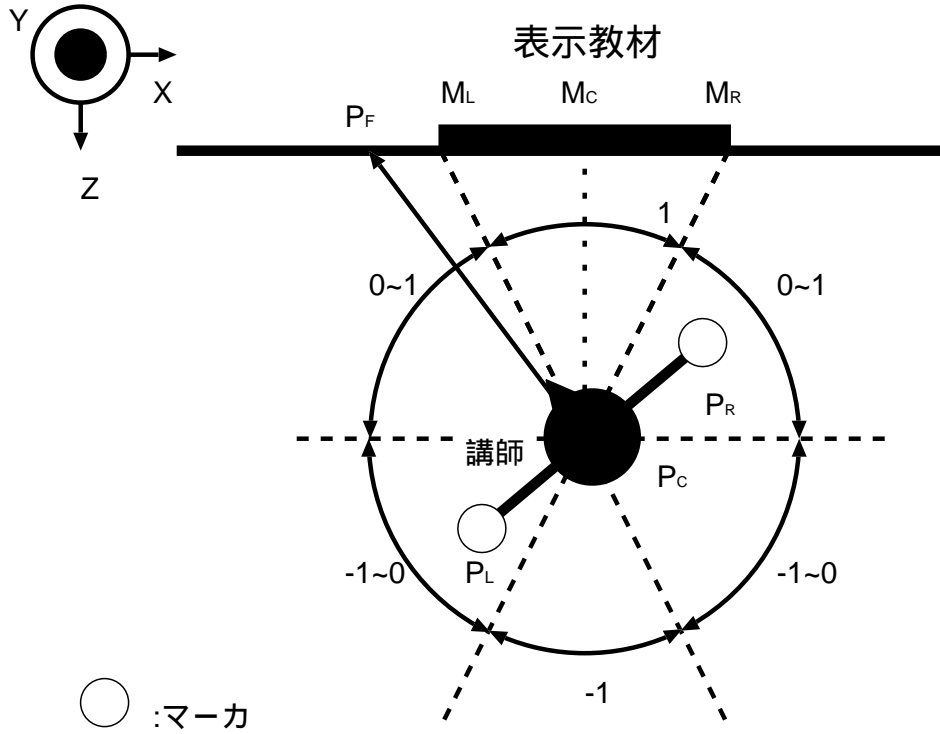


図 16: 講師が体を向けている方向の特定 (講義室の俯瞰図)

これにより講師の向いている方向を定義する．そして，以下の手順で講師が何も指示していない状態で向いている方向よりも表示教材側にどれだけ体を向いているかの評価関数 $O_B(t_s, n)$ を定義する．

1. 講義全体 (時間区間 $[0, (n_{total} - 1)\delta\tau]$) を通じて講師が指示していないときの講師の向き \bar{o}_B を求め，それを講師が何も指示していないときに向いている方向とする．なお，講師が時刻 t に指示しているときに 1 を返す関数 $p(t)$ と，講義中に指示していないと検出できる回数 n_{normal} とする．

$$\bar{o}_B = \frac{1}{n_{normal}} \sum_{k=0}^{n_{total}-1} o_B(k\delta\tau) \cdot p(k\delta\tau)$$

2. 時間区間中の時刻 t における講師の体の向き $o_B(t)$ と通常 \bar{o}_B から講師がどれだけ資料側に体を向いているかの評価値 $D_{o_B}(t)$ を $0 \leq D_{o_B}(t) \leq 1$ になるよう正規化するため下記のように定義する．

$$D_{o_B}(t) = \begin{cases} \frac{o_B(t) - \bar{o}_B}{2(1 - \bar{o}_B)} + 0.5, & \text{if } o_B(t) > \bar{o}_B \\ \frac{o_B(t) + 1}{2(\bar{o}_B + 1)}, & \text{if } o_B(t) \leq \bar{o}_B \end{cases}$$

3. 時間区間 $[t_s, t_s + (n - 1)\delta\tau]$ 中の $D_{o_B}(t)$ の平均から評価値 $O_B(t_s, n)$ を計算

する．

$$O_B(t_s, n) = \frac{1}{n} \sum_{k=0}^{n-1} D_{o_B}(t_s + k\delta\tau)$$

3.3.6 講師の体の位置に関する評価関数

最後に講師の体の位置に関する評価関数について説明する．この関数では指示のためにどれだけ講師が表示教材の脇の方へ体を移動させるかで評価する．そこで，まず講師の位置について取得する必要があるが，それは図 16 のように両肩にマーカをつけ，その中心 P_C を講師の体の位置とすることで解決し，図 16 の X 軸方向でどれだけ脇によっているかで判断する．そのためある時刻における講師の位置の X 座標を $x(t)$ とし，通常時よりも表示教材脇にどれだけ移動しているかの評価関数 $P_B(t_s, n)$ を以下の手順で定義する．

1. 一つの教材資料を説明している時間区間 $[0, (n_{total} - 1)\delta\tau]$ を通じて講師の位置の平均 \bar{p}_B を求め，それを講師が通常いる位置とする．またその間に講師の移動範囲も考え分散 v_{p_B} も計算する．

$$\bar{p}_B = \frac{1}{n_{total}} \sum_{k=0}^{n_{total}-1} x(k\delta\tau)$$

$$v_{p_B} = \frac{1}{n_{total}} \sum_{k=0}^{n_{total}-1} (x(k\delta\tau) - \bar{p}_B)^2$$

2. 時間区間中の時刻 t において講師がどれだけ資料脇に移動しているかの評価値 $D_{p_B}(t)$ を $0 \leq D_{p_B}(t) \leq 1$ という形で正規化するために，下記のように定義する．

$$D_{p_B}(t) = \begin{cases} 1 - \frac{1}{1 + \frac{|p_B(t) - \bar{p}_B|}{\sqrt{v_{p_B}}}} & , (p_B(t) \text{ が通常より資料脇にある場合}) \\ 0 & , (p_B(t) \text{ が資料前側にある場合}) \end{cases}$$

3. 時間区間 $[t_s, t_s + (n - 1)\delta\tau]$ 中の $D_{p_B}(t)$ の平均から評価値 $P_B(t_s, n)$ を計算する．

$$P_B(t_s, n) = \frac{1}{n} \sum_{k=0}^{n-1} D_{p_B}(t_s + k\delta\tau)$$

3.3.7 指示意図領域と指示領域の変更時刻の決定

指示動作の決定と指示領域の変更時刻の検出

これまで，指示動作を規定している特徴からの，講義の時間区間 $[t_s, t_s + (n - 1)\delta\tau]$ における指示領域の変更時刻の検出のための評価関数と指示動作分類のための評価関数について説明してきた．ここでは，それらの評価関数と表 2 からの，時間区間 $[t_s, t_s + (n - 1)\delta\tau]$ における各指示動作の評価関数の定義について述べる．

