

平成 21 年度
学士学位論文

**e-Learning における学習者の存在を
リアルタイムで検出するシステムの構築**

**Development of Real-time Learner Detection System
for e-Learning**

1100349 松本 直樹

指導教員 妻鳥 貴彦

2010 年 3 月 1 日

高知工科大学 情報システム工学科

要 旨

e-Learning における学習者の存在を リアルタイムで検出するシステムの構築

松本 直樹

近年，多くの高等教育機関で WBT (Web Based training) に代表される e-Learning を利用した遠隔教育が行われている。遠隔教育は主にインターネットを利用して行われることが多く，高等教育機関では主としてインターネットを利用した遠隔講義として実施されている。遠隔講義は大学設置基準等で単位認定が認められており，各高等教育機関で様々な取り組みが行われている。

単位認定を伴う遠隔講義では，対面での講義と同様に大学設置基準等で定められている学習時間を確保する必要がある。現状では LMS (Learning Management System) に記録される情報から学習状況を把握することができる。しかし，LMS に記録される情報からでは間接的に，そして一時的にしか学習状況を把握することができないため，学習者がコンピュータの前にはいない場合や，学習を行っていない場合を判断することができないという問題がある。

そこで本研究では，この問題を解決するために WBT に代表される e-Learning において，学習者の存在をリアルタイムで検出するシステムを構築した。そして，最後に評価実験を行い学習者の存在を正しく検出できているか確認を行った。

キーワード e-Learning, WBT, LMS, 遠隔講義, 単位認定, 大学設置基準

Abstract

Development of Real-time Learner Detection System for e-Learning

Naoki MATSUMOTO

In recent years, WBT (Web Based Training) and other e-Learning are introducing at higher education institutions such as distance learning. Distance learning mainly apply the internet. Distance lecture is one of distance learning style. Higher education institutions is conducting distance lecture using the internet. Distance lecture can accredit credit based on university establishment standards. Each higher education institutions are addressing the distance lecture.

It is necessary that distance lecture secure learning time based on university establishment standards to accredit credit. Currently, Learner's progress figure out from recorded information on LMS (Learning Management System) such as learning time and result of test. However, administrator of distance lecture can not figure out learner's situation from recorded information on LMS at directly and real-time. Therefore, administrator of distance lecture can not figure out the Learner's situation such as learn and not learn.

In this paper, we developed real-time learner detection system for WBT and other e-Learning. Moreover, we evaluated this system to confirm correctly detect the presence of learner.

key words e-Learning, WBT, LMS, Distance Lecture, Certification Unit, University Establishment Standard

目次

第 1 章	はじめに	1
第 2 章	研究の背景と目的	2
2.1	e-Learning を利用した遠隔教育	2
2.1.1	WBT (Web Based Training)	3
2.1.2	LMS (Learning Management System)	3
2.1.3	遠隔講義 (同期型, 非同期型)	4
2.1.4	遠隔講義の単位認定状況	4
2.2	単位認定を伴う遠隔講義の現状と問題点	5
2.3	本研究の目的	6
第 3 章	学習者の存在検出方法の検討	8
3.1	システムを設計する上での要求条件	8
3.2	各画像処理手法の特徴	9
3.2.1	2 値化	9
3.2.2	エッジ抽出	10
3.2.3	肌色抽出	10
3.2.4	顔検出 (物体検出)	11
3.2.5	動体検出 (フレーム間差分)	11
3.2.6	背景差分法	12
3.3	各画像処理手法の検討	13
3.3.1	2 値化とエッジ抽出を用いた手法の検討	13
3.3.2	背景差分法を用いた手法の検討	13
3.3.3	動体検出を用いた手法の検討	14

目次

3.3.4	肌色情報, 顔検出を用いた手法の検討	14
3.4	本研究での実現方法	15
3.5	本研究における学習者の検出手順	16
第 4 章	学習者の存在を検出するシステムの構築	18
4.1	システムの概要	18
4.2	基準値の取得	18
4.3	顔検出	20
4.4	肌色情報による検出	21
4.4.1	肌色情報の抽出方法	21
4.4.2	画像全体の肌色情報からの検出	22
4.4.3	肌色部分の動体検出	23
第 5 章	評価と考察	24
5.1	評価の概要	24
5.2	評価実験結果	24
5.3	評価実験の考察	29
5.3.1	被験者 A と B の評価結果に対する考察	29
5.3.2	被験者 C の評価結果に対する考察	30
	カメラに学習者の顔が入っていなかったとき	30
	学習者が大きくのけぞったとき	30
	学習者が大きくカメラに近づいたとき	31
第 6 章	比較検証と考察	33
6.1	比較検証の概要	33
6.2	カメラの種類による検証	33
6.3	照明環境の変化による検証	35

目次

6.4	装飾品の有無による検証	36
6.5	比較検証の考察	38
6.5.1	カメラの違いと照明環境の違い	38
	肌色情報の抽出結果について	38
	学習者がいるときといないときで肌色情報が変化したことについて .	39
6.5.2	装飾品の有無	39
第7章	おわりに	41
	謝辞	42
	参考文献	43

目次

3.1	本研究における学習者の存在の検出手順	16
4.1	スナップショットの撮影画面	19
4.2	基準値の取得 (肌色情報の抽出例)	19
4.3	顔検出例	20
4.4	全体の肌色情報からの検出例 (学習者がいるとき)	22
4.5	全体の肌色情報からの検出例 (学習者がいないとき)	23
4.6	肌色部分の動体検出例	23
5.1	被験者 A のシステムの検出結果	26
5.2	被験者 A の実際の存在状況	26
5.3	被験者 B のシステムの検出結果	27
5.4	被験者 B の実際の存在状況	27
5.5	被験者 C のシステムの検出結果	28
5.6	被験者 C の実際の存在状況	28
6.1	各カメラで撮影した画像	38

表目次

3.1	各方法の特徴まとめ	15
5.1	評価実験環境	25
5.2	Logicool 1.3-MP Webcam C300 の仕様	25
6.1	基準とする環境	34
6.2	Logicool 1.3-MP Webcam C200 の仕様	35
6.3	Microsoft LifeCam VX-3000 の仕様	36
6.4	各カメラごとの肌色抽出結果	37
6.5	各照明ごとの肌色抽出結果	37
6.6	眼鏡とマスクを着用した際の肌色抽出結果	37

第 1 章

はじめに

近年，多くの高等教育機関で WBT に代表される e-Learning を利用した遠隔教育が行われている。主に遠隔教育はインターネットを利用した遠隔講義として実施されている。学習者は Web ブラウザからコンテンツにアクセスすることで時間や場所の制約を受けることなく学習することができる。近年では，社会人の受け入れや企業や他大学との連携が促進されることによって，教育機会の拡大が期待される等の理由から政府もインターネットを利用した遠隔教育を推進している。また，遠隔講義は 1999 年から大学設置基準 [1] 等で単位認定が認められており，現在各高等教育機関で様々な取組みが行われている。

単位認定を伴う遠隔講義では，対面での講義と同様に大学設置基準等で定められている学習時間を確保する必要がある。現状の WBT をはじめとする e-Learning システムは，一般的に学習コンテンツと学習全体を管理する LMS (Learning Management System) から構成されている。この LMS に記録される学習時間や成績等の情報から学習状況を把握することができる。インターネットを利用した遠隔講義を行っているサイバー大学では，学習の合間に小テストやレポート，書き込み等のアクションを起こさせることで学習時間の確保を行っている。[2] しかし，これらの情報からでは間接的に，そして一時的にしか学習状況を把握することができない。そのため，学習者がコンピュータの前にはいない場合や，学習を行っていない場合を判断することができないという問題がある。

そこで本研究では，WBT に代表される e-Learning において，学習者がコンピュータの前に存在し，学習を行っているか把握するために，学習者の存在をリアルタイムで検出するシステムを構築する。本研究では，Web カメラを用い，この Web カメラから入力された映像に対して画像処理を施すことで学習者の存在を検出する。

第 2 章

研究の背景と目的

本章では、高等教育機関における e-Learning を利用した遠隔教育の現状と課題を述べたあと、現状の e-Learning システムの問題点とその解決方法を述べる。

2.1 e-Learning を利用した遠隔教育

近年、情報通信技術 (ICT : Information Communication Technology) の著しい発展と普及によって、教育分野にも情報通信技術が利用されている。情報通信技術を利用して学習を行うことを e-Learning といい、大学を始めとする多くの高等教育機関において導入されている。高等教育機関における e-Learning は主に対面での授業の補助教材や自習用の教材として提供されている。また近年では、平成 18 年度に策定された IT 新改革戦略 [3] において、「インターネット等を用いた遠隔教育を行う学部・研究科の割合を 2 倍以上にすることを目指し、大学におけるインターネットを用いた遠隔教育等の推進により、国内外の大学や企業との連携、社会人の受け入れを促進する」とあるように、政府もインターネットを利用した遠隔教育を推進しており、インターネット上で学習を行う WBT (Web Based Training) に代表される e-Learning を利用した遠隔教育に関心が高まっている。

遠隔教育 (Distance Learning) は、遠隔地にいる人に対して、郵便、テレビ、ラジオ、インターネット、衛星等を利用して教育を行うことをいう。従来は主に郵便を利用した遠隔教育が行われていたが、情報通信技術の発達と普及に伴い現在ではテレビ、ラジオ、インターネット、衛星等を利用した遠隔教育が行われている。高等教育機関においては、例えば放送大学のようにテレビやラジオを利用して遠隔講義を実施している機関もあれば、SCS

