

大学における実運用のための講義自動アーカイブシステムの開発

西口 敏司[†] 亀田 能成^{††} 角所 考^{†††} 美濃 導彦^{†††}

Automatic Lecture Archiving System for Practical Use at Universities

Satoshi NISHIGUCHI[†], Yoshinari KAMEDA^{††}, Koh KAKUSHO^{†††},
and Michihiko MINOH^{†††}

あらまし 本論文では、大学における講義室での実運用が可能な講義自動アーカイブシステムの構成法を提案する。講義の自動アーカイブ化においては、講義室から様々な種類の情報を獲得して講義室の状況を推定する必要があるため、機能のモジュール化、機器の仕様や制御に関する情報の一元管理、また講義状況の推定に必要なデータのキャッシュデータベースからの獲得、の三つの特徴をもつアーカイブシステムの構成法を提案する。これらの構成法に基づく講義自動アーカイブシステムを実際の講義室に実装し、約1年半にわたり実運用を行った。その結果、ハードウェアの追加や状況認識アルゴリズムの更新のしやすさなど、運用上の様々な状況に応じて柔軟な運用が可能であることを確認した。

キーワード 講義の自動アーカイブ化、マルチモーダルシステム、モジュール構造、実用システム

1. ま え が き

日本の大学教育では、講義室に集まった受講者に対して講師が教科内容を教示するという講義形態をとることが多く、日々何百という数の講義が行われている。このような講義は、同じ時刻・場所にいる学習者のみに利用可能であり、それ以外の時刻・場所で再利用することはできない。このため、復習や遠隔地での学習に再利用することを目的として、講義中の講師や受講者の様子を撮影した映像や、表示されたスライド教材を電子的に記録・保存してアーカイブ化し、いつでも、どこでも講義を聴講できる形で利用者に提供するサービスが、国内外の大学で始まりつつある [1], [2]。

講義のアーカイブ化では、講師の身振りや質問者の表情など、講義の様子を伝える上で撮影すべき被写体が状況とともに変わることから、それに応じて撮影カ

メラを操作できることが望ましい。このとき、講義における様々な状況を記録するには、複数台のカメラで講義の様子を撮影することが望ましいが、1人で複数台のカメラを操作することは難しい。

このため、講義の状況を認識して複数台のカメラで自動撮影することを目指した研究が行われている [3] ~ [6]。これらの研究では、観測カメラやマイクロホンアレーなどのセンサ機器を用いて人物位置や話者位置を推定して自動撮影を行っている。しかしながら、これらの研究で構築されているシステムは、人物位置推定や話者位置推定などに関する提案手法の有効性を評価するための実験システムであるため、観測する空間も小さく、利用するセンサ機器の数は比較的少ない。また、位置推定や撮影の対象となる講師や受講者の数が合わせて4~5人、多くとも10人程度の規模にとどまっており、実運用を考慮したシステムにはなっていない。

そこで本論文では、マイクロホンアレーなどのセンサ機器や、パン・チルト・ズームの自動制御が可能なカメラなどの制御機器等、多種多様な機器を使用し、かつ個々の様々な技術を利用して講義自動アーカイブシステムを構築することを考える。このとき、アーカイブに対する要求に応じて柔軟に機能を組み合わせることを可能とし、かつ毎日行われる実際の講義を継続的に漏れなくアーカイブ化することが可能な、実運用

[†] 京都大学大学院法学研究科，京都市
Graduate School of Law, Kyoto University, Yoshidahonmachi,
Sakyo-ku, Kyoto-shi, 606-8501 Japan

^{††} 筑波大学大学院システム情報工学研究科，つくば市
Graduate School of Systems and Information Engineering,
University of Tsukuba, 3M304, The 3rd Cluster, 1-1-1
Tennodai, Tsukuba-shi, 305-8573 Japan

^{†††} 京都大学学術情報メディアセンター，京都市
Academic Center for Computing and Media Studies, Kyoto
University, Yoshida Nihonmatsu-cho, Sakyo-ku, Kyoto-shi,
606-8501 Japan

のための講義自動アーカイブシステムを構築する。これを実現するために、一時的な実験のためのアーカイブシステムとは異なり、以下のような点を考慮したシステムを提案する。

第一に、大学における講義室にアーカイブシステムを導入する際には、講義室の広さや形、また座席の配置など、利用するセンサ機器の数や種類、それに基づく状況認識方法を変える必要がある。また、同じ講義室でも、講師位置の認識精度の向上や、講師や受講者を撮影する複数の撮影カメラの制御ルール、また撮影した複数映像の中から記録する映像を一つ選択するルールなどに関しては、どのような映像を記録したいかによって、様々な要求が存在する。このような要求にこたえつつ、実運用を可能とするためには、本論文で述べる多種多様なセンサ機器や制御機器を用いたシステムでは、ある部分の変更が、他の部分に与える影響をできるだけ小さくし、次のアーカイブ運用に影響を与えない仕組みをもつ必要がある。

第二に、センサ機器からデータを獲得し、制御機器を制御するためには、これら多数の機器がどの計算機などのインタフェースに接続されているかという情報が必要である。また観測対象の位置を推定する場合など、センサ機器や制御機器自身の位置や内部パラメータに関する情報が必要な場合もある。このような、それぞれの機器に関する様々な情報が個々に管理されていると、それらの情報間の整合性を保つことが困難となる。したがって、アーカイブシステムで用いる多数の機器に関する情報を統一的に管理できる枠組みが必要である。

第三に、講義室内の状況推定は、様々なセンサ機器で獲得される複数のデータを利用して行われるが、各センサ機器においてデータの獲得にかかる時間は、それぞれのセンサ機器の能力やデータの種類に依存して様々である。すなわち、状況認識に必要な複数種類のデータを獲得してから状況認識を完了するまでにかかる時間は、データ獲得に最も時間がかかるセンサ機器に依存する。したがって、時間同期していない複数のセンサ機器からのデータを直接利用する方法では、撮影カメラなどの機器の制御はリアルタイムに行うことができない。

そこで、時間同期していない複数のセンサ機器で獲得されるデータに基づいてリアルタイムに状況認識を行うための枠組みを導入する必要がある。

本論文では、以上の三つの問題に対処するための方

法として、それぞれ、各種機能のモジュール化、機器情報の一元管理化、及び情報共有データベースのキャッシュサーバ化、の3点を施したシステム構成により、実運用可能な講義自動アーカイブシステムを実現する。