表 2 より，指示動作を規定する特徴の組み合わせがわかるので，規定する特徴すべての評価関数が高いものが指示動作に選ばれるように各指示動作の評価関数を設定する．そこで，各指示動作を規定する特徴の評価関数の積により，どの指示動作かを評価すると共に，指示領域の変更時刻検出のための評価関数も利用して，時間区間 $[t_s, t_s + (n - 1)\delta\tau]$ 中の Emphasizing, Highlighting, Pointing, Outlining の評価値 $(V_E(t_s, n), V_H(t_s, n), V_P(t_s, n), V_O(t_s, n))$ を以下のように定義する．そのため，時間区間内の評価関数の値が高いものが，その時間区間内の指示動作として選ぶことができ，値が小さくなれば，指示領域の変更時刻として検出することができる．

$$V_E(t_s, n) = C_O(t_s, n) \times \max[T_P(t_s, n), T_L(t_s, n), T_C(t_s, n)] \times \max[R_S(t_s, n), R_P(t_s, n)] \times \max[((1 - O_B(t_s, n)) \times (1 - P_B(t_s, n))), (O_B(t_s, n) \times (1 - P_B(t_s, n)))]$$

$$V_H(t_s, n) = C_O(t_s, n) \times \max[(T_C(t_s, n) \times R_P(t_s, n)), (T_L(t_s, n) \times R_P(t_s, n)), (T_L(t_s, n) \times R_S(t_s, n))] \times O_B(t_s, n) \times P_B(t_s, n)$$

$$V_P(t_s, n) = C_O(t_s, n) \times T_P(t_s, n) \times \max[R_S(t_s, n), R_P(t_s, n)] \times \max[O_B(t_s, n), (1 - O_B(t_s, n))] \times P_B(t_s, n)$$

$$V_O(t_s, n) = C_O(t_s, n) \times \max[T_L(t_s, n), (1 - T_P(t_s, n))] \times R_P(t_s, n) \times O_B(t_s, n) \times P_B(t_s, n)$$

$$V_*(t_s, n) - V_*(t_s, n + 1) > D_{TH*}$$

$$V_*(t_s, n + 1) < V_{TH*}$$

以上により指示領域の変更時刻を検出すると共に、その区間内で評価値がもっとも高いものが時間区間 $[t_s, t_s + (n - 1)\delta\tau]$ の指示動作とする。

指示動作分類結果からの指示意図領域の決定

2.2.2 節で述べたように、各指示動作により講師が意図する指示領域は決定できる。そのため、以下に示すように各指示動作に応じて指示意図領域を決定・保存する。

- Emphasizing で意図される指示領域は一つの文領域・図領域であるので、その時間区間中にもっとも長く指示点が乗っていた領域の情報(領域の ID)を求める。すなわち、下式の i の値を返す。

$$C_O(t_s, n) = \frac{1}{n} \max_i c_{(t_s, n)}(i)$$

- Highlighting で意図される指示領域は文領域や図領域の一部であるので、その時間区間中に指示点が分布していた領域を囲む最小矩形領域の左上と右下の頂点座標を保存する。すなわち、時間区間 $[t_s, t_s + n\delta\tau]$ の間の指示点 $P(t) = (x(t), y(t))$ に対し、 $\max_t x(t), \min_t x(t), \max_t y(t), \min_t y(t)$ を求める。
- Pointing で意図される指示領域は語句・点であるので、その時間区間中の指示点の重心座標を求める。
- Outlining で意図される指示領域は形状であり、それは指示点の軌跡で記述されるのでその時間区間中の指示点の座標をすべて求める。

そして、時間区間、指示動作の種類、指示動作ごとに求めたデータの組を保存する指示意図領域の推定が完了する。

3.4 指示領域の決定

3.2 節と 3.3 節により得られた構成要素の領域分割と時間区間内の指示意図領域より、講師の指示領域は図 17 のように求められる。まず、指示意図領域推定の時に得られる時間区間動作の種類動作ごとの指示意図領域のデータの組から、その時間区間に提示されている表示教材の構成要素の領域分割の結果を持ってくる。

次に動作の種類に応じて、指示領域を決定する。Emphasizing の場合には領域の ID が与えられるので、構成要素の領域分割の結果から構成要素の領域の左上と右上の座標値が得られる。Highlighting の場合は、矩形領域が与えられるので、構成要素の領域分割の結果とその矩形領域との積領域である矩形領域の

代表座標値を得ることができる。Pointing の場合はそのままの座標値を指示領域として得ることができる。Outlining はその区間中の指示点をすべて結んだ図形と構成要素の領域との積領域が指示領域として得ることができる。以上のようにして指示領域が決定できる。

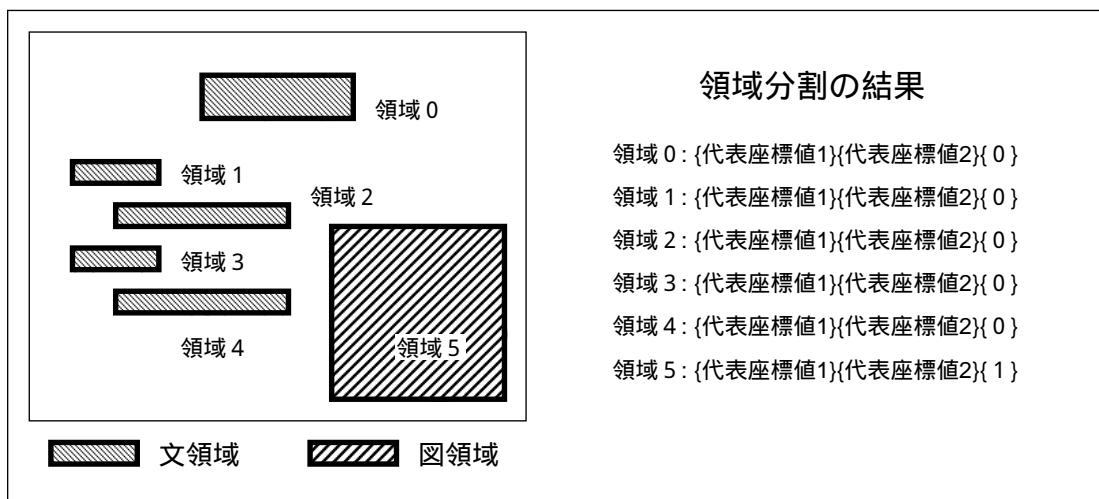
3.5 アーカイブに付与された教材指示情報の提示

この節では教材指示情報の提示について述べる。本研究における教材指示情報の提示は 2.2.3 節で記述したように、従来の表示教材映像に教材指示情報を強調表示させることにより行う。強調表示は教材が指示されている時間区間の間だけ、指示範囲を明確にする形で指示動作の種類に応じた表現を用いて行い、その流れを図 18 に示す。