2.1 e-Learning を利用した遠隔教育

(Space Collaboration System) に代表される衛星通信やインターネットを利用した遠隔講義も実施されている。特に WBT のようにインターネットを利用して行う遠隔講義は、放送大学のようなテレビやラジオを利用したものと比較して時間や場所の制約を受けないインタラクティブな講義を提供することができる。また、遠隔講義は 2001 年 3 月より大学設置基準 [1] 等により単位認定が認められ、現在多くの高等教育機関で遠隔講義に対して様々な取組みが行われており、今後も遠隔講義を行う高等教育機関は増加していくものと予想される。

2.1.1 WBT (Web Based Training)

WBT はイントラネットやエクストラネット、あるいはインターネット上で Web ブラウザを用いて学習を行う e-Learning の手法の一つである。他の e-Learning の手法に教材を CD-ROM や DVD に収録して配布する CBT (Computer Based Training) がある。しかし、CBT は学習教材の作成と更新に多くのコストがかかるという問題点があった。WBT はこの CBT の問題点である学習教材の作成・更新コストを Web を用いることで軽減し、ネットワークを使用することでインタラクティブ性を向上させている。加えて、コンテンツをサーバにアップロードすることで学習コンテンツを多くの学習者に配信することができる。学習者にとっては、Web ブラウザを利用してコンテンツにアクセスするだけで学習ができるため、時間や場所に制約されず、いつでもどこでも学習することができる。

2.1.2 LMS (Learning Management System)

LMS (Learning Management System) は学習者の成績・個人情報や学習の進捗状況など学習管理に必要な情報の記録や学習コンテンツの管理を行うシステムである。また、単純に学習コンテンツをストックするだけの機能のものもあり、それらは CMS (Course Management System) と呼ばれることもある。WBT をはじめとする e-Learning では、LMS を利用した手法が主流となっており、LMS と学習コンテンツで相互に通信を行うこと

2.1 e-Learning を利用した遠隔教育

で、よりインタラクティブ性を高めることができる。

2.1.3 遠隔講義 (同期型, 非同期型)

遠隔講義は、同期型の遠隔講義と非同期型の遠隔講義の2つのタイプに分けることができる。前者の同期型の遠隔講義は、一般的にテレビ会議システム等を利用したリアルタイム型の遠隔講義を指す。ある地点で行われている講義をカメラとマイクにより撮影し、インターネットや衛星を利用して別の教育機関の会場にリアルタイムで配信することで他地点から講義を受講することが可能となる。また、受講側の会場にもカメラとマイクを用意することで、リアルタイムに複数の地点間を接続し、双方向の講義を行う。これにより、遠隔地にいながら講師に質問したり、他会場の受講生と議論することができるなど、実際の対面での講義に近い環境で講義を受講することができる。

これに対し、非同期型の遠隔講義は、一般的にインターネットを利用した VOD (Video On Demand) タイプの映像配信型の遠隔講義を指す。VOD タイプの映像配信型の遠隔講義は映像コンテンツの一時停止や早送り、巻き戻し等の操作が可能なサービスが多く、学習コンテンツを使って繰り返し学習できることや、聞きそびれた箇所をすぐに聞き直すことができるなど、利便性が高いという特徴がある。この VOD タイプの映像配信型の遠隔講義は、主にインターネットを利用して提供されており、学習者は Web ブラウザを用いてコンテンツにアクセスし、学習を行う。そのため、学習者は対面での講義や、リアルタイム型の遠隔講義に比べ、時間や場所の制約をうけることなく学習を行うことができる。

2.1.4 遠隔講義の単位認定状況

高等教育機関における遠隔講義は、大学設置基準 [1] 等に「多様なメディアを高度に利用して、当該授業を行う教室等以外の場所で履修させることができる」とあるように、現行の制度で単位認定が認められている。遠隔講義は、大学設置基準 [1] 等の改正により、1999年3月に、同時性、双方向性を確保した遠隔講義の単位が60単位まで認定されるようになり、

2.2 単位認定を伴う遠隔講義の現状と問題点

また、2000年11月に大学審議会から答申された「グローバル化時代に求められる高等教育の在り方について」を受けて、2001年3月には同時性、双方向性がなくても、対面での講義と同様、あるいはそれ以上の学習効果を確保されると評価された場合には、遠隔講義として単位を認定できるようになった [4].

現行の制度では、通学制の大学では大学設置基準 [1] により卒業要件である 124 単位中 60 単位まで、短期大学設置基準 [5] では卒業要件である 62 単位中 30 単位まで、高等専門学校設置基準 [6] では卒業要件である 167 単位中 30 単位まで e-Learning を含めた「メディアを利用した授業」で単位認定できる。また、通信制の高等教育機関においては、大学通信教育設置基準 [7] 等により、卒業要件であるすべての単位を e-Learning を含めた「メディアを利用した授業」で取得することができるようになっている。

2.2 単位認定を伴う遠隔講義の現状と問題点

単位認定を伴う遠隔講義では、対面での講義と同様に、大学設置基準等で定められている学習時間を確保する必要がある。対面での講義であれば、講師は講義中に出席をとったり、学習中の様子を目視により確認することができるが、WBT のようにインターネットを利用して行う遠隔講義では、受講者がインターネットを介した場所にいるため、対面での講義のように講師が直接学習者の状況を確認することはできない。

現状の WBT をはじめとする e-Learning システムは、学習コンテンツと LMS (Learning Management System) から構成されている。インターネットを利用して行う遠隔講義においても、この LMS に記録される学習開始時間や終了時間、成績等の情報から学習状況や学習の進捗状況を把握することができる。インターネットを利用した遠隔講義を行っているサイバー大学では、学習の合間に定期的に小テストやレポート、書き込み等のアクションを起こさせることで学習時間の確保を行っている [2]。これらの学習者が起こしたアクションはすべて LMS に細かく記録される。これらの情報をチェックすることで学習者の学習状況を詳細に把握することができる。しかし、これらの情報からでは、間接的に、そして一時的な

2.3 本研究の目的

情報からしか学習状況を把握することができない。

現状の e-Learning システムには学習者の学習状況を直接的に把握できる機能は備わっていないため、学習者がコンテンツを再生しているにも関わらず、コンピュータの前にいない状態や、学習を行っていない状態を検出することができない。そのため、学習時間を確保することができないという問題がある。

2.3 本研究の目的

単位認定を伴う遠隔講義では、対面での講義と同様に大学設置基準 [1] 等で定められている学習時間を確保する必要がある。現状では LMS に記録される情報から学習状況を把握することができるが、これらの情報からでは、間接的に、そして一時的にしか学習状況を把握することができず、対面での講義のように直接学習状況を把握することはできない。そのため、学習者がコンテンツを再生しているにも関わらず、コンピュータの前にいない場合や、学習を行っていない場合は把握することができない問題がある。

そこで、本研究では、WBT のようなインターネットを利用して行う e-Learning において、学習者がコンピュータの前に存在し、学習を行っていることを把握するために学習者の存在をリアルタイムで検出するシステムを構築する。

本研究では学習者の存在を検出するために Web カメラを用いる。近年低価格化が進んでいる Web カメラを用いることで、導入コストを低く抑えることができる。また、Web カメラは USB (Universal Serial Bus) によりコンピュータに接続するタイプが一般的となっている。現代ではほぼすべてのコンピュータに USB が装備されているため、Web カメラを用いることで学習を行うコンピュータを選ばずに学習者の存在を検出することが可能となる。さらに、学習者の存在を検出するために使用するだけでなく、学習中の監視にも使える。学習者の学習中の様子を Web カメラにより撮影し LMS に保存することで学習ログの一つとして使うことができる。

これらの理由から Web カメラを用いる。本研究では、この Web カメラから入力された

2.3 本研究の目的

映像に対し画像処理を施すことで学習者の存在の検出を行う。

第 3 章

学習者の存在検出方法の検討

学習者の存在をリアルタイムで検出するために、本研究では Web カメラを用いる。なお、デスクトップ PC の場合は Web カメラを USB により接続し、ディスプレイの上部に固定して使用することを想定している。また、ノート PC の場合は組み込まれているカメラを使用するか、USB により接続し、ディスプレイの上部に固定して使用することを想定している。本研究では、この Web カメラから入力された映像に対し、画像処理を施すことで学習者の存在を検出する。

本章ではまず、システムを設計する上での要求条件を述べ、その後一般的な画像処理技術の特徴を列挙し、WBT に代表される e-Learning においてどの方法を用いればよいか考察する。そして、最後に本研究ではどの方法を用いて学習者の存在の検出を実現したか述べる。

3.1 システムを設計する上での要求条件

現在、高等教育機関における単位認定を伴う遠隔講義は、WBT のように主にインターネットを利用して行っている。そのため、本研究においてもインターネットを利用した学習を想定した上で学習者の存在検出方法の検討を行う。本研究では以下の条件を考慮に入れ学習者の存在検出方法を検討する。

- Web ブラウザ上で処理が行えること
- 学習を行うコンピュータに特殊なアプリケーションを導入する必要がないこと
- リアルタイム処理が可能であること
- 一般の家庭にあるコンピュータでもあまり処理負荷をかけることなく動作すること