本論文は以下のような構成である。まず2.では、講義自動アーカイブシステムの一般的な構成を述べ、これに対する実運用上の要求条件について議論する。3.では、この要求条件を満足するための具体的なシステムの構築方法について提案し、4.でこれに基づくシステムの具体的な実装例及び実運用の結果について述べる。そして最後に5.でまとめと今後の課題を述べる。

2. 実運用のための講義自動アーカイブシステム

2.1 講義の自動アーカイブ化

従来の講義アーカイブシステムでは、撮影カメラの操作や、映像切替操作は人間が行っていた。また、講師に自動追従する単独カメラによる撮影システムでは、被写体は講師のみであり、しかも視点が固定されるため、講師の正面に近い位置から撮影した映像を撮影することなどができない。一方、複数台の撮影カメラを設置し、講師や受講者の状況に応じた撮影や映像選択等ができれば、アーカイブの利用者は話者の表情や身振りを確認しながら説明を聞くことなどが可能となり、より高い臨場感を得ることが可能となる。

このような、講義室における様々な状況に応じたアーカイブ化を自動化するには、多種類のセンサを講義室に設置し、画像処理技術や音声処理技術を用いて講師や受講者の状況等を認識しつつ、対応する機器を制御しなければならない。すなわち、講義データ記録用の機器に加えて、位置センサ、マイクロホンアレー、観測用カメラなど、講義状況の認識にかかわる観測データ獲得のための多様な複数のセンサ機器が必要となる。例えば、位置センサは、講師に装着して、その位置と向きを精度良く検出するために用いられる[9]。また、マイクロホンアレーは講義室内の音源の位置を推定するために用いられる[10]。観測カメラは獲得された画像から講師や受講者の講義室内の位置を推定するために用いられる。このほかにも、魚眼レンズ付きカメラや明度制御可能な照明機器など、講義室の広さや費用、講義状況認識のアルゴリズムなどに応じた機器が利用される[12]。

このように、講義自動アーカイブシステムでは、講

義室内に設置した複数のセンサ機器で獲得されるデータに基づいて講義の状況を推定し、それに基づいて撮影カメラや映像切替器を制御する。

2.2 実運用のための要求

2.2.1 機器やソフトウェアの更新

従来研究のシステムは、個々の手法を評価するための実験システムであるため、一般に単一構造のソフトウェアとして開発されてきた。

一方、大学の講義室の広さは数十人規模から数百人規模まで様々であり、また、その形や座席の配置なども様々である。したがって、大学の講義室に講義自動アーカイブシステムを導入するには、講義室の広さや形、座席の配置に応じて、機器の種類や個数及び認識手法を柔軟に設定できる必要がある。

これらの処理を、計算機 1 台で動作させることを想定した単一構造のソフトウェアで実現すると、センサ機器や制御機器の更新によりセンサ機器のデータ獲得方法や制御機器の制御方法が変更された場合、または位置推定や話者状況認識などの処理がより優れたものに更新されたりした場合に、ソフトウェア全体を再構築しなければならない。更に、単一構造のソフトウェアは 1 台の計算機上で実行されるので、複数のセンサ機器を利用する場合や、撮影カメラなどの制御機器の制御に必要な処理をリアルタイムに行うには、接続インタフェースが多く、かつ高速度処理が可能な高性能な計算機を必要とするなど、入手のしやすさ及び金銭的コストの観点から、機器選択上の制約が強い。このため、規模が大きなソフトウェアは追加変更を容易とするためにモジュール化されることが多い。その際、どのようにモジュール化するかは、ソフトウェアが扱う問題に依存する。

そこで本論文では、実運用の観点から機器やソフトウェアの構成に関してモジュール構造を採用し、どのようなモジュール化を行えばよいかについて検討する。

2.2.2 機器に関する情報の参照

センサ機器からデータを獲得し、制御機器を操作するには、それらの機器が接続されている計算機と接続インタフェースの情報（接続情報）が必要である。また、獲得されたデータから特徴量を抽出するには、機器の位置情報（外部パラメータ情報）や内部仕様情報（内部パラメータ情報）が必要な場合がある。例えば、複数の観測カメラで撮影された画像を用いてステレオ視を行う場合には、カメラ自身のキャリブレーション情報が必要である。更に、該当機器からのデータの獲

得や、該当機器の制御を目的として構築されるソフトウェアがソケット間通信を行うための情報（サーバポート情報）が必要である。このように、ある機器に関して利用する機能に応じて様々な情報が存在する。これらを総称して機器情報と呼ぶ。

ある機器を利用する際には、その機器の利用目的によって、どの機器情報を参照する必要があるかが異なる。そこで、これらの情報を必要とするソフトウェアそれぞれが自分に必要な情報のみを保持していると、機器を接続する計算機やインタフェース、または機器自体を変更した場合に、機器に関する情報が更新されるソフトウェアと更新されないソフトウェアが混在する状況が生じるため、モジュール化されたそれぞれのソフトウェアにおいて機器情報の整合性を保つことが困難となる。

そこで本論文では、それぞれの機器情報を、1 個のファイル（機器情報ファイル）に格納して、複数のソフトウェアで機器情報を共有し、整合性を保つ枠組みを導入する。

2.2.3 非同期特徴量を用いた状況認識

2.1 で述べたように、ある時刻における講義室の状況の推定を行うためには、複数のセンサ機器からデータを獲得する必要がある。ところが、各センサ機器がデータを獲得するのにかかる時間は異なるため、完全に同時刻にすべてのセンサデータを獲得することは困難である。したがって、データ獲得に時間がかかるセンサ機器を利用する場合には、制御機器をリアルタイムに制御することは難しい。

そこで本論文では、特徴量を抽出するモジュールがそれぞれ常にデータを獲得して特徴量を抽出し続け、特徴量を出力するモジュールがその最新の抽出結果を利用することにより、特徴量の要求から制御機器に制御命令を伝えるまでの時間を短縮する仕組みを提案する。

3. 実運用に必要な機能の実現法

3.1 機能の 4 群モジュール化

3.1.1 モジュール化の単位

講義自動アーカイブシステムは、講義室内の状況を観測し、状況に応じた制御や記録を行う情報処理システムである。状況を観測するには何らかのセンサが必要となり、状況に応じた記録を行うには、パン・チルト・ズームが可能なカメラや、映像切替器などの機器を制御する必要がある。したがって、各センサからの