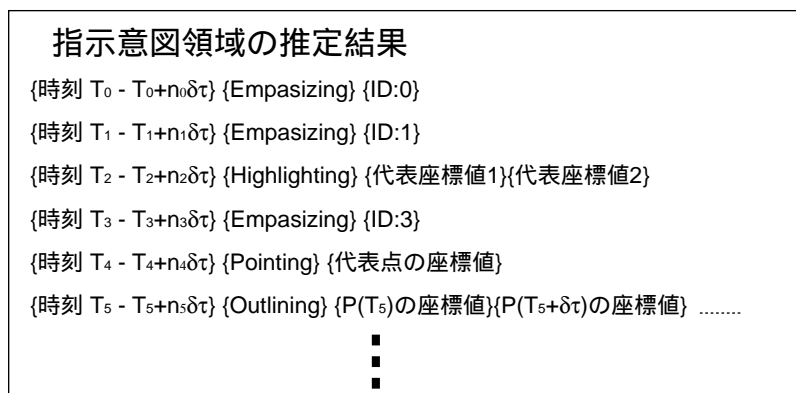
具体的な手順としては以下のような流れになる。

1. 図 17 のように指示領域は時間区間と指示動作の種類、指示領域の組で与えられる。まず、時間区間に表示されている表示教材に関してその画像を抽出する。
2. 次に抽出された画像を指示動作の種類と指示領域の組を基に以下のように描き換える。
 - Emphasizing の場合の場合、指示領域の座標情報を用い、領域中に含まれる表示教材中の文字・図の色を変化させる。
 - Highlighting の場合、指示領域の座標情報から、その矩形領域に内接する楕円を表示教材の画像に描く。
 - Pointing の場合、表示教材中で代表点の位置するところに点を描画する。
 - Outlining の場合、表示教材の画像に指示領域として与えられた図形を描く。
3. 描き換えられた資料画像をもとの資料映像の同一時間区間に相当する部分の画像と置き換える。