3.2 各画像処理手法の特徴

- 学習環境の変化に対応すること

WBT をはじめとする e-Learning では、Web ブラウザを用いてコンテンツへアクセスし、学習を行う方法が一般的となっている。そのため、Web ブラウザ上で学習者の存在を検出する処理を行う必要がある。また、WBT をはじめとする e-Learning では、インターネットカフェや勤務先のコンピュータなど普段自分が学習に使用しているコンピュータ以外のコンピュータを使用して学習を行う場合が考えられる。そのため、インターネットを利用することでいつでもどこでも学習することができるという利点を活かすためにも、学習を行うコンピュータに特殊なアプリケーションを導入する必要なく、学習者の存在を検出できる必要がある。また、学習者の状況は常に変化しているためリアルタイム性を持たせることが必要であり、一般の家庭にあるようなコンピュータや、スペックの低いノート PC でもあまり処理負荷をかけることなく動作する必要がある。さらに、学習者の学習環境は一様ではなく、場所や時間などによって変化することから、学習環境の変化に対応することが求められる。

本システムでは、これらの条件を考慮にいたした上でシステムを設計する。

3.2 各画像処理手法の特徴

本節では、以下の一般的な画像処理手法について述べる。以下では、2 値化、エッジ抽出、肌色抽出、顔検出 (物体検出)、動体検出 (フレーム間差分)、背景差分法の特徴を述べる。

3.2.1 2 値化

2 値化とは、濃度情報や色情報を持つ画像中の各画素を、しきい値を境に白と黒に変換する処理のことをいう。この 2 値化された画像のことを 2 値画像という。グレースケール画像やカラー画像を 2 値画像にすることで画像中の多くの情報を失うことになるが、2 値化は最も単純に前景と背景を分離することができる手法である。また、画像内の物体の輪郭を抽出する際にも有効な手法であり、実用的な画像処理システムで広く用いられている。主な利

3.2 各画像処理手法の特徴

点，欠点としては以下のものがある．

- 利点

- － 明度値等の情報を元に 2 値化を行うことで前景と背景を分離することができる．
- － しきい値を基準に白と黒に変換するだけなので処理が軽く，処理速度が速い
- － 情報量を削減することによりその後の処理速度が速くなる

- 欠点

- － 中間色を表現することができない
- － 画像が変わってしまった場合や，前景にある物体の距離が変わってしまった場合には，その都度しきい値を設定し直さなければならない．

3.2.2 エッジ抽出

エッジ抽出は，物体の輪郭などに見られる，濃度が大きく変化する点を抽出する処理をいう．物体と背景の境目は画素値が急激に変化することが多い．このような箇所にエッジ検出を用いることで，画像内に存在する物体の輪郭を抽出することができる．主な利点，欠点としては以下のものがある．

- 利点

- － 画像内に存在する物体の輪郭を抽出することができる

- 欠点

- － 背景にエッジがない場所であれば，対象とする物体の輪郭を忠実に抽出することができるが，背景にエッジが多ければ，物体の輪郭を忠実に抽出することができない

3.2.3 肌色抽出

肌色抽出は，映像内や，画像内から肌色の部分を抽出する処理のことをいう．肌色情報は，人間が必ず所持している色情報であり，人間の顔や，皮膚，手を検出する際に多く用いられている．主な利点，欠点としては以下のものがある．

3.2 各画像処理手法の特徴

- 利点
 - 誰もが持っている肌色を抽出することで、人間の顔や手を検出することができる
 - 処理が軽く、処理速度が速い
- 欠点
 - 人物以外の肌色情報、例えば木や段ボール等の色が肌色として抽出されてしまうことがある
 - 暗い場所や明るすぎる場所では安定して肌色情報を抽出することができない

3.2.4 顔検出 (物体検出)

顔検出を始めとする物体検出では、パターンマッチングによる手法が提案されている。パターンマッチングでは、あらかじめ学習された顔パターン等の学習パターンと画像や映像を照合する。このとき、学習パターンと同じパターンが出現した場合に物体として検出する。また、顔検出を始めとする (物体検出) は背景や照明環境や明るさの変化に強く、また物体の大きさや色に影響されないという特徴を持つ。主な利点、欠点としては以下のものがある。

- 利点
 - 映像内、画像内に存在する顔を検出することができる
 - 背景や照明環境や明るさの変化に強い
 - 顔の大きさや肌の色に関係なく検出ができる
- 欠点
 - 画像や映像内に顔と似た形状のものがあれば、顔と認識してしまうことがある
 - 2値化や、肌色抽出に比べ、計算コストが多くかかってしまう

3.2.5 動体検出 (フレーム間差分)

動体検出は映像、画像内から、動いている物体を検出する処理のことをいい、フレーム間差分とも呼ばれる。動体検出は、連続で取得した画像から前フレームの画像と現フレームの

3.2 各画像処理手法の特徴

画像の各画素の差分を取得する。その各画素の差分とあらかじめ設定しておいたしきい値を比較し、設定しておいたしきい値より大きくなった場合に、動体として検出する。主な利点、欠点としては以下のものがある。

- 利点
 - － ひとつ前のフレームと現在のフレームを比較することから、環境の変化に強い
 - － 映像内で人間が動いている場合には比較的簡単に検出することができる
- 欠点
 - － 対象としている物体以外の動きも検出してしまう
 - － 処理が軽く、処理速度が速い

3.2.6 背景差分法

背景差分法はカメラが固定されていて視界の移動がない場合に、画像中から注目すべき物体を抽出する定番の方法として用いられている。最も単純な背景差分法は、抽出したいものが入っていない画像を背景画像として保持しておき、現在のフレームの色情報と背景画像における色情報との差分を計算することで実現できる。主な利点、欠点としては以下のものがある。

- 利点
 - － 背景画像がある場合には、比較的簡単に注目すべき物体を抽出することができる
- 欠点
 - － 事前に背景画像を撮影しておく必要がある
 - － 照明の変化や、時間変化に伴う外からの光の変化等の環境の変化に弱い
 - － カメラは固定されている必要がある

3.3 各画像処理手法の検討

各画像処理手法の特徴をふまえ、WBTをはじめとするインターネットを利用して行う e-Learning の学習環境下においてどの手法を用いるのが適当か検討を行った。また、その方法が本当に適用可能か調べるために実際に各画像処理手法を試し確認を行った。

3.3.1 2 値化とエッジ抽出を用いた手法の検討

学習者が存在しているときとしていないときで、映像中の何らかの情報が増減することから、まず 2 値化とエッジ検出を試みた。これらは事前に確認を行った結果、2 値化では学習者がいるときといないときとで白と黒の合計ピクセル数から情報の増減が確認することができた。また、エッジ検出では学習者がいるときといないときとでエッジの量から情報の増減を確認することができた。

しかし、これらの方法は上記の 3.2.1 と 3.2.2 でも述べた通り、学習を行う環境の明るさや背景の情報、学習者とカメラの間の距離の変動に大きく左右される。確認を行った結果からも、2 値化であれば学習環境に応じてしきい値を定める必要があることがわかった。また、学習者とカメラの間の距離が変化したときにも、学習者がいるときといないときと同様の変化を抽出することがあり、安定して検出を行うことが困難であった。またエッジ検出では、エッジをとることで学習者の輪郭を抽出することができるが、本棚などの前で学習を行う場合には、本棚にあるすべてのエッジを検出してしまう。また、学習者とカメラの間の距離の変化にも大きく影響を受ける。これらの結果から安定して検出を行うことが困難であった。

これらの理由から 2 値化とエッジ検出から学習者の存在を検出することは困難である。

3.3.2 背景差分法を用いた手法の検討

背景差分法により学習者の存在を検出することを試みた。本研究ではカメラは固定されていると想定しており、視界の移動もないので、背景差分法は有効な手段といえる。しかし、背景差分法は 3.2.6 でも述べたように、事前に背景画像を取得しておく必要がある。WBT

3.3 各画像処理手法の検討

のようなインターネットを利用した e-Learning では、普段学習を行う場所以外（インターネットカフェや、勤務先）での学習が考えられることや、学習を行う場所の明るさが逐次変化する場合が考えられるため、事前に背景画像を取得するだけではその変化に対応することができない。

また、学習開始時に背景画像を撮影することで学習環境の変化に対応できる。しかし、背景画像を撮影するために学習者にコンピュータの前から離れてもらうことは、学習を開始しようとしている学習者の学習意欲を低下させてしまう可能性がある。これらのことから背景差分法により学習者の存在を検出することは困難である。

3.3.3 動体検出を用いた手法の検討

学習者は学習中にノートを取る、参考書を見るなどの行動をとることが考えられる。動体検出は映像内で動きがあったものを検出するため、これら学習者の動きを検出することで学習者の存在を検出することができると考えられる。学習者の動体だけを検出できれば、学習者の存在を検出することが可能である。しかし、映像内に学習者が存在していて、さらにその背景に風に揺れるカーテンやテレビの映像があった場合、これらも動体として検出されてしまうため、学習者の動体だけを検出することは困難である。

また、学習者が学習を行っている環境に学習者以外の人がいる可能性がある。そのため、動体検出は学習者ひとりしか存在していない環境においては有効であるが、学習者以外の人間が存在している環境においては有効な手段ではないといえる。