データを入力とし、何らかの機器を制御するデータを出力とする情報処理システムを考えたとき、講義アーカイブシステムが扱うデータは、以下に挙げるように、5種類に分類することができる。

センサデータ (sensor data: sd) 観測カメラから得られる画像や、機器固有の位置データなど、機器やメーカー固有の形式で獲得されるデータ

制御データ (controlling data: cd) パン・チルト・ズームが可能なカメラや映像切替器など、機器やメーカーに固有のコマンド形式をもつ制御データ

一次特徴量 (primary feature: pf) 位置センサの座標値や背景差分の値など、センサデータの種類に応じて初めて解釈された原始的な特徴量

二次特徴量 (secondary feature: sf) 講師の向きや、話者の位置など、一回以上の特徴量の変換で表現された特徴量

整形特徴量 (formatted feature: ff) 右に 30 度パン、何番のカメラ映像に切り換えるかなど、制御データに変換する前の機器固有の依存性を低減した特徴量、及びメディアとして記録するための特徴量

講義アーカイブシステムでは、センサからのデータを入力とし、機器を制御するデータを出力とするシステム上のデータは上記のいずれかに分類可能であるため、データの種類としては、上記定義で十分であるといえる。また、上記データの種類のいずれかが欠ければ、本論文で提案する講義自動アーカイブシステムは構築できないため、上記5種類のデータは必要条件でもあり、結果として、上記の5種類のデータは必要十分であるといえる。

センサデータは、処理を加えられることによって変換されていくが、どの種類のデータからどの種類のデータへ変換されるかは、上記の定義から、表1に示すような制約が存在する。例えば、一次特徴量 (pf) は、二次特徴量 (sf) または整形特徴量 (ff) に変換され得る。

この表から、最も詳細な単位では、6種類の変換の

表1 データ変換の組合せの制約
Table 1 Constraint of conversion combination.

Converted from	Converted to				
	sd	pf	sf	ff	cd
sensor data	sd	—	—	—	—
primary feature	pf	—	—	—	—
secondary feature	sf	—	—	—	—
formatted feature	ff	—	—	—	—
controlling data	cd	—	—	—	—

組合せが存在することが分かる。しかしながら、一般に、情報処理システムは、どのような種類の出力を得たいかという要求に応じて設計されるため、本論文においては、表1における変換先のデータの種類の4種類あることから、変換先の種類が等しい以下の4種類の機能単位群を考え、これをモジュール群と定義する。

特徴量抽出 (feature extraction): pf を出力

特徴量変換 (feature conversion): sf を出力

特徴量出力 (feature output): ff を出力

機器制御 (device controlling): cd を出力

また、モジュール群に含まれる個々の機能単位をモジュールと呼ぶ。

図1に、4種類に分類されたモジュール群と、表1に基づくモジュール間のデータの流れを示す。この図に示すように、本論文で考えるシステムでは、センサから入力されたデータは、最終的には制御データとして機器を制御するか、または整形特徴量として記録される。

このような枠組みのもとで新しい機能を実現するモジュールを追加する際には、その入出力の種類によってどの群に追加すればよいかを決定することができる。

3.1.2 モジュール間の通信

3.1.1 で提案したモジュール群間の通信手段としては、特徴量の共有のしやすさを考慮し、抽出されたすべての一次特徴量 (pf)、二次特徴量 (sf) を保持するデータベースサーバ (Lecture Room DataBase Server: LRDBS) を構成し、特徴量抽出モジュール群、特徴量変換モジュール群及び特徴量出力モジュール群の各モジュールをクライアントとするサーバクライアント型の通信モデルを採用する。

このとき、クライアントとなる各モジュールにおい

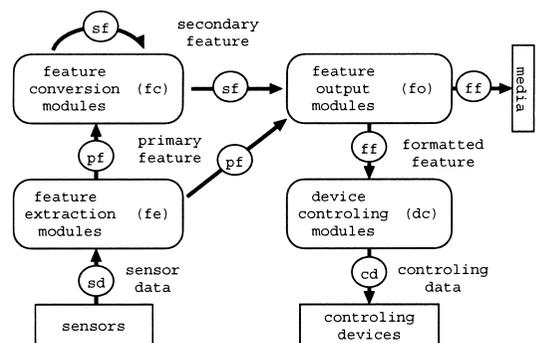


図1 モジュール群のアーキテクチャ
Fig.1 Architecture of groups of modules.

ては、LRDBS から複数のデータを獲得して何らかの処理を行い新たなデータを出力する場合、本来必要とするデータの一部が何らかの理由で獲得できないときに、獲得できたデータのみを用いて処理を続けるような実装を行う。これにより、例えば複数種類のデータを用いて精度を高める処理を行っているモジュールでは、精度は悪くなるが、処理を続けることが可能となる。

一方、特徴量出力モジュールで生成され、機器制御モジュールに渡される整形特徴量は、それ以外のモジュールで利用されることはないので、LRDBS を経由せずに、機器制御モジュールに直接伝達する。機器制御モジュールは、特徴量出力モジュールから整形特徴量を受け取り、機器に依存した制御データに変換してシリアルポートを介して制御機器を制御するサーバプログラムとして実装する。すなわち、特徴量出力モジュールは、LRDBS からデータを獲得するクライアントであり、かつそれを機器制御モジュールに伝達するクライアントでもある。このように、機器制御モジュールをサーバプログラムとして実装することにより、どの特徴量出力モジュールの出力を受け取るかなどの変更が容易となる。

図 2 に個々のモジュールの例とサーバクライアント型通信方式のデータの流れを示す。図に示す各データの流れのうち、実線は機器が接続されたシリアルポートを介してデータが流れる部分であり、破線は LAN 上をデータが流れる部分であることを示す。特徴量抽出モジュールと機器制御モジュールは、ネットワーク負荷の低減を考慮して、それぞれセンサ機器、制御機器が接続された計算機で稼動しており、特徴量抽出モジュール、特徴量変換モジュール、特徴量出力モジュール間の通信は、LRDBS を介して行われる。また、整形特徴量は、特徴量出力モジュールから機器制御モジュールまたは記録媒体に直接送られる。

3.1.3 モジュール間の独立性

階層的なモジュール化によってモジュール間の独立性がどの程度になったかを評価するために、モジュールの実装例として図 2 に示した個々のモジュール（全 28 個）のいずれか 1 個が欠損した場合に、処理を続行できないモジュール（fatally affected modules）及び何らかの影響が及ぶモジュール（partially affected modules）を表 2 に示す。表 2 における各モジュールの識別子及びデータの識別子は、図 1 及び図 2 に示してあるものと同じである。