以上の流れにより、教材指示情報を表示教材映像上で強調表示させることで、本研究では講義コンテンツに付与された教材指示情報の提示を実現する。



+



↓

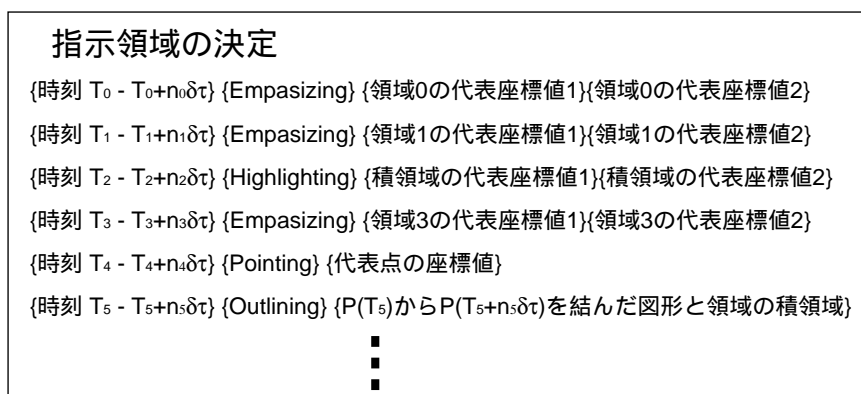


図 17: 構成要素の領域分割と指示意図領域による指示領域特定

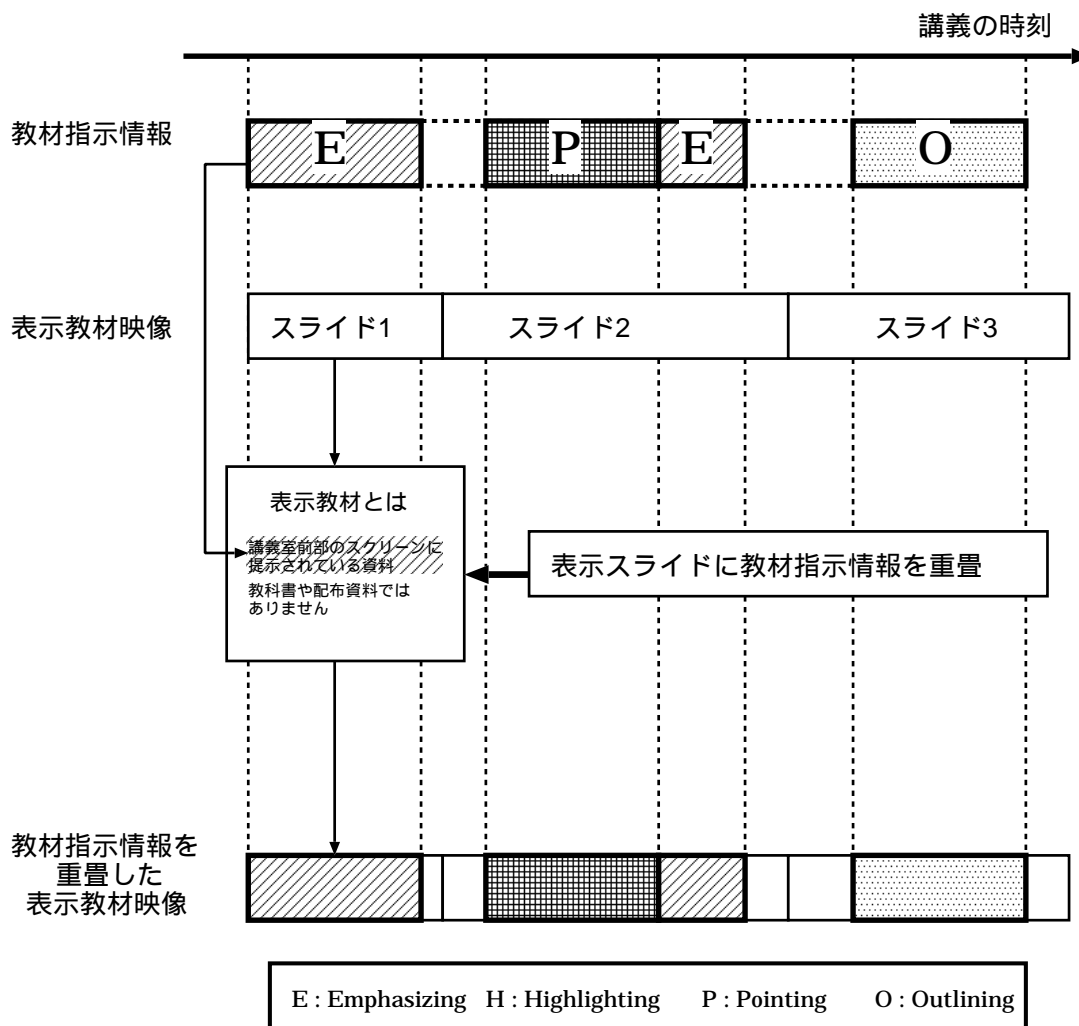


図 18: 教材指示情報を重畳した表示教材映像作成

第4章 実験と評価

本章では，本研究で提案した教材指示情報の付加した講義コンテンツの作成に関して，予備実験と，教材指示情報の抽出と提示の有効性を確認するために行った実験について紹介する．

4.1 実験の目的

本実験は講義コンテンツへの教材指示情報の付加に関する有効性を確認するために行ったものであり，その目的は教材指示情報の抽出が実際に人手で行った場合のものと同様であるか，また教材指示情報を提示したことが教示内容の理解に役立ったかについての検証を行うことである．

4.2 実験環境

この節では実験を行った際の実験環境について説明する．本研究で行う実験のために使用した講義室の大まかな俯瞰図を図 19 に示す．この講義室には講義室前面に3面の教材の表示用のスクリーンがあり，また講義中の様子を観測，撮影するために13台のカメラが設置されており，そのうちの4台で講師を追跡した映像を撮影している．また講義室の前天井部には超音波位置センサを設置し，マーカ位置に関してはこの超音波位置センサによって計測する座標情報を利用する．

なお以降行う実験では，設定を簡単にするため，スクリーンはスクリーン2のみ使用する．

4.3 予備実験

4.3.1 手動抽出した教材指示情報の評価実験

本節では，指示点の位置のみを離散点軌跡として重畳表示させた場合と，本研究で考えている指示領域を明確化して提示させた場合とを比較し，本研究で考えている教材指示情報を提示することの有効性について確かめるための予備実験を行った結果を述べる．

実験は以下の手順で行った．まず一人の講師に短時間の講義を行ってもらい，その様子を講義コンテンツとして保存すると共に，講師と表示教材が同時に捉えられるようにビデオカメラで撮影する．次に撮影した映像から本研究におけ

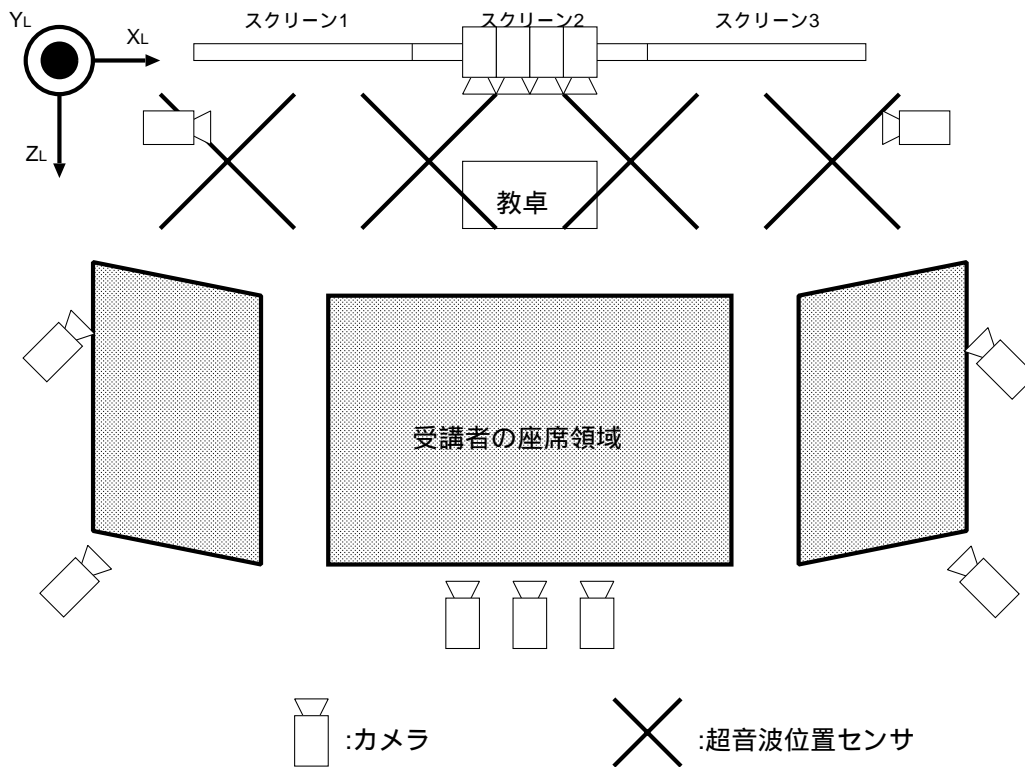


図 19: 実験に用いた講義室の俯瞰図

る教材指示情報を手動で抽出する．最後に講義コンテンツに対し，指示点だけを重畳させたものと，手動で抽出した教材指示情報を強調表示させたものを作成し，それらの両方を被験者に見てもらい，アンケートに答えてもらうことで本研究で考えている教材指示情報を提示することの有効性について確かめた．

アンケートは6人の被験者に対し行い，そのうち3人(Aグループ)は指示点だけを重畳表示させた映像を先に見てもらい，残りの3人(Bグループ)には手動抽出した教材指示情報を強調表示させた映像を先に見てもらい表3のアンケートに答えてもらった．なお，その時に見てもらった分類に応じた映像のカットを図20に示す．