3.3.4 肌色情報、顔検出を用いた手法の検討

肌色情報は 3.2.3 でも述べたように、人間誰しもが持っている色情報である。事前に確認を行った結果、学習者がいるときといないときで、画像内、映像内の肌色情報の増減から変化を抽出することができた。また、学習者とカメラの間の距離や照明等の明るさが変化した場合においても多少抽出結果に増減は見られるが、他の手法に比べて安定して変化を抽出す

3.4 本研究での実現方法

表 3.1 各方法の特徴まとめ

	肌色抽出	2 値化	エッジ抽出	顔検出	動体検出	背景差分法
明るさの変化	△	×	○	○	○	×
カメラとの距離の変化	○	×	×	○	○	○

ることができた。

顔検出は上記の 3.2.4 で述べたように、明るさや、背景の情報、学習者とカメラの間の距離に影響されずに、検出を行うことができる。また、顔と似た形状のものも検出してしまうことが考えられるが、検出された顔の位置に肌色情報が含まれていることを調べることで、これを顔と判断することが可能であり、誤検出を防止することができる。

これらの理由から、肌色抽出を用いた手法と顔検出と肌色抽出を組み合わせた手法は有効な手段である。

3.4 本研究での実現方法

上記で検討を行った各手法の特徴を以下の表 3.1 にまとめる。肌色抽出は、明るさが変化した場合やカメラと学習者の間の距離が変化した場合においても抽出結果に多少の増減は見られたが、安定して抽出を行うことができた。また、顔検出と動体検出も明るさの変化やカメラと学習者の間の距離の変化に強く、有効な手段であるといえる。これらの結果から肌色抽出および顔検出と動体検出は有効な手段である。

また、顔の検出範囲に肌色抽出を施すことで、検出された顔が人間の顔か判断することができる。さらに本研究では、肌色抽出と動体検出を組み合わせる。これらの 2 つの手法を組み合わせることで映像内に動いている人間を検出する。

本研究では顔検出と動体検出と肌色抽出により学習者の存在の検出を行う。

3.5 本研究における学習者の検出手順

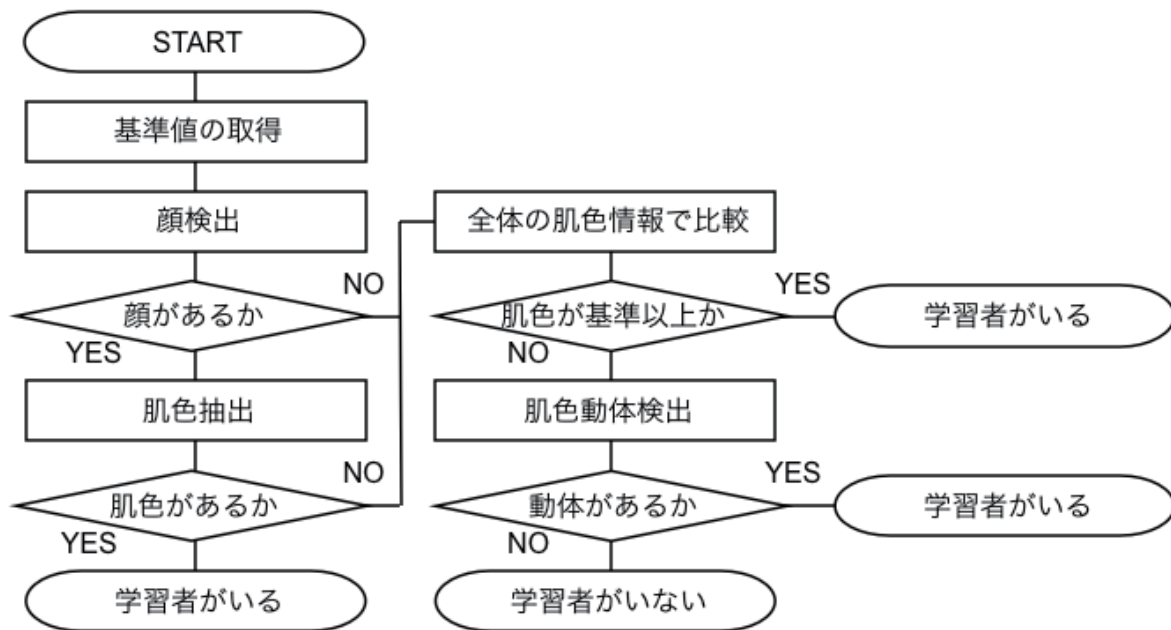


図 3.1 本研究における学習者の存在の検出手順

本研究では、肌色抽出と顔検出と動体検出を用いて学習者の存在を検出する。本研究での学習者の存在の検出手順を図 3.1 に示す。

まず、検出を行う前に顔領域内の肌色情報の合計ピクセル数と全体の肌色情報の合計ピクセル数、および背景の合計ピクセル数を取得する。ただし、学習を行う場所の明るさや照明環境によって抽出できる肌色情報に差が生じる可能性がある。これは、肌色情報を元に検出を行っているため、肌色情報の抽出結果に差がでてしまうと正確に学習者の存在を検出できない可能性がある。そこで、学習環境に応じた肌色情報の合計ピクセル数を学習開始時に取得することで、学習環境の変化に対応する。これらの数値を基準値とすることで学習環境の変化に対応する。

次に顔検出を行う。顔検出は照明等の明るさの変化、カメラと学習者の間の距離の変化に強いという特徴を持つため、安定して検出ができる。そこで、本研究では安定して検出を行うことができる顔検出を最初に行う。また、顔検出を行った際に、顔と認識された領域内に

3.5 本研究における学習者の検出手順

肌色情報が存在しているか調べることで誤検出を防止する。顔検出に関しては、多くの学習パターンから学習者の顔を厳密に検出することも可能であるが、そうすると計算量が大きくなるため、本研究では、学習者の正面を向いた顔だけを検出するようにしている。

顔検出ができなかった場合は全体の肌色情報により検出を行う。本システムでは学習者の正面を向いた顔のみ検出するようにしているので、学習者が下を向いたり横を向いたりした際には検出することができない。全体の肌色情報と学習開始時に取得した肌色情報を比較することでこれらの学習者の状態を検出する。

最後に、肌色部分の動体を検出する。全体の肌色情報から学習者の存在を検出できなかった場合でも学習者がまだコンピュータの前に存在している可能性がある。背景に学習者以外の人が存在していた場合、誤検出が発生する可能性があるが、本研究では学習者の存在をより厳密に検出するために、肌色部分の動体検出を最後に施す。

これら学習者の存在の検出方法に関しては次の4章で詳しく述べる。

第 4 章

学習者の存在を検出するシステムの構築

本研究では、顔検出、全体の肌色情報による検出、肌色動体検出を用いて学習者の存在をリアルタイムで検出するシステムを構築した。本章ではまずシステムの概要を述べる。その後、図 3.1 に示す検出の処理の流れに沿って、基準値の取得、顔検出、肌色情報による検出、肌色動体検出について述べる。

4.1 システムの概要

本研究では、WBT のようなインターネットを利用して行う e-Learning を想定し、Web ブラウザ上で動作するよう ActionScript3.0 により構築した。また、Web サーバとして Apache を使用し、学習者を検出するために Web カメラを端末に接続して使用した。そして Web カメラから入力された映像に対して顔検出、肌色情報による検出、肌色動体検出を施すことで学習者の存在の検出を行った。また、クライアントに対し低負荷で動作するように、フレームレートを 1 FPS (Frames Per Second) に設定した。

4.2 基準値の取得

学習を行う場所の明るさや証明環境によって抽出できる肌色情報に差が生じる。本システムでは、肌色情報を元に検出を行っているため、肌色情報の抽出結果に差がでてしまうと正確に学習者の存在を検出できない可能性がある。そこで本システムでは、まず学習開始時

4.2 基準値の取得

に学習環境に応じた肌色情報の合計ピクセル数を取得することで、学習環境の変化に対応する。

学習者に学習開始時に図 4.1 のように指定された枠内に顔を入れて、スナップショットを一枚撮影してもらおう。このとき、撮影された画像に対して顔検出を施し、顔領域内の肌色情



図 4.1 スナップショットの撮影画面

報の合計ピクセル数を取得する。その後、画像全体の肌色情報の合計ピクセル数を取得する。また、これら 2 つの値に加えて背景の肌色情報の合計ピクセル数を取得する。背景の肌色情報の合計ピクセル数は、全体の肌色情報の合計ピクセル数から顔領域内の肌色情報の合計ピクセル数を差し引くことで取得する。これらの値を基準値として学習開始時に取得しておく。顔領域内の肌色情報、画像全体の肌色情報の抽出例、背景の肌色情報の抽出例を図 4.2 に示す。



図 4.2 基準値の取得 (肌色情報の抽出例)

肌色情報の抽出方法に関しては、以降の 4.4.1 で詳しく述べる。

4.3 顔検出

4.3 顔検出

基準値を取得した後に、学習者の顔を検出する。顔検出は学習環境の変化や学習者の顔とカメラの間の距離の変化の影響を受けにくい。そのため、本研究では安定した検出結果を得ることができる顔検出を最初に行い学習者の存在を検出する。