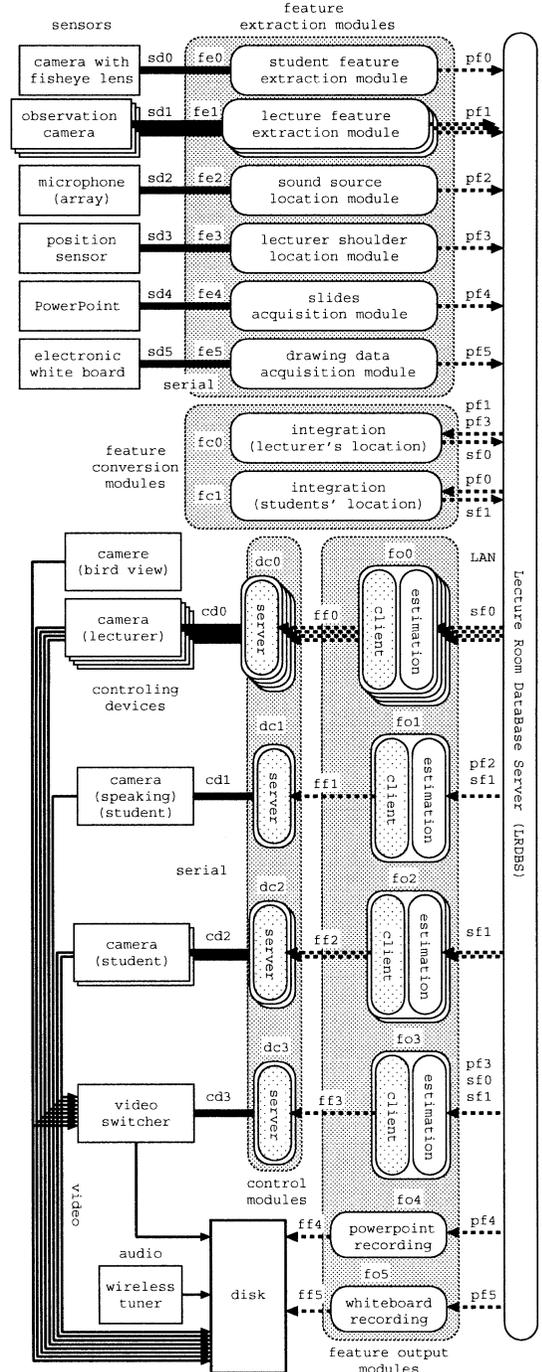


図 2 モジュール群の通信方式
Fig. 2 Transmission method of modules.

例えば、図 2 におけるモジュール fe0 が欠損した場合、fe0 の出力 (pf0) のみを受け取って処理するモジュール fc1 が処理を続行できなくなって sf1 を出力

表 2 モジュールの欠損が他に与える影響
Table 2 Interdependency between modules.

deficit module	out data	fatally affected	partially affected
fe0	pf0	fc1→fo2→dc2	fo1→dc1, fo3→dc3
fe1	pf1	-	fc0→fo0→dc0, fo3→dc3
fe2	pf2	-	fo1→dc1
fe3	pf3	-	fc0→fo0→dc0, fo3→dc3
fe4	pf4	fo4	-
fe5	pf5	fo5	-
fc0	sf0	fc0→fo0→dc0	fo3→dc3
fc1	sf1	fo2→dc2	fo1→dc1, fo3→dc3
fo0	ff0	dc0	-
fo1	ff1	dc1	-
fo2	ff2	dc2	-
fo3	ff3	dc3	-
fo4	ff4	-	-
fo5	ff5	-	-

できず、その結果、sf1 のみを受け取って処理するモジュール fo2 及びその出力 (ff2) を利用する dc2 が処理を続行できなくなる。一方、sf1 だけでなく他のデータを利用するモジュール fo1 や fo3 は、処理の精度等は落ちるものの、sf1 が受け取れないという事実及び他のデータを利用して処理を続行する。

表 2 から、図 2 における実装例では、全モジュールの数が 28 個あるのに対し、1 個のモジュールの欠損が、処理の続行が不可能になるモジュールの数はたかだか 3 個 (全体の 10.7%) であり、更に影響を受けるモジュールの数を合わせても、たかだか 7 個 (全体の 25%) であることが分かる。

3.2 多種多様な機器情報の一元管理化

2.2.2 で述べた機器情報ファイルに基づいて、モジュールが参照する機器情報の整合性を保つ手法について述べる。

センサ機器や制御機器に関して、異なるメーカーや型番の機器であっても、同じ種類であり、かつ類似の機能が利用可能であるならば、機器情報ファイルに記述する機器情報の種類は同じである。一方、機器の種類が異なれば、内部パラメータに関する情報は大きく異なるため、異なる機器の機器情報も含めた機器情報ファイルを構成すると、機器情報ファイルを更新する人間にとって、一覧性が悪くなる。また、このような複数の種類の機器情報がまとめて記述されたファイルを読み込むパーサも機器情報の種類に応じて構築したものを機器情報の種類に応じて切り換えながら利用するという形態をとる必要があり、新しい種類の機器を追加した場合などに、読み込むパーサ自体を更新する

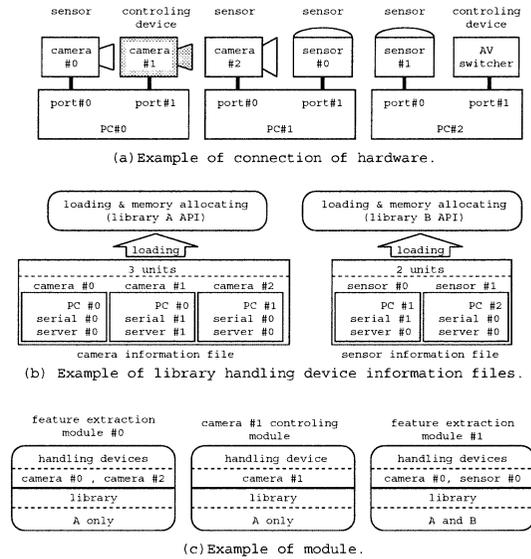


図 3 機器情報ファイルの一元管理化
Fig. 3 Consolidation of hardware information.