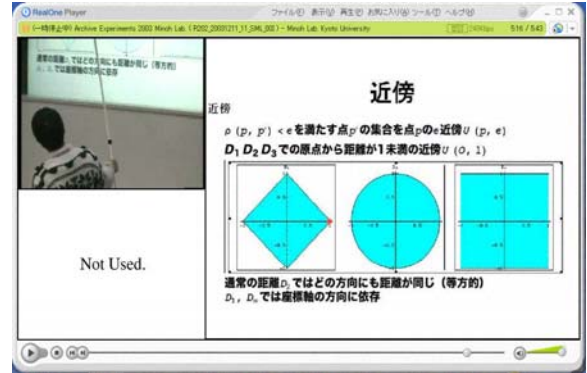
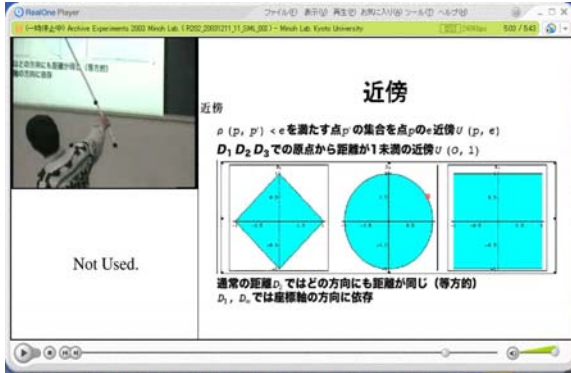
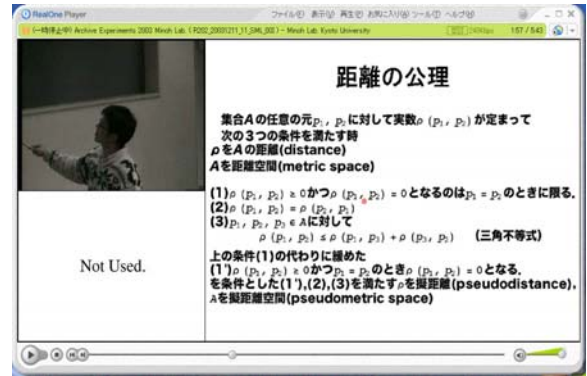
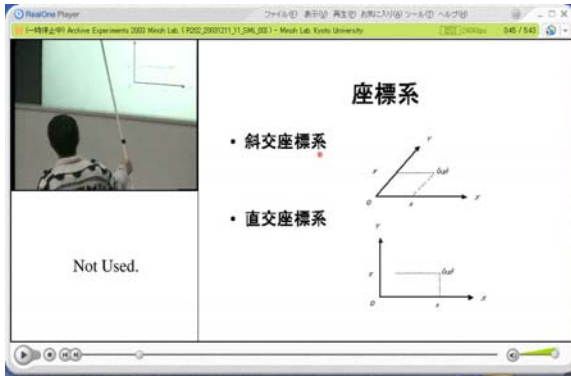
アンケートの結果は表4のようになった．表中の数字は評価項目の点数を平均したものである．この結果，すべての項目に対し，本研究の教材指示情報を強調表示させた結果のほうが良い評価が得られていることが確認できた．また，どちらのほうが良いかという質問に対しても6人中5人が本研究の教材指示情報を強調表示させたほうが良いと答えており，本研究で想定する教材指示情報の提示が教示内容の理解に役立っているといえる．

表 3: アンケート項目

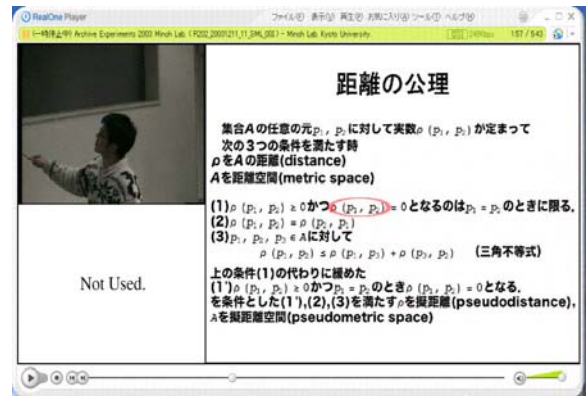
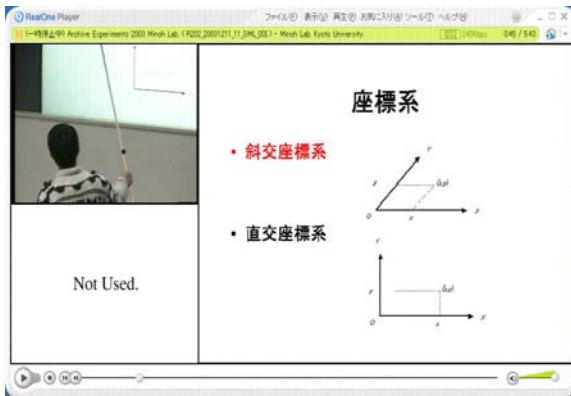
<一つ目の映像に関して>		良				悪
画面は見やすかったか？	:	5	4	3	2	1
教材指示情報が理解を助けたか？	:	5	4	3	2	1
指示情報の位置がわかりやすかったか？	:	5	4	3	2	1
集中して受講することができたか？	:	5	4	3	2	1
<二つ目の映像に関して>						
画面は見やすかったか？	:	5	4	3	2	1
教材指示情報が理解を助けたか？	:	5	4	3	2	1
指示情報の位置がわかりやすかったか？	:	5	4	3	2	1
集中して受講することができたか？	:	5	4	3	2	1
<両方の映像に関して>						
わかりやすかったのは？	:	一つ目		同じ		二つ目

表 4: アンケート結果

<指示点のみを重畳した映像に関して>	Aグループ	Bグループ	総合
画面は見やすかったか？	3	3.3	3.2
教材指示情報が理解を助けたか？	4.3	3	3.7
指示情報の位置がわかりやすかったか？	3	4.3	3.7
集中して受講することができたか？	3.3	3	3.2
<教材指示情報を強調表示した映像に関して>	Aグループ	Bグループ	総合
画面は見やすかったか？	4	4.3	4.2
教材指示情報が理解を助けたか？	4.3	4.3	4.3
指示情報の位置がわかりやすかったか？	3.3	4.7	4
集中して受講することができたか？	3.3	4.3	3.8