顔検出はパターンマッチングにより行う。パターンマッチングとは、画像内に特定のパターンが出現するかどうかを特定する手法のことをいう。本システムでは、パターンマッチングの手法のひとつである Haar-like 特徴に基づいた検出を行った。この Haar-like 特徴を用いた物体を検出するアルゴリズムは、学習パターンとなる分類器をあらかじめ作成しておき、この分類機によって物体を検出する。物体を検出するために用いる特徴が、Haar-like 特徴である。この Haar-like 特徴は、矩形領域の平均輝度差であり、隣接した矩形領域で輝度差を求めることで、物体の検出を可能としている。

本研究では、顔検出に画像認識ライブラリである OpenCV の ObjectDetection (物体検出) を ActionScript3.0 に移植したオープンソースの Marilena を使用し、顔の学習パターンとなる分類器には、OpenCV に備わっている“haarcascade_frontalface_alt.xml”を使用した。実際に顔を検出した例を図 4.3 に示す。赤線で囲われた領域が顔と判別された部分である。

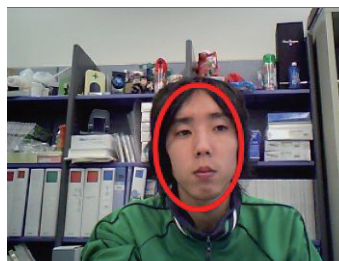


図 4.3 顔検出例

本システムでは、学習者の顔が検出できた場合に、“学習者がいる”と判定している。また、顔が検出された際に、基準として取得しておいた顔領域内の肌色情報と同じ割合の肌色情報があるか調べることで誤検出を防止している。顔を検出できなかった場合には、映像中

4.4 肌色情報による検出

の肌色情報から検出を行う。

4.4 肌色情報による検出

顔が検出できなかった場合、また顔は検出されたが基準値として取得していた顔領域内の肌色情報と同程度の肌色情報を取得できなかった場合には、人間誰しもが持っている色情報である肌色情報から検出を行う。

4.4.1 肌色情報の抽出方法

コンピュータは RGB により色を表現している。RGB は光の三原色である、R (赤 : Red), G (緑 : Green), B (青 : Blue) を元にした表色系である。この RGB はビットマップ形式をはじめとする多くの画像形式で用いられている。

しかし、この RGB 表色系で表現された色は、コンピュータが処理しやすいように作られた表色系であるので、人間にはわかりにくいという欠点がある。また、RGB 表色系では空間の光度が変化した場合に R と G と B の値それぞれが変化してしまうため、環境に応じて正確に肌色情報を抽出することができないという問題がある。そこで本研究では、肌色情報の抽出は HSV 表色系により行う。

HSV 表色系とは、色相である H (Hue : 色合い) と、彩度である S (Saturation : 鮮やかさ)、明度である V (Value : 明るさ) の 3 つの数値で表す表色系である。RGB 表色系からこの HSV 表色系へ変換することで、色合いである色相情報を分離し画素ごとの色相を区別することができる。そのため、RGB と違い H (色相) の値のみを変化させることで色を表現できるため、肌色情報抽出を行う際のしきい値を設定しやすい。また、HSV 表色系は実在する空間の光度が変化したとしても色相の値は変化しにくいという特徴を持つ。これらから RGB により肌色情報を抽出するより、HSV のほうが安定して抽出できる。そして、人間の肌色は、この HSV 表色系の H (色相) の値が 0 から 30 の範囲に収まるということが証明されている [8]。そこで、本研究においてもこの HSV 表色系により肌色情報の抽出を行う。

4.4 肌色情報による検出

また、3.3.4 で事前に確認を行った結果、蛍光灯の光や白色に近い壁の色も肌色として抽出してしまうことがあることがわかった。一般的に、人間の肌色は比較的彩度、明度ともに高い色が用いられているのに対し、室内の蛍光灯や壁の色等は、彩度、明度ともに低い色が多く用いられている。そこで、本研究では、彩度と明度のしきい値を 0.1 以下とすることで、これらの人物以外の肌色抽出の誤検出を減らしている。

4.4.2 画像全体の肌色情報からの検出

今回顔の分類器として使用した“haarcascade_frontalface_alt.xml”は学習者の正面を向いた顔しか検出することができない。そのため、学習者が横や下を向いた場合、顔を傾げた場合や口元を手で隠してしまった場合には学習者の存在を検出することができない。そこで、顔を検出できなかった場合にも学習者の存在を検出するために、画像全体の肌色情報から検出を行う。

全体の肌色情報からの検出例（学習者がいるとき、いないとき）を図 4.4 と、図 4.5 に示す。画像全体の肌色情報からの検出では、まず、現フレームの画像全体の肌色情報を取得する。次にその肌色情報から学習開始時に記録した顔領域内の肌色情報の半分の値を差し引く。この半分の値というのは、学習者の顔の向きと、カメラからの距離、学習者の背後の肌色情報を考慮して値を設定している。これらの現フレームの肌色情報と基準値として取得しておいた顔領域内の肌色情報の差分が、学習開始時に取得した背景の肌色情報より多かった場合に、“学習者がいる”と判定している。このとき、図 4.4 では、“学習者がいる”と判定され、図 4.5 では、“学習者がいない”と判定される。



図 4.4 全体の肌色情報からの検出例（学習者がいるとき）

4.4 肌色情報による検出



図 4.5 全体の肌色情報からの検出例 (学習者がいないとき)

4.4.3 肌色部分の動体検出

顔検出、画像全体の肌色情報より学習者の検出が行えなかったときに、さらに厳密に学習者の存在を検出するために、肌色部分の動体検出を施す。画像全体の肌色情報による検出は、学習者がディスプレイから遠く離れてしまった場合には検出されない可能性がある。そこでこのような状態も厳密に検出を行うために、肌色部分の動体検出を実施する。肌色部分の動体検出の例を図 4.6 に示す。

動体検出は映像中から動いている物体を検出する処理のことをいい、フレーム間差分とも呼ばれる。動体検出は、まず連続で取得した画像から前フレームの画像と現フレームの画像の各画素の差分を取得する。その各画素の差分とあらかじめ設定しておいたしきい値を比較し、設定しておいたしきい値より大きくなった場合に、動体として検出する。

本システムではこの動体検出にさらに肌色抽出を組み合わせることで画像内で動く肌色、すなわち学習者の動きを検出することで学習者の存在を検出する。また、学習者は常に動いているとは限らないため、本システムでは、3 秒間という時間を設定し、この時間内で動体が検出された場合に“学習者がいる”と判定し、3 秒以上動体が検出されない場合に“学習者がいない”と判定する。

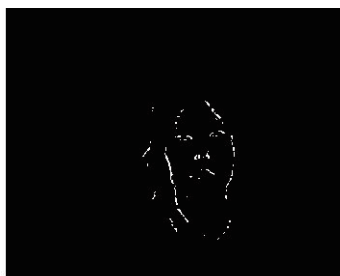


図 4.6 肌色部分の動体検出例

第 5 章

評価と考察

本研究で構築したシステムで学習者の存在を正しく検出できているか確認するために評価実験を行った。本章では、評価実験の概要を述べたあと評価実験の結果を述べ、その後結果に対して考察を行う。

5.1 評価の概要

学習者の存在を正しく検出できているか確認するために本学の学生 3 名 (A, B, C) に対して評価実験を行った。評価実験を行った環境を表 5.1 に示す。学習者には、4.2 で述べたように指定された領域に顔を入れスナップショットを撮影してもらったあと本研究で構築したシステムをバックグラウンドで動作させる。同時にフォアグラウンドで本学で行われた人工知能基礎の講義ビデオを再生し、学習を行ってもらう。そして、学習終了時に、ビデオカメラにより撮影しておいた学習中の映像と本システムの検出結果を比較し、学習者の存在を正しく検出できているか確認した。

Web カメラは、ロジクールの Logicool 1.3-MP Webcam C300 を使用した。ロジクールの Logicool 1.3-MP Webcam C300 の仕様を以下の表 5.2 に示す。

5.2 評価実験結果

評価実験を行った結果、被験者 3 人 (A, B, C) のうち、A, B に関しては正しく学習者の存在を検出することができた。被験者 A のシステムの検出結果と実際の存在状況を図 5.1、図 5.2 に、被験者 B のシステムの検出結果と実際の存在状況を図 5.3、図 5.4 に示す。

5.2 評価実験結果

表 5.1 評価実験環境

場所	室内 (本学の研究室)
時間帯	夜 (明るさを一定にするため)
照明環境	蛍光灯の光 (明るい)
背景	一定 (背景に人が通らない場所)

表 5.2 Logicool 1.3-MP Webcam C300 の仕様

フォーカス	2cm - ∞ (マニュアルフォーカス)
画角 (映る広さ)	60°
画像センサ	130 万画素
フレームレート	最大 30 フレーム / 秒
露出	自動
ホワイトバランス	自動
対応 OS	WindowsXP(SP2 以上), WindowsVista, Windows7
CPU	クロックスピード 1GHz 以上の CPU(1.6GHz 以上推奨)
メモリ	512MB 以上
接続インターフェース	USB

これらの図からわかるように、2人の被験者は学習者がいるときといないときで正しく存在を検出することができた。

しかし被験者 C に関しては概ね正しく存在を検出することができたが、3回の誤検出があった。被験者 C のシステムの検出結果と実際の存在状況を図 5.5、図 5.6 に示す。図 5.5 のシステムの検出結果と図 5.6 の実際の学習者の存在状況と比較した結果、学習を開始してから 96 秒後と、426 秒後と、555 秒後に誤検出が起きたことを確認した。