必要が生じる。更に、すべての機器に関する機器情報を一つのファイルに記述する方法では、あるモジュールで読み込んだ機器情報に基づいて、それぞれの機器に応じたメモリ空間を割当てたとき、利用しない種類の機器に関するメモリ空間を無駄に消費するという問題がある。

以上のような理由から、機器情報ファイルは機器の種類ごとに一元化することとする。そして、それぞれの機器情報ファイルから読み込むためのパーサを含むライブラリを、機器の種類ごとに構築しておく。図 3 (a) のような機器構成の場合の機器情報ファイルと読み込みライブラリを図 3 (b) に示す。機器に関する情報を利用するモジュールは、必要とする種類の機器の機器情報を読み込むパーサを含むライブラリのみ利用すればよい。例えば、図 3 (c) において、撮影カメラ #1 の機器制御モジュールは、カメラに関する機器情報のみを参照できればよいので、図 3 (b) に示すライブラリ A のみを利用すればよい。

それぞれの機器に関する機器情報ファイルは共有ディレクトリ内に格納し、センサ機器、制御機器を利用するモジュールは、この共有ディレクトリ内の必要な機器情報ファイルを参照する。

3.3 LRDBS のキャッシュサーバ化

データ獲得速度が異なる複数のセンサ機器やデータ処理速度が異なる複数のモジュールが存在する環境で、

ある時刻における講義状況をリアルタイムに推定する手法について述べる。

まず、一次特徴量 (pf), 二次特徴量 (sf) を保持する LRDBS をキャッシュサーバ化する。LRDBS に登録された特徴量を一時的に蓄積 (キャッシュ) して、状況認識のための特徴量変換モジュールが LRDBS からキャッシュデータを獲得することにより、複数の特徴量の獲得にかかる時間や周期の差異を吸収する。

特徴量抽出モジュールは、各センサ機器のデータ獲得間隔をできるだけ短くするために、センサ機器でデータを獲得し、一次特徴量を求めて LRDBS に登録した直後に再びセンサ機器からデータを獲得するという動作を繰り返す。

特徴量変換モジュールは、特徴抽出モジュールで抽出された一つ以上の一次、二次特徴量を統合して、新たな二次特徴量を求めて LRDBS に登録する。

LRDBS は、複数の特徴量抽出モジュールや特徴量変換モジュールから特徴量を受け取る。このとき、それぞれのモジュールごとにキャッシュ用メモリを確保し、過去に登録された複数回分の一次、二次特徴量を保持しておく。また、特徴量統合モジュールや、特徴量出力モジュールから特徴量の要求があったとき、要求された時刻の最も近い過去に登録された特徴量を返す。また、特徴量出力モジュールから過去に抽出、統合された特徴量の要求があったときにも、キャッシュ

された特徴量を返す。

特徴量出力モジュールは、ある特徴量を利用する必要のあるとき、特徴量抽出モジュールや特徴量変換モジュールに直接要求するのではなく、LRDBS に問い合わせる。これにより、LRDBS にキャッシュされた特徴量を即時に得ることが可能となる。

図 4 に、3.1 で述べた 4 群のモジュール構成のうち、特徴量抽出モジュール、特徴量変換モジュール、特徴量出力モジュール及び LRDBS と、それぞれの間のデータの流れの概要を示す。図 4 において、 t_0, t_1, t_2 が、講義状況の推定を開始する時刻である。処理速度の異なる特徴量抽出モジュールや特徴量変換モジュールから得られる特徴量を LRDBS 上にキャッシュすることにより、特徴量出力モジュールが、認識処理を開始する時刻までに得られている最新の特徴量を用いてリアルタイムに状況を認識することが可能となり、その結果、撮影カメラや映像切換器の制御をリアルタイムに行うことが可能となる。

4. 講義アーカイブシステムの実装

4.1 使用した機器

アーカイブシステムを実装した講義室の大きさは、幅 15.5 m、奥行 9.6 m、天井の高さ 3.2 m で、座席数は 120 席である。講義室に設置した機器の配置例を図 5 に示す。これらの機器は配置例であり、本論文で提案した講義自動アーカイブシステムの構成手法によって、使用する機器等を容易に増減することができる。

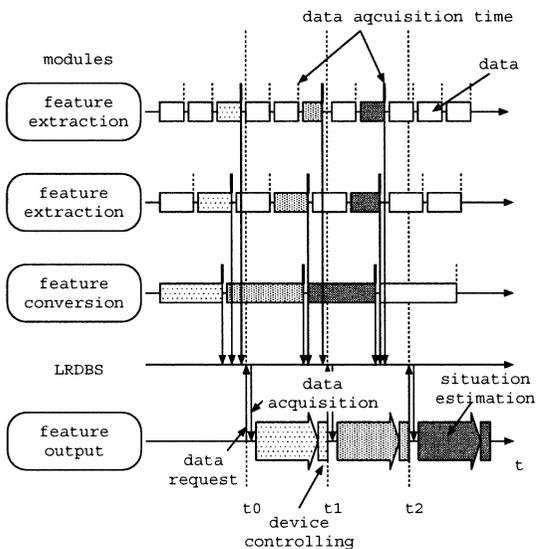


図 4 講義状況推定のための非同期データ獲得

Fig.4 Recognition using asynchronous data from multiple sensors.

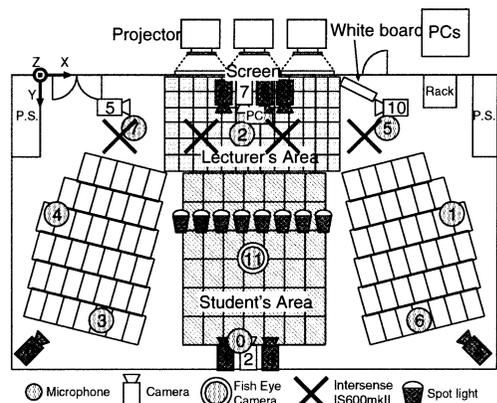


図 5 複数の機器の配置例

Fig.5 Example of installing multiple sensors and controlling devices.