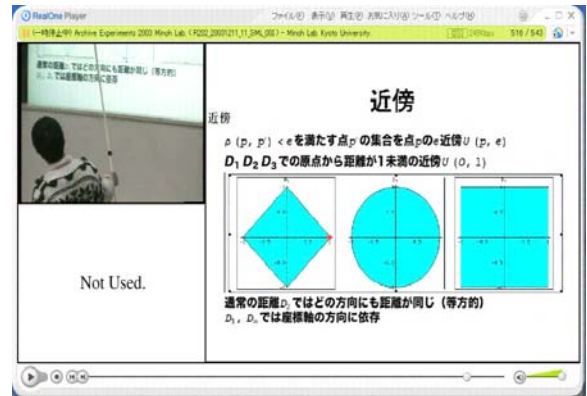
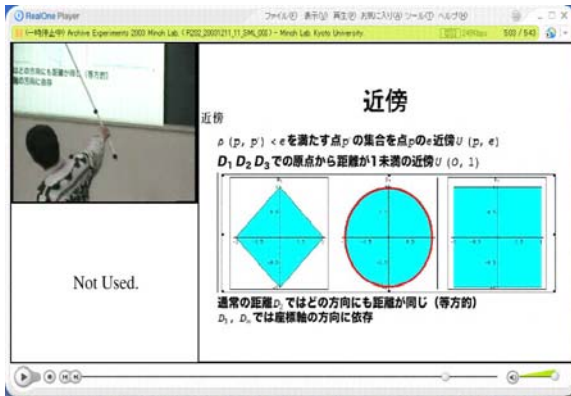


(a)指示点をそのまま表示した提示



(1)Emphasizingの場合の表示

(2)Highlightingの場合の表示



(3)Outliningの場合の表示

(4)Pointingの場合の表示

(b)手動分類した結果を反映させた提示

図 20: 評価のために用いた映像のカット

4.3.2 講義中の講師の指示に関する実験と検討

本節では実際の講義における講師の指示動作に関する情報を得るため、講義の様子を撮影し、確認した時に得られた知見に関して述べる。実験手順は以下の通りである。異なる学部の講義を担当する4人の講師について、1回90分の講義を2回ずつビデオに撮影し、講義中での講師の指示対象・指示時の視線・指示動作の継続時間についてビデオを再生しながらそれらの情報を手動で検出した。なお、撮影したビデオは講師・指示対象間の関係が明確に観測できるように講義室後方の固定カメラから撮影したものである。

その結果、以下のようなことがわかった。

- 指示には指示棒がよく使われる
- 講師が資料を指示する場合、指示棒の先端を指示対象に接触させる
- 講師の指示動作は点・線・楕円の3種類が観測された

以上の点より3.3.3節で述べた指示点検出において指示点を検出するタイミングは先端位置が対象に接触した時点で行うものとする。そのうえ、4.2節の実験環境であってもLilian[8]らの研究で示されている指示動作分類(emphasizing, highlighting, pointing, outlining)のそれぞれが観測されているため、この分類を適用することには問題がないと考えられる。

4.4 教材指示情報抽出に関する実験と評価

この節では本研究で提案した手法で抽出した教材指示情報が手動で抽出した場合と比較して、どの程度手動で抽出した結果と一致しているかについて評価を行う。

実験のため、3人の講師に10分程度の講義を2種類ずつ計6種類の講義をしてもらい、その様子を従来の講義コンテンツの形で保存すると共に、講師と教材資料の関係がわかるように全体を捉えるようにビデオで撮影を行う。手動による分類は、撮影されたビデオを確認して行った。手動で分類した結果と、本手法により分類した結果との比較は図21で示すようにして行う。図は手動で分類した結果と自動で分類した結果を時間軸をそろえて並べた様子であり、自動で分類した一つの指示動作による時間区間が手動で分類したときの時間区間に含まれ、かつ分類の種類が同じときに正しく分類できたと判断する。例えば分類結果が同じでも、手動で取り出した時間区間に正しく含まれなかった場合は分類がうまくいかなかったとする。すなわち図中の網掛け部分はうまく分類で

きなかった部分を示している．ここで分類成功率 S は，指示が検出できた時間 T_{point} と正しく分類された時間 $t_{correct}$ を用いて以下のように定義する．

$$S = \frac{T_{correct}}{T_{point}}$$

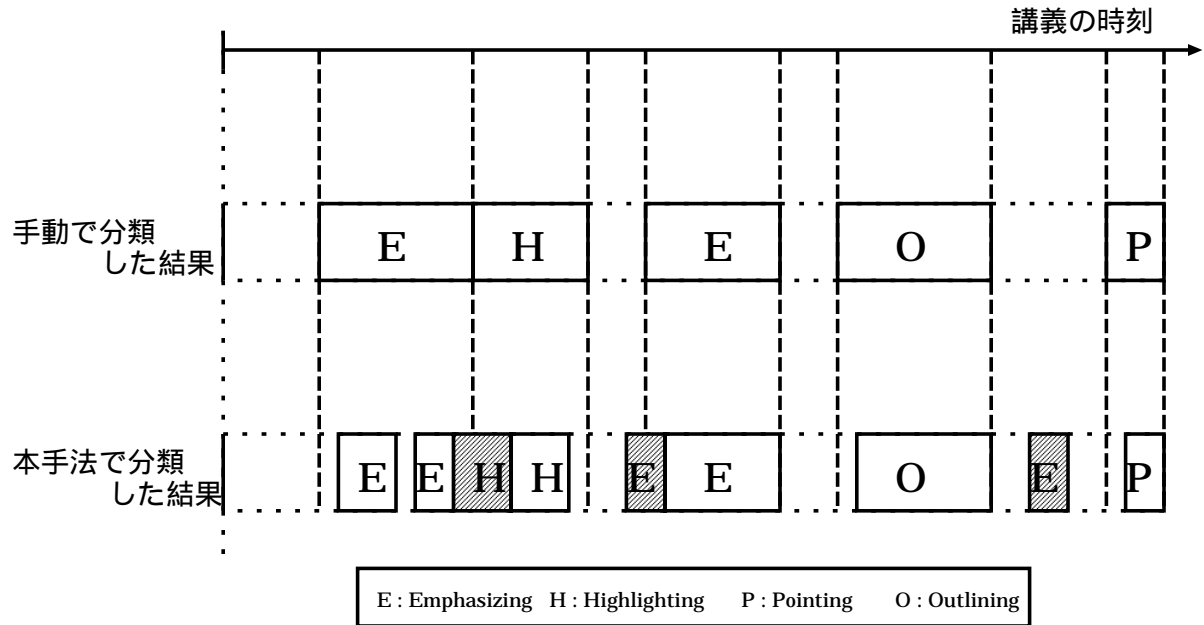


図 21: 比較方法の概要

上述の 6 種類の講義に関して分類成功率を求めた結果が表 5 である．

表 5: アンケート結果

	講義の種類	授業時間	分類成功率
1	講師 A による一つ目の講義	10:21	77.4%
2	講師 A による二つ目の講義	11:36	84.0%
3	講師 B による一つ目の講義	8:14	82.2%
4	講師 B による二つ目の講義	10:01	94.6%
5	講師 C による一つ目の講義	8:55	85.8%
6	講師 C による二つ目の講義	10:54	72.8%