5.2 評価実験結果

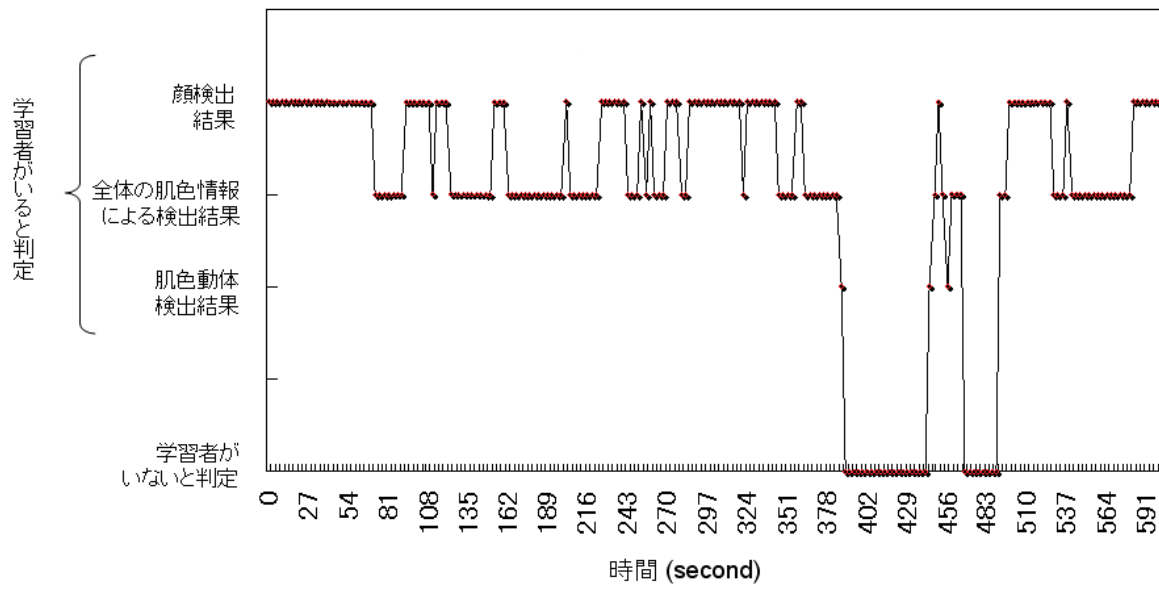


図 5.1 被験者 A のシステムの検出結果

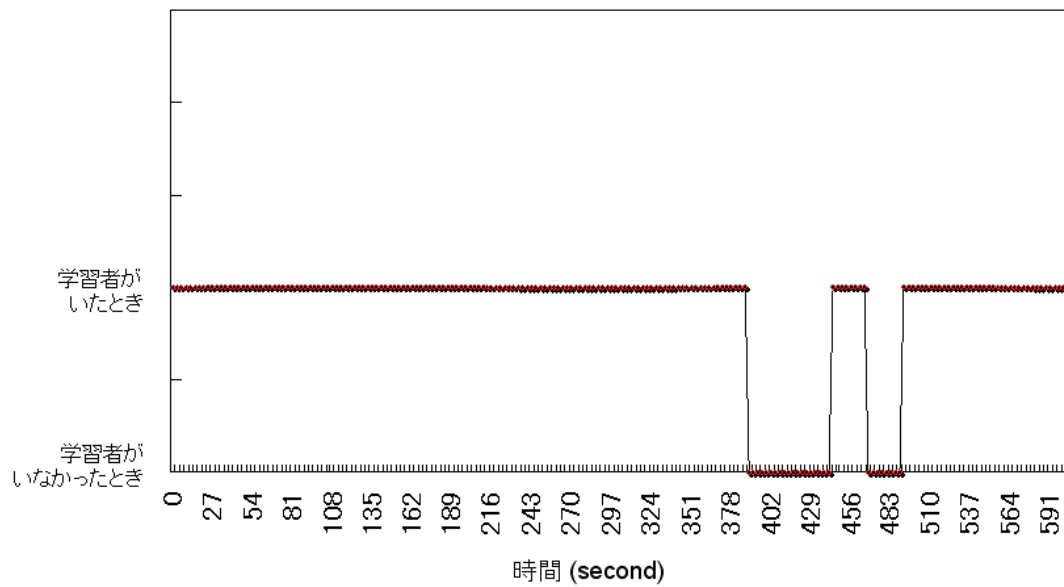


図 5.2 被験者 A の実際の存在状況

5.2 評価実験結果

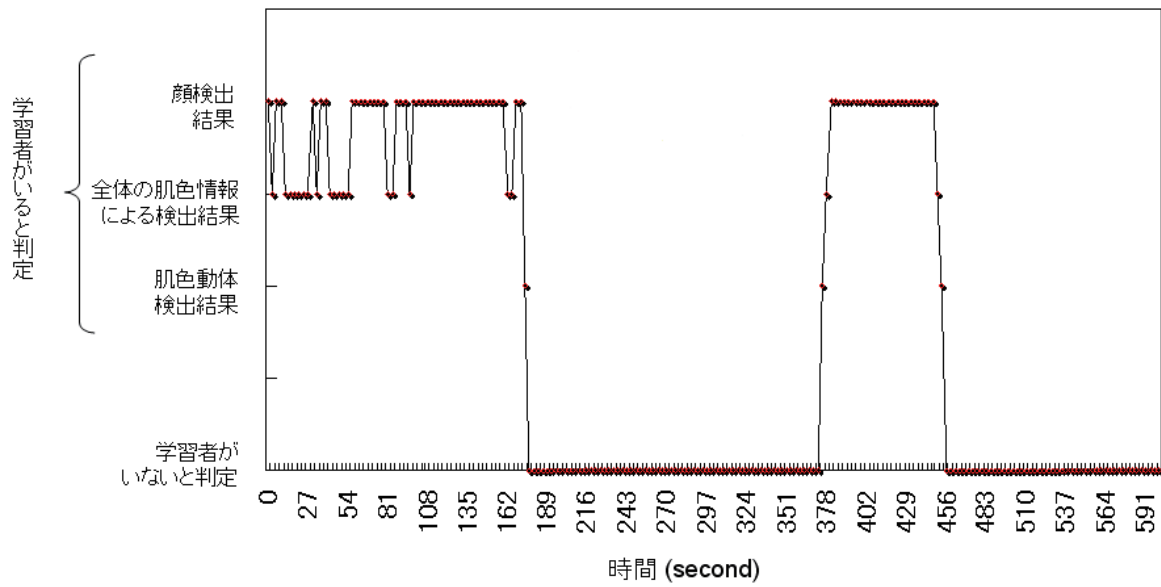


図 5.3 被験者 B のシステムの検出結果

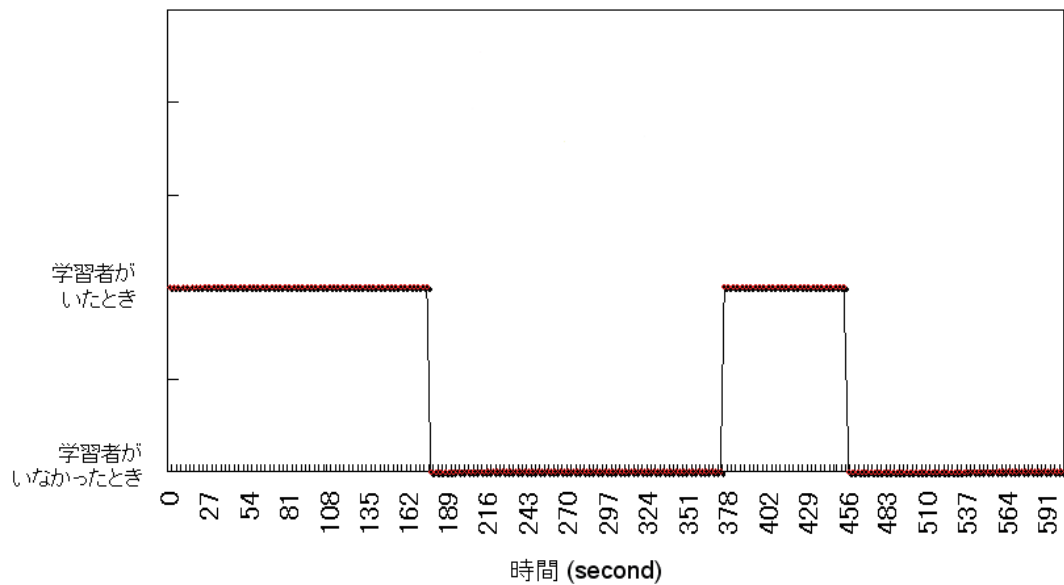


図 5.4 被験者 B の実際の存在状況

5.2 評価実験結果

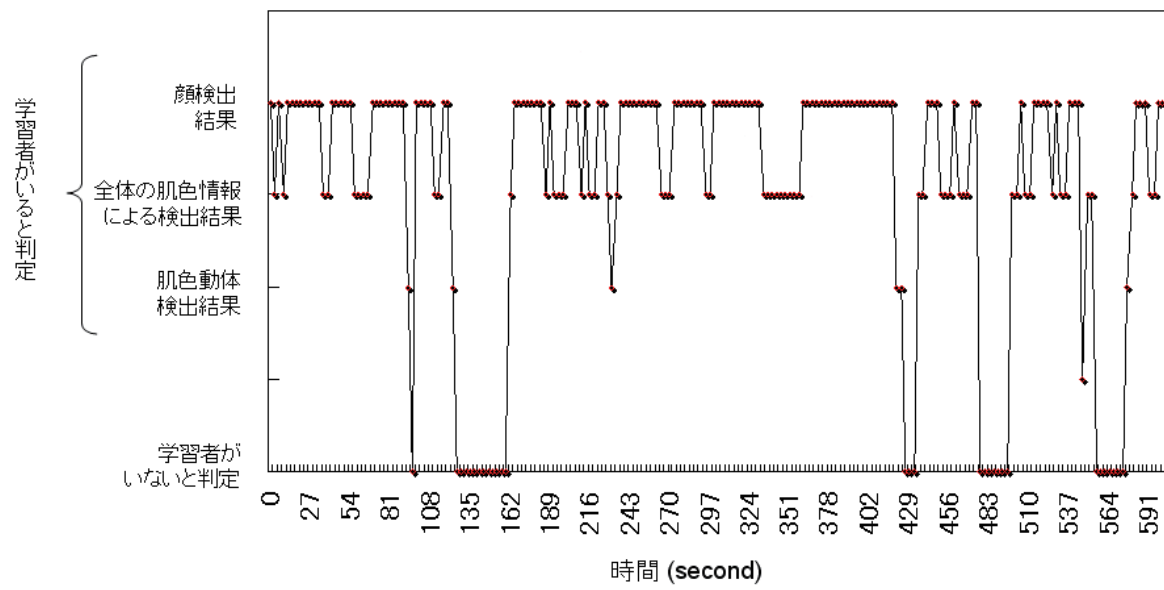


図 5.5 被験者 C のシステムの検出結果

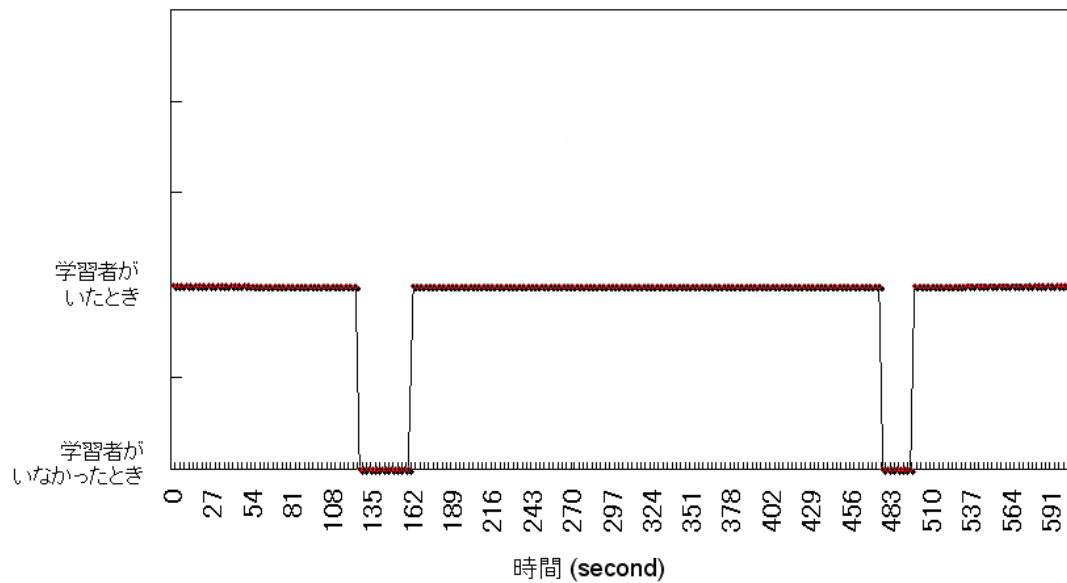


図 5.6 被験者 C の実際の存在状況

5.3 評価実験の考察

評価実験の結果に対して考察を行った。本節では、被験者 A と B に対して行った評価実験ではなぜ正しく存在を検出することができたのか、理由を明らかにする。そして、被験者 C に対して行った評価実験ではなぜ誤検出が発生してしまったのか、A と B と何が違ったのか原因を明らかにし、その原因から今後どのような対策が必要となるか考察する。

5.3.1 被験者 A と B の評価結果に対する考察

被験者 A と B は評価実験の結果から、正しく存在を検出できていることを確認した。正しく存在を検出することができた理由として、これら 2 人の被験者が評価実験中に大きく姿勢を変えることがなかったことが挙げられる。そのため、顔検出を安定な距離で行うことができ、さらに全体の肌色情報による検出も学習開始時に基準として取得していた数値から大きく変わることがなかったため、誤検出なく正しく存在を検出することができたと考えられる。評価結果からも確認できるように、ほぼ顔検出と全体の肌色情報による検出だけで学習者を検出することができている。このことから本システムでは、学習者が大きく姿勢を変えない場合、すなわち、学習者とカメラの間の距離が大きく変動しない場合は正しく存在を検出できたと言える。

被験者 A に関しては 456 秒に顔検出と全体の肌色情報からの検出が行えず、肌色動体検出により学習者の存在を検出している。これは被験者 A がコンピュータの前に戻ってきて再度またコンピュータの前を離れようと立ち上がった際に一定量の肌色情報を抽出することができなかったためだと考えられる。その後被験者 A が立ち上がったからコンピュータの前を離れるまでの間に全体の肌色情報により学習者を検出しているが、これは被験者 A の手の肌色情報を抽出したことにより、存在を検出したという結果になってしまったと考えられる。

5.3 評価実験の考察

5.3.2 被験者 C の評価結果に対する考察

被験者 C は評価実験の結果から、概ね正しく存在を検出することができていたが、大きく姿勢を変えた場合に誤検出が発生した。誤検出は合計で 3 回発生した。以下では、誤検出が発生した際に学習者がとった行動をビデオカメラにより確認し原因を明らかにしたあと、どのような対策が必要となるか考察する。

カメラに学習者の顔が入っていなかったとき