観測カメラ#2, #5, #10 は、講義室の前方で歩き回る講師の頭部の位置を観測するのに利用する [4]。また、観測カメラ#11 は魚眼レンズ付きカメラであり、講義室の天井に光軸を垂直下方向に向けて受講者席全体が見渡せるように設置し、座席に着席した受講者の位置を推定するのに利用する [7], [9]。また、撮影カメラ#0, #1, #3, #4 を用いて講師を追跡撮影する [11]。撮影カメラ#8 は声を出している受講者を撮影し [10], 撮影カメラ#6, #9 を使って、受講者の動きの大きさに基づく撮影を行う [9]。また、撮影カメラ#7 は、受講者全体を固定撮影するために用いる。

マイクロホン#0~#7 は、講義室の中央の座席部分を取り囲む位置の天井付近に設置して、CSP 法を利用して講義室内の音源の位置を推定するのに用いる [8]~[10]。更に、超音波式位置センサは教卓付近の天井に設置し、講師の両肩に装着した 2 個の超小型ビーコンの位置を検出するために用いる。

また、電子白板 1 台、明度制御可能な照明機器を 8 個設置した。映像切換器は 32 入力 × 16 出力のマトリックススイッチャを用い、特徴量出力モジュール群に含まれる映像切換を担うモジュールによって、講師映像 4 本及び受講者映像 4 本の合計 8 本の映像から 1 本を選択する。

すべての観測カメラは、640 × 480 画素の解像度で動画像を獲得し、すべての撮影カメラは、320 × 240 画素の解像度で動画像を獲得する。

4.2 各機器の接続

各機器を接続する計算機として、互いに通信可能な 100 MByte の LAN に接続された PC/AT 互換機を 11 台 (それぞれ PCI スロット 3, シリアルポート 2) 使用した。それぞれの仕様を表 3 に示す。また、これらの計算機に接続するセンサ機器及び制御機器を表 4 と表 5 に示す。各計算機にはビデオキャプチャボード 2 枚、サウンドキャプチャボード 1 枚を装着して、それぞれに各カメラの映像出力、無線マイク用チューナーの音声出力を接続した。すべてのカメラを 1 台の計算機に接続することは現実的ではないので、できるだけ負荷が分散するように、計算機にカメラの映像ケーブル、シリアル制御ケーブルをそれぞれ 2 本ずつ接続した。

3.2 で提案した機器情報ファイルの記述形式は、一般に広く利用されている XML 形式とした。例えば、機器の種類がカメラの場合、カメラは複数台利用するので、利用しているカメラの台数 n に続けて、カメ

表 3 使用した計算機
Table 3 Use of computers.

use	number	CPU	clock
Controlling camera & Capturing video	6	P-III	1.0 GHz
Recording selected video	1	P-III	1.0 GHz
Controlling whiteboard	1	P-4	2.53 GHz
Controlling PowerDAQ	1	P-4	2.53 GHz
PowerPoint Presentation	1	P-III	800 MHz
Console & LRDBS	1	P-III	1 GHz
Total	11		

表 4 使用したセンサ機器
Table 4 Use of sensors.

No.	model	spec
-	IS-600 markII	position sensor serial (9600 bit/s)
#2,#5,#10	EVI-G20	serial (VISCA)
#11	TS2V114E VK-C555R	fish-eye lens C-mount camera
-	SoftBoard	serial (9600 bit/s)
#0~#7	ATL-9	microphone
-	PowerDAQ	synchronous multiple audio data aquisition

表 5 使用した制御機器
Table 5 Use of controlling devices.

No.	model	spec
#0,#1,#3,#4 #6,#7,#8,#9	EVI-D30	serial (VISCA)
-	LPL (500 W)	halogen lamp
-	Effect Arts	light control

ラの内部パラメータ、外部パラメータ、接続されている計算機、シリアル端子、制御用サーバのポート番号、キャプチャボードのデバイス名等をカメラの台数分を、機器情報ファイルに記述する。

カメラの情報を管理するライブラリは、このような情報が記述されたファイルを読み込んで、 n 台のカメラの情報をメモリ上に保持する。各カメラの情報の獲得は、カメラごとにメモリに割り当てられた構造体へアクセスすることによって行う。

4.3 講義自動アーカイブシステムの実運用

本論文で述べた手法に基づいて構築した講義自動アーカイブシステムを、2002 年度後期 6 コマ、2003 年度前期 4 コマ、2003 年度後期 7 コマの計 3 期間、約 1 年半にわたって運用した。各講義では、アーカイブシステムの動作を監視することを目的として、1 人の補助員が付いた。補助員は、講義の開始時と終了時に

システムの起動・停止を行い、障害が起こった際には、その原因に関する情報収集を行った。

表 6 に、本論文で提案した講義自動アーカイブシステムでどのようなメディアの組合せでアーカイブ化したかの種類を示す。また、2002 年度後期から 2003 年度前期までに記録した全講義アーカイブの数を表 7～表 9 に示す。約 1 年半の間に総計 181 回の講義（1 回約 90 分）、すなわち合計約 272 時間の講義が自動的に

表 6 アーカイブの種類
Table 6 Archived media.

種類	映像	電子スライド	電子白板	割合
A	自動	使用	0 台	12.1%
B	自動	使用	1 台	50.0%
C	自動	使用	2 台	6.9%
D	自動	不使用	0 台	1.7%
E	自動	不使用	1 台	24.1%
F	自動	不使用	2 台	5.2%

表 7 2002 年度後期（2002 年 10 月～2003 年 3 月）
Table 7 Second semester of 2002.

講義名	講義数	障害数	アーカイブの種類					
			A	B	C	D	E	F
情報メディア論	9	0	2	3	0	0	1	3
計算機アーキテクチャ 1	12	0	4	8	0	0	0	0
言語情報論 B	10	0	1	9	0	0	0	0
マルチメディア	11	0	1	10	0	0	0	0
パターン認識	12	2	0	10	0	0	0	0
基礎物理化学 B	9	1	0	0	7	0	0	1
合計	63	3	8	40	7	0	1	4

表 8 2003 年度前期（2003 年 4 月～2003 年 9 月）
Table 8 First semester of 2003.

講義名	講義数	障害数	アーカイブの種類					
			A	B	C	D	E	F
都市経営論 A	12	2	0	0	0	1	9	0
情報処理入門	10	1	1	7	0	1	0	0
画像処理論	10	1	0	1	0	0	8	0
パターン認識特論	10	0	0	8	0	0	2	0
合計	42	4	1	16	0	2	19	0

表 9 2003 年度後期（2003 年 10 月～2004 年 3 月）
Table 9 Second semester of 2003.