実験の結果，良い結果では 94.6% の分類成功率を，平均として 82.8% の分類成功率で分類を行うことができることを確認した．1 や 6 のような成功率が低

い講義について詳しく調べた結果、以下のような原因が明らかになった。

- 指示動作を行っている区間の抽出が不完全
- 表示されているスライドの構成が複雑

上記の原因のうち最も成功率の減少に影響を与えたものは一つ目の原因である。4.3.2 節でも述べたとおり、講師の指示は指示棒先端が表示教材にほぼ接触しているときに行われるものとして、本研究では指示動作を行っている時間は指示棒先端が表示教材にほぼ接触していると判断できる時刻を検出していた。しかし、講師が指示棒を移動させるときにも指示棒の先端を表示教材にほぼ接触させていた場合、その移動軌跡中に通過した領域も本研究の手法では検出してしまい、通常なら指示していないと思う場合に対しても、指示を行ったと解釈してしまう問題が生じる。表 5 の結果に対しても、手動分類では指示していないと判断できる時間区間を除いて成功率を計算すると値は向上するため、指示動作を行っている区間の抽出は今後の課題として重要な点の一つである。二つ目の原因に関しては、特にスライドの構成が多く小さな領域が密接していたりすることが問題になる。このような場合には指示点が短時間に複数の領域上を移動し、軌跡や単一の領域上にある時間の計算で安定した値が出にくく、分類がうまくいかないことが考えられる。

上述のような成功率の減少の要素はあるものの、結果として最大 94.1% の成功率を出すことができたので、ある程度本手法の有効性は確認できた。

4.5 教材指示情報の提示の実験と評価

この節では本研究で提案した手法で抽出した教材指示情報を図 20 の場合と同様の方法で表示した場合に、指示情報の把握が容易であるかと教示内容の理解がどの程度であるかについてアンケートにより評価を行った。

実験内容は以下の通りである。4.4 節で撮影した 6 種類の映像のうち、正解率の高かった講義に関して、従来の講義コンテンツによる提示映像 M と本手法による教材指示情報を付加した講義コンテンツの提示映像 M' を作成した。次に被験者 6 人を 2 つのグループ (A, B) に分け、A グループにはまず M を、次に M' を見てもらい、B グループはまず M' を見てもらい、次に M を見ってもらうということを行い表 6 のアンケートに回答してもらった。

アンケートの結果は表 7 のようになった。表中の A, B はそれぞれグループを、数字は評価項目の点数を平均したものである。

表 6: アンケート項目

	良					悪
<一つ目の映像に関して>						
講義内容はわかりやすかったか？	:	5	4	3	2	1
映像として見やすかったか？	:	5	4	3	2	1
集中して受講することができたか？	:	5	4	3	2	1
何を指示しているかがわかりやすかったか？	:	5	4	3	2	1
<二つ目の映像に関して>						
講義内容はわかりやすかったか？	:	5	4	3	2	1
映像として見やすかったか？	:	5	4	3	2	1
集中して受講することができたか？	:	5	4	3	2	1
何を指示しているかがわかりやすかったか？	:	5	4	3	2	1
<指示に関する情報がついた映像に関して>						
指示している位置がわかりやすかったか？	:	5	4	3	2	1
指示していると示されている範囲は適切だったか？	:	5	4	3	2	1
指示していると示されているタイミングは適切だったか？	:	5	4	3	2	1
指示していると示されている表示時間は適切だったか？	:	5	4	3	2	1

表 7: アンケート結果

< 従来の講義コンテンツ映像 M に関して >	A	B	総合
講義内容はわかりやすかったか?	2.6	2.4	2.5
映像として見やすかったか?	2.8	2.2	2.5
集中して受講することができたか?	2.8	2.8	2.8
何を指示しているかがわかりやすかったか?	1.4	1.4	1.4
< 本研究の講義コンテンツ映像 M' に関して >	A	B	総合
講義内容はわかりやすかったか?	3.0	3.0	3.0
映像として見やすかったか?	2.4	2.4	2.4
集中して受講することができたか?	2.6	3.2	2.9
何を指示しているかがわかりやすかったか?	2.4	2.6	2.5
< 教材指示情報を強調表示した映像に関して >	A	B	総合
指示している位置がわかりやすかったか?	2.4	3.0	2.7
指示していると示されている範囲は適切だったか?	2.2	2.6	2.4
指示していると示されているタイミングは適切だったか?	2.2	2.8	2.5
指示していると示されている表示時間は適切だったか?	2.2	3.8	3.0

まず従来の講義コンテンツ映像と本研究の講義コンテンツ映像を比較すると、講義のわかりやすさに関しては多少の改善が確認された。また、指示に関する情報については大幅に改善が見られたため、本研究での目的はある程度達成されたと考えられる。ただ、映像としての見易さに関しては提示方法のせいかわきがやや下がっていることが確認できた。

次に本研究の教材指示情報を強調表示させたコンテンツ映像に関しては全項目に対して評価が低めであり、提示方法について問題があることが考えられる。しかし4.3.1節の場合の映像作成と比べ、すべてを自動で行った今回の場合、例えば人なら一度指示された場所ならばしばらくはそこが指示されている状態が続いていると判断するような場合でも、自動分類ではセンサ情報でそこに指示点がないので指示していると見なさないというように、人の知覚で捕らえる教材指示情報とセンサ情報のみによる教材指示情報との間になんらかのギャップが存在し、指示情報付加による影響がさほど出なかったことも考えられる。

なお、アンケート後に、任意に被験者に意見を求めたところ以下のような意見が寄せられた。

- 指示情報の表示の頻繁な切り替わりが気になる
- 対象によっては表示が邪魔をする
- スライドの色の使用状況により強調表示に使う色使いも考えたほうが良い
- 文か図かに応じて見せ方を変えたほうが良い

上述の意見からもわかるように、教材指示情報の提示法に関してもより検討の余地があることがわかる。

第5章 結論

本稿では講義コンテンツ視聴時における教示内容の理解を助けるために講義から教材指示情報を抽出し、講義コンテンツ上で教材指示情報の強調表示を行う手法について説明した。本研究では、教材指示情報の抽出を、ある時間区間における教材中で指示されている領域の特定と位置づけ、そのために以下の処理を行った。