ビデオカメラで録画した映像で確認したところ、96 秒後に発生した誤検出は学習者の顔がカメラに入っていないため発生したことがわかった。検出の対象となる顔と肌色情報が映像内に存在していなかったため、学習者が存在しているにも関わらず“学習者がいない”と判定された結果となった。この結果から、映像から得た検出結果だけで学習者の存在状況を判断せず、検出できない時間を利用する方法が考えられる。

学習者の顔がカメラに入っていなかった時間は 3 秒程度であったが、学習者はその後カメラ内にすぐもどってきた。ビデオカメラで撮影した映像を確認しても、学習者はコンピュータの前から離れようとしたわけではなく、姿勢を変えただけであった。このことから学習者の存在を検出できないからとすぐ“学習者がいない”と判定するのではなく、5 秒間学習者の存在を検出できない場合に“学習者がいない”と判定するなどの対策が必要である。

学習者が大きくのけぞったとき

ビデオカメラの映像を確認した結果、学習開始から 426 秒後に発生した誤検出は学習者が大きくのけぞったため発生したことがわかった。顔検出に関しては、学習者がのけぞったことによって検出を行える顔の大きさより学習者の顔が小さくなってしまったため検出することができなかった。また、全体の肌色情報からの検出では、学習者がのけぞったことによって、検出の対象となる顔領域内の肌色情報が基準値として取得していた数値よりも小さく

5.3 評価実験の考察

なってしまったため検出することができなかった。また、肌色部分の動体検出は、学習者がのけぞった際に動かなかったため検出することができなかった。

これらについて以下の対策が必要である。本システムでは、カメラから 90 cm 以内であれば安定して顔を検出できることを確認した。今回のように学習者の顔が遠ざかってしまった場合にも対応できるように、さらに探索を行う顔の範囲を小さくすることが考えられる。しかし、探索を行う顔の範囲を小さくしてまうとそれだけ探索を行う回数が増えてしまうため、計算コストが大きくなってしまうという問題がある。また、全体の肌色情報から検出を行う場合においても、現フレームの肌色情報から顔領域内の肌色情報の半分の値を差し引くとしているが、この値をさらに半分より小さくすることでより厳密に学習者の存在を検出することができる。しかし、学習者の後ろには肌色情報が存在している可能性があるため、半分より値を小さくすることで誤検出が発生してしまう可能性がある。これらから、学習者が大きくのけぞってしまい、対象とする顔と肌色情報が小さくなってしまった場合には、顔や肌色情報だけでなく、着用している服の情報も検出の対象にするなどの対策が必要である。

また肌色動体検出に関しても、現在 3 秒間の間に肌色部分の動体が検出できれば学習者がいると判定しているが、このとき学習者は 3 秒以上動かなかったため、動体として検出することができなかった。そのため、動体として検出する時間を 3 秒間よりさらに長い時間を設定する必要がある。

学習者が大きくカメラに近づいたとき

ビデオカメラの映像で確認した結果、学習開始から 555 秒後に発生した誤検出は、学習者がカメラに大きく近づいたため発生していたことがわかった。学習者がカメラに大きく近づいたことで映像が暗くなり、設定していたしきい値では肌色情報を抽出することができなかった。本システムの顔検出と全体の肌色情報による検出と肌色動体検出はすべて肌色情報を検出の対象としているため、肌色情報を取得することができない場合には検出することができなくなるため対策が必要となる。