講義名	講義数	障害数	アーカイブの種類					
			A	B	C	D	E	F
比較政治学	10	0	0	0	0	0	10	0
情報メディア論	10	0	1	8	0	0	1	0
統計学 B	12	0	0	0	0	1	11	0
マルチメディア	13	0	2	11	0	0	0	0
パターン認識	11	0	0	11	0	0	0	0
基礎物理化学 B	10	0	0	0	5	0	0	5
IT ベンチャービジネス論	10	0	9	1	0	0	0	0
合計	76	0	12	31	5	1	22	5

アーカイブ化された。これらの表から、様々な回生を対象とした理科系科目及び文科系科目をアーカイブ化できていることが分かる。

講義自動アーカイブで記録する各メディアには時刻同期情報を付与し、講義終了後に SMIL で記述された同期提示用ファイルを自動生成して、すべての講義アーカイブを学内からの視聴に限定してインターネット上に公開した。

4.4 講義自動アーカイブシステムの評価

4.4.1 リアルタイム性能に関する評価

表 10 に、LRDBS と通信を行う各モジュールにおいて、LRDBS からデータを受け取って処理を行い、再び LRDBS にデータを問い合わせるまでにかかる時間の平均を示す。この表に示すように、LRDBS にデータを登録する fe 群の処理時間の平均は 0.340 秒であるが、制御モジュールに制御指示を出す fo 群の処理時間の平均は 0.218 秒であり、全体として、センサ機器でデータを獲得する時間間隔よりも早い間隔で、制御機器を制御することができた。また、最もデータ獲得に時間がかかるモジュールは音源位置獲得モジュールであり、0.509 秒間隔でデータを獲得することができるが、3.3 で提案した手法により、0.241 秒間隔で話者を撮影するための指示を送ることができた。

4.4.2 安定性に関する評価

表 6 に、アーカイブ化した講義で用いられたメディアの種類に基づくアーカイブの種類を示す。表 7～表 9 における障害数は、獲得すべきメディアのうち、障害により映像などのメディアに欠損が生じた講義の数である。アーカイブ化された講義のうち、メディアに欠損が生じたアーカイブは、全部で 7 講義であり、全体に対する割合は、3.9%であった。データに欠損が生

表 10 LRDBS の性能評価
Table 10 Performance evaluation of LRDBS.

群	モジュール	回数	処理時間 (秒)	回数/秒
fe	講師情報獲得	3	0.270	3.704
	講師位置獲得	2	0.282	3.546
	音源位置獲得	1	0.509	1.965
	動受講者獲得	1	0.300	3.333
	平均	-	0.340	-
fc	講師統合	1	0.699	1.431
	動受講者統合	1	0.300	3.333
	平均	-	0.500	-
fo	講師位置推定	4	0.286	3.497
	話者位置推定	1	0.241	4.149
	動受講者位置推定	2	0.122	8.197
	映像切替	1	0.221	4.525
	平均	-	0.218	-

じた主な原因は、講義アーカイブシステム以外の実験用プログラムが同時に起動していたことによるリソース競合によるものであり、障害検出の仕組みが脆弱であったことによる。これらの障害以外は、各講師追跡撮影カメラ、受講者撮影カメラ、電子白板、及び電子スライドに関して、これまでの章で提案した手法に基づくアーカイブ化が安定的に問題なく行われた。

4.4.3 柔軟性に関する評価

表 7～表 9 に示したアーカイブの種類分布から、様々な種類の講義アーカイブを記録できたことが分かる。このように、実際の運用では、講義を担当する講師などから、受講者のみを撮影したい、スライドは記録しないでほしいなど、様々な要求が出される状況があった。これらの要求に対しても、各機能をモジュール化していることにより、受講者位置認識モジュール及び受講者撮影カメラの制御にかかわるモジュールを起動するなど、それぞれの要求を満たすのに必要な機能を担うモジュール群のみを使用することによって、柔軟に対処することが可能であった。本項では、具体的なモジュールの追加事例を挙げる。

システムを運用開始した当初は、電子白板を 1 台用いて運用していたが、記述面積を広く使いたいという講師の要望に応じるため、3.1.3 で述べた方法に基づいて更にもう 1 台の電子白板を以下の手順で増設した。

- (1) 新規の電子白板を講義室内に設置する
- (2) 電子白板のシリアル端子を計算機に接続する
- (3) 機器情報ファイルに増設した電子白板の情報を #2 として追加する
- (4) システム起動時に、増設した電子白板からデータを獲得するモジュール（特徴抽出モジュール）及び、及びデータを記録するモジュール（特徴量出力モジュール）を、電子白板の番号（ここでは #2）を引数として与えて起動する

このように、電子白板の追加においては、データ獲得に関するモジュールプログラム自身には変更を加えることなく、新しい機能を追加することが可能であった。

制御機器としての撮影カメラを増設する場合には、機器情報ファイルに該当カメラのエントリを追加し、該当カメラを制御するアルゴリズムを記述した特徴量出力モジュール、及び該当カメラを制御するための機器制御モジュールを追加する。

また、発話する人物の有無や講師、受講者の位置に関する特徴量を用いてどの撮影カメラからの映像に切

り換えて保存するかに関するモジュールに関しても、運用期間中のアンケートなどを参考にして、切替方法や頻度などのアルゴリズムを改良していったが、特徴量出力モジュールにおけるアルゴリズムの部分のみを変更すればよいため、更新作業はスムーズに行うことが可能であった。

5. む す び

本論文では、実際の講義室に講義自動アーカイブシステムを構築して運用する際の問題点について検討し、それらに対処する方法を提案した。

自動アーカイブ化ソフトウェアを構築する際に、機器の更新やソフトウェアの更新等に対処するために、それぞれの機能をモジュール化することによって、ソフトウェアの他の部分に影響を与えずに柔軟に更新可能とする手法を提案した。

また、実際の様々な規模の講義室で自動アーカイブシステムを運用するにあたっては、多様かつ多数の機器を利用する必要があるため、システムを構成するそれぞれのモジュールが機器を利用する際に一元的に管理された機器情報ファイルをすべてのモジュールが参照する共有ディレクトリ内に配置し、それを参照して整合性を保つ手法を提案した。

更に、状況認識にかかわるモジュールが、データ獲得能力の異なる複数の種類のセンサ機器から得られるデータに基づいて講義室の状況をリアルタイムに判断するために、それぞれの特徴量をキャッシュサーバに随時登録し、このサーバに保存された特徴量を用いて状況認識を行う枠組みを提案した。