- 表示されているスライドの構成要素ごとへの領域分割
- 指示動作が行われている時間区間の抽出
- 指示動作の種類に応じた講師の指示意図領域の推定と指示領域の変更時刻の検出

また、これらの以上の情報を提示するための強調表示の方法についても定義した。

そして教材指示情報の抽出の成功率を確かめるために、本研究の手法により自動で抽出した結果と手動で抽出した結果との比較を行ったところ、平均で82.8%、最高で94.6%という成功率を示した。また、教材指示情報を提示したことによる効果を確認するため、教材指示情報を強調表示したコンテンツ映像と従来のコンテンツ映像とをアンケートにより比較したところ、提示に関しては何を指示しているかという点が従来よりも明確になったという結果が得られた。

本研究では講義をアーカイブ化し、提示する枠組みの下で、視聴者の教示内容の理解を助けるために、講義中の講師の指示動作に着目し、表示教材に対する講師の指示の情報の自動抽出と、その情報の講義コンテンツ上での提示を行った。その過程において明らかになった課題は以下のとおりである。

- 指示動作が行われている時間区間の抽出が指示棒の接触時とした場合、指示棒の移動時間も指示と認識してしまうことがあるので、指示動作に関してさらに調査し、指示動作が行われている時間区間の抽出に関して改善する必要がある
- 教材指示情報の提示方法に関しても、実験で何を指示しているかについては従来よりも良いという結果が示されたが、提示された映像自体への評価は低かったため、教材指示情報の表現方法についてさらに検討する必要がある。

また、一般に指示は動作によるものだけでなく、「この」「あの」といった指示語を用いた音声によっても行われる。とくに音声と動作を協調させた指示を

用いることの効果は Kelly らの研究 [5] により示されている。そこで、視聴者による教示内容の理解を助ける場合、講師の音声から指示語を検出し、指示語と指示動作双方により指示されている対象を特定、提示することは効果的であると考えられる。また、音声による指示の場合、指示対象は指示動作だけでなく発話内容からもある程度推定されるため、発話中のキーワードと、スライド中のキーワードとのマッチングのような形で指示対象の推定を行うことができる。そのため、将来的には教材指示情報の抽出のためには

- 講師の発話からの指示語やキーワードの自動抽出
- 表示スライド中の文領域における語句解析，キーワード検出

のような課題についても取り組むことが有用であると考えられる。

謝辞

最後に、本研究を進めるに当たり、御指導を賜りました美濃導彦教授に深く感謝の意を表します。本研究の審査員を務めていただき、多くの助言を下された奈良先端大学の木戸出正繼教授と本学美濃研究室の角所考助教授にも深く感謝の意を表します。そして、日頃より熱心に指導して下さい、本論文の作成に対しても多大なる助言を頂きました八木啓介助手と西口敏司助手に心から感謝致します。また、実験のアンケートに関して相談にのっていただいた京都外国語大学の村上正行講師にも感謝致します。最後に、研究会等、様々な機会に数多くの助言を頂いた美濃研究室の皆様にも感謝致します。

参考文献

- [1] 前田 香織,相原 玲二,大槻 説乎:”遠隔講義のためのマルチメディア教材提示システム”, 情報処理学会論文誌,vol.40,No.1,Jan,1999
- [2] 先山 卓郎,大野 直樹,椋木 雅之,池田 克夫:”遠隔講義における講義状況に応じた送信映像選択”, 電子情報通信学会論文誌,vol.84D-,No.2,pp.248-257,Feb,2001
- [3] Sugata Mukhopadhyay,Brian Smith:”Passive Capture and Structuring of Lectures”, Proceedings of the 7th ACM international conf. on MM (Part1),1999,pp.477-487
- [4] 椋木雅之,西口敏司,美濃導彦:”複数台カメラによる講義の観測と撮像”, コンピュータビジョンとイメージメディア,131-11,pp.77-84,Jun,2002
- [5] Kelly,S.D,Barr,D.J,Church,R.B,Lynch,K:”Offering a hand to pragmatic understanding:The role of speech and gesture in comprehension and memory”, Journal of memory and Language,40,577-592
- [6] McNeil D :”Hand and Mind:What gestures reveal about thought”, The University of Chicago Press.
- [7] WIDE project, School of Internet workgroup, <http://www soi.wide.ad.jp/>
- [8] Lilian Leivas Pozzer-Ardenghi,Wolff-Michael Roth :”Gestures:Helping Students to Understand Photographs in Lectures”, Connections '03,pp.1-30.
- [9] Roth, W-M.,Lawless,D.V.:”When up is down and down is up:Body orientation,proximity,and gestures as resources”, ,Language in Society,31,1-28,2002
- [10] 先山卓郎 :”講義映像の撮影および遠隔講義室への送信映像に関する研究”

予備審査報告書に対する回答

修正が必要な点(1)

この研究では指示に関する情報を講義アーカイブに対応付けることを目的としているが、指示に関する情報と、それを反映させた表示の関係に関して明確に述べられていないため、提案手法の妥当性が不明であり、その点を明確にする必要がある。

(回答):指示動作は各時刻での教示内容と表示教材とを関連付けるものであり、各時刻でどこを指示していたかがわかることでその関連が明確になる。本研究ではどの時間帯にどのような指示動作をしたかを解析し、それを提示することで教示内容と表示教材との関連を明確にする。

修正が必要な点(2)

指示に関する情報を反映させた表示について具体的に示す必要がある。

(回答):実装し、その結果を提示し、実験によりその有効性を評価する。

修正が必要な点(3)

講義における指示動作を4つに分類する必要性について、明確に示す必要がある。

(回答):2.2.2節で述べたように、4つの指示動作により講師の指示で意図される領域が異なってくる。講師の発話内容と資料の関わりにおいて講師の指示が意図する領域の明確化は教示内容理解の上でも、分類がなされる必要がある。

修正が必要な点(4)

この研究で提案する分類手法の有効性に関して、手法による分類結果と人手による分類結果を比較し、評価する必要がある。

(回答):実験の章で評価する。

修正が必要な点(5)

従来研究との違いを明確にする必要がある。

(回答):従来の講義コンテンツの枠組みでは、表示教材映像は提示されていたものの、教示内容との関係が不明瞭であり、その点の理解が負担であった。それに対し、本手法でいつどの領域を指示しているかを明確にすることで、関係を把握し、教示内容の理解を助けることが期待できる。