本論文で提案した機器構成及びソフトウェア構成に基づいて実際の講義室で運用可能な講義自動アーカイブシステムを実装し、約 1 年半にわたって実運用した。その結果、機器の追加・変更、アルゴリズムの更新や実際の講義における必要な機能の選択等が柔軟に行えるシステムとして構築できていることが確認できた。

一方、本論文で記録の対象とした講義はすべて一斉講義型の対面授業であり、講師による講義のみで受講者による質問がない場合は、質問をしている受講者の映像は撮影されず、また、講義中に受講者が咳をしたりトイレに立ったり私語をしたりする状況も、記録されることがあった。そこで、今後の課題としては、本論文で述べた講義自動アーカイブシステムの枠組みのもとで、より詳細な講義状況の推定に必要な、画像からの講師位置推定の精度向上や、講師の指示情報を推

定するためのアルゴリズムなどをモジュールとして実装し、利用者にとってより有用な講義アーカイブを生成する方法の検討などが挙げられる。

文 献

- [1] 大川恵子, “教育用計算機環境の事例 — SOI (School of Internet) での事例,” IPSJ Magazine, vol.45, no.3 pp.259–262, 情報処理学会, 2004.
- [2] N. Bousdira, S. Goldfarb, E. Myers, H.A. Neal, C. Severance, K.M. Storr, and G. Vitaglione, “WLAP the web lecture archive project: The development of a web-based archive of lectures, tutorials, meetings and events at CERN and at the University of Michigan,” CERN-OPEN-2001-066; Geneva: CERN, 1 Sept. 2001. (URL:http://www.wlap.org)
- [3] 大西正輝, 泉 正夫, 福永邦雄, “情報発生量の分布に基づく遠隔講義撮影の自動化,” 信学論 (D-II), vol.J82-D-II, no.10, pp.1590–1597, Oct. 1999.
- [4] 亀田能成, 石塚健太郎, 美濃導彦, “状況理解に基づく遠隔講義のための実時間映像化手法,” 情処学研報, 2000-CVIM-121-11, pp.81–88, 2000.
- [5] 大西正輝, 村上昌史, 福永邦雄, “状況理解と映像評価に基づく講義の知的自動撮影,” 信学論 (D-II), vol.J85-D-II, no.4, pp.594–603, April 2002.
- [6] 先山卓朗, 大野直樹, 棕木雅之, 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,” 信学論 (D-II), vol.J84-D-II, no.2, pp.248–257, Feb. 2001.
- [7] S. Nishiguchi and M. Minoh, “A camera controlling method for lecture archive,” Proc. International Workshop on Pattern Recognition and Understanding for Visual Information Media in Cooperation with ACCV 2002, pp.21–26, 2002.
- [8] 國田政志, 有田大作, 谷口倫一郎, “音源位置同定によるカメラの首振りに関する研究,” 情処学研報, 2002-CVIM-233-6, pp.41–48, May 2002.
- [9] S. Nishiguchi, Y. Kameda, K. Kakusho, and M. Minoh, “Automatic video recording of lecture’s audience with activity analysis and equalization of scale for students observation,” J. Advanced Computational Intelligence and Intelligent Informatics, vol.8, no.2, pp.181–189, 2004.
- [10] 西口敏司, 東 和秀, 亀田能成, 角所 考, 美濃導彦, “講義自動撮影における話者位置推定のための視聴覚情報の統合,” 電学論, vol.124, no.3, pp.729–739, 2004.
- [11] 東 和秀, 亀田能成, 美濃導彦, “パン・チルトスピード及びズーム制御による動物体の自動撮影手法,” 第6回知能情報メディアシンポジウム予稿論文集, pp.101–106, 2000.
- [12] 新宮 淳, 亀田能成, 西口敏司, 美濃導彦, “複数人物の撮影における照明光量の動的制御,” 信学論 (D-II), vol.J87-D-II, no.3, pp.829–839, March 2004.
- [13] 美濃導彦, 角所 考, “環境メディア: コミュニケーション環境として機能する情報メディア—遠隔講義を具体例として,” システム制御情報学会誌, vol.47, no.10, pp.1–6, 2003.
- [14] 村上正行, 西口敏司, 亀田能成, 美濃導彦, “講義自動撮

影システムの導入に伴う講師・受講生への影響,” 信学技報, MVE-2002-127, 2003.

(平成 16 年 3 月 10 日受付, 7 月 15 日再受付)



西口 敏司 (正員)

平 8 京大・工・情報工学卒・平 13 同大学院情報学研究科博士後期研究指導認定退学・同年同大学法学部助手・現在複数のセンサ情報を利用した講義室空間の状況把握, 講義の自動撮影及びアーカイブ化に関する研究に従事・博士(情報学)・電気学会, 人工知能学会各会員。



亀田 能成 (正員)

平 3 京大・工・情報工学卒・平 8 同大学院工学研究科博士後期研究指導認定退学・同年同大学工学部助手, 平 10 同大学総合情報メディアセンター助手・平 13~14 マサチューセッツ工科大学客員研究員・平 14 京都大学学術情報メディアセンター助手・平 15 筑波大学機能工学系講師・平 16 同大学院システム情報工学研究科助教授・モデルベーストビジョン, 画像認識とその応用, 人間を対象とするコンピュータビジョン・マルチメディア処理, 遠隔講義における IT システムの研究に従事・博士(工学)・情報処理学会, IEEE 各会員。



角所 考 (正員)

昭 63 名大・工・電気卒・平 5 阪大大学院工学研究科通信工学専攻博士課程了・平 4~6 日本学術振興会特別研究員・平 5~6 スタンフォード大学ロボティクス研究所客員・平 6 大阪大学産業科学研究所助手・平 9 京都大学総合情報メディアセンター助教授・平 14 同大学学術情報メディアセンター助教授・視覚メディア処理, コミュニケーション, インタラクションに関する研究に従事・博士(工学)・IEEE, ACM, 情報処理学会, 人工知能学会各会員。



美濃 導彦 (正員)

昭 53 京大・工・情報工学卒・昭 58 同大学院博士課程了・同年工学部助手, 昭 62~63 マサチューセッツ州立大学客員研究員, 平元京都大学工学部附属高度情報開発実験施設助教授, 平 7 同教授, 平 9 京都大学総合情報メディアセンター教授, 平 14 京都大学学術情報メディアセンター教授・画像処理, 人工知能, 知的コミュニケーション関係の研究に従事・工博・IEEE, ACM, 情報処理学会, 画像電子学会, 日本ロボティクス学会各会員。