5.3 評価実験の考察

本評価で使用したカメラでは、カメラと学習者の顔の間の距離が約 20 cm 以上であれば、安定して肌色情報を抽出することができたが、学習者の顔が約 20 cm 以内になったときに、抽出できる肌色情報が激減してしまった。このことから学習者の顔とカメラの間の距離が短くなってしまった場合や映像が暗くなってしまった場合に、映像の明るさに応じて肌色のしきい値を変化させるなどの対策が必要がある。

第 6 章

比較検証と考察

評価実験の結果から本システムにより学習者の存在を概ね正しく検出することが可能であることがわかった。しかし本研究で構築したシステムは、検出を行う環境、例えば照明環境や背景の状況、使用するカメラ等によって検出結果に違いがでる可能性がある。そこで比較実験を行い、検出を行う環境が変わることによって検出結果にどのような違いが出るか、また顔検出と肌色抽出それぞれにどのような違いがでるか調べた。本章では、比較検証実験の概要を述べたあと、比較検証実験の結果を述べ、その後結果に対して考察を行う。

6.1 比較検証の概要

カメラの種類の違い、照明環境の違い、装飾品の有無から比較検証を行った。これらの条件を変えることによって顔検出と肌色抽出にどのような違いがでるか調べた。本比較検証実験では、はじめに基準となる環境を設定する。本比較検証実験で設定した基準となる環境を表 6.1 に示す。この基準とする環境における顔検出の結果と肌色抽出の結果と比較することで、どのような違いがでるか調べた。以下では比較検証を行った結果を述べる。なお、比較検証を行った際の画像サイズは 160×120 pixel とし、肌色情報は抽出できたピクセルの合計値として取得した。

6.2 カメラの種類による検証

使用するカメラを変えることにより、どのような違いがでるか、比較検証を行った。使用したカメラは、Logicool 1.3-MP Webcam C300 と、ロジクールの Logicool 1.3-MP

6.2 カメラの種類による検証

表 6.1 基準とする環境

カメラ	Logicool 1.3-MP Webcam C300
場所	室内 (本学の研究室)
時間帯	夜 (明るさを一定にするため)
照明環境	蛍光灯の光 (明るい)
背景	一定 (背景に人が通らない場所)
装飾品	なし
服	肌色と認識されない色のもの

Webcam C200 と、マイクロソフトの Microsoft LifeCam VX-3000 の 3 種類である。これらのカメラの仕様を、表 5.2 と表 6.2 と表 6.3 に記載する。

検証を行った結果、使用するカメラにより抽出できる肌色情報の合計ピクセル数が大きく違うことと、学習者がいなくなったときに正しく学習者の存在を検出できないカメラがあることがわかった。まず、本検証に使用した各カメラごとの肌色情報の抽出結果を表 6.4 に示す。これら比較検証に使用したカメラすべてで肌色情報の抽出結果に大きな違いがでた。とくに Logicool C300 と Microsoft LifeCam VX-300 では肌色抽出の結果に約 1000px もの違いがあった。また、学習者がいなくなったときの肌色情報の抽出結果を比較した結果、Logicool 1.3-MP Webcam C300 と Microsoft LifeCam VX-300 では、肌色情報の抽出結果には差が出たものの学習者の存在を正しく検出することができていた。しかし、Logicool 1.3-MP WebCam C200 では、学習者がいなくなったときに 2088px もの肌色情報を抽出し、学習者の存在も正しく検出することができなかった。

また、顔検出はカメラを変わっても 90cm 以内であればどのカメラでも安定して検出できることを確認した。

6.3 照明環境の変化による検証

表 6.2 Logicool 1.3-MP Webcam C200 の仕様

フォーカス	2cm - ∞ (マニュアルフォーカス)
画角 (映る広さ)	60°
画像センサ	30 万画素
フレームレート	最大 30 フレーム / 秒
露出	自動
ホワイトバランス	自動
対応 OS	WindowsXP(SP2 以上), WindowsVista, Windows7
CPU	クロックスピード 1GHz 以上の CPU(1.6GHz 以上推奨)
メモリ	512MB 以上
接続インターフェース	USB
サイズ (幅 × 奥行 × 高さ)mm	48 × 114 × 73 (机上設置時)

6.3 照明環境の変化による検証

照明環境の違いにより検証を行った。比較した照明環境は、基準と同様の明るい (室内の蛍光灯をすべてつけた状態) 環境, 薄暗い (半分つけた状態) 環境, 暗い (一箇所だけつけた状態) 環境の 3 パターンである。検証した結果, 顔検出, 全体の肌色情報による検出は問題なく行うことができたが, 明るさに応じて抽出できる肌色情報に差が生じた。各照明環境における肌色情報の抽出結果を表 6.5 に示す。

比較検証に使用したカメラは Logicool 1.3-MP Webcam C300 である。このカメラは表 6.2 の結果から, 他のカメラに比べ抽出できる肌色情報の合計ピクセル数が少ないという特徴があった。しかし, 今回比較を行った照明環境のうち, 基準とする明るい (蛍光灯全て点灯) 環境においては全体の肌色情報を 1808px 抽出することができたが, 薄暗い (蛍光灯を半分点灯) 環境では明るい環境における抽出結果よりも多い 2473px という結果になった。

6.4 装飾品の有無による検証

表 6.3 Microsoft LifeCam VX-3000 の仕様

フォーカス	2cm - ∞ (マニュアルフォーカス)
画角 (映る広さ)	55°
画像センサ	30 万画素
フレームレート	最大 30 フレーム / 秒
露出	自動
ホワイトバランス	自動
対応 OS	WindowsXP SP2 日本語版以上, WindowsVista, Windows7
CPU	Windows Vista : Pentium4 2.8GHz (Pentium Dual Core 1.8 GHz 以上推奨) Windows XP SP2 以上 : Pentium4 1.8GHz (Pentium4 3GHz 以上推奨)
メモリ	WindowsVista : 1G 以上の RAM (2GB 推奨) Windows XP SP2 以上 : 256MB 以上の RAM (500MB 推奨)
接続インターフェース	USB
寸法 (幅 × 奥行 × 高さ)mm	57 × 57 × 68

また暗い環境 (蛍光灯をひとつ点灯) においては全体の肌色情報を 535px しか抽出することができなかった。

6.4 装飾品の有無による検証

装飾品の有無により検証を行った。検証は眼鏡とマスクを装着しているときとしていないときでどのような差が生じるか確認を行った。検証した結果、眼鏡を装着した際には顔検出と全体の肌色情報による検出は問題なく行うことができたが、マスクを着用した場合は顔検

6.4 装飾品の有無による検証

表 6.4 各カメラごとの肌色抽出結果

	顔領域	全体	背景	全体 (学習者がいない)
基準となる環境 (Logicool C300)	1454px	1808px	354px	390px
Logicool C200	2041px	2844px	803px	2088px
Microsoft LifeCam VX-300	2390px	2707px	317px	693px

表 6.5 各照明ごとの肌色抽出結果

	顔領域	全体	背景	全体 (学習者がいない)
明るい (蛍光灯全て点灯)	1454px	1808px	354px	390px
薄暗い (蛍光灯半分点灯)	2129px	2473px	344px	289px
暗い (蛍光灯ひとつ点灯)	328px	535px	207px	296px

出ができなかったが、全体の肌色情報による検出はできたという結果になった。眼鏡を装着した際の肌色情報の抽出結果とマスクを装着した際の肌色情報の抽出結果を表 6.6 に示す。

眼鏡を装着した場合は、ほぼ基準として設定したときと同じ検出結果が得られた。顔検出も 90cm 以内であればほぼ安定して顔検出を行うことが確認できた。また、全体の肌色情報による検出も眼鏡を装着することで若干抽出結果に減少が見られたが、正しく存在を検出することができた。しかし、マスクを装着した際には顔検出ができず、全体の肌色情報による検出はできたが、学習者が存在しているときに少しでもカメラからの距離が遠くなっただけで、検出できない結果になった。

表 6.6 眼鏡とマスクを着用した際の肌色抽出結果

	顔領域	全体	背景	全体 (学習者がいない)
眼鏡装着時	1332px	1654px	322px	380px
マスク装着時	0px	939px	939px	339px

6.5 比較検証の考察

6.5 比較検証の考察

比較検証の結果に対して考察を行った。本節では、なぜカメラの違いや、照明環境の違い、装飾品の違いから検出結果および顔検出、肌色抽出結果に違いがでたのか理由を明らかにする。そして、これらの検出結果の違い、抽出結果の違いから今後どのような対策が必要となるか考察を行う。

6.5.1 カメラの違いと照明環境の違い

肌色情報の抽出結果について

本検証から使用するカメラによって色の発色が違うということがわかった。実際に使用したカメラで撮影した画像を図 6.1 に示す。



図 6.1 各カメラで撮影した画像

これらの画像からわかるように、使用するカメラによって明るさが異なり、肌色にも違いが見られる。本研究では HSV 表色系の H (色相) の値が 0 から 30 のときに肌色として抽出するようにしているが、このように、使用するカメラによって発色が違うため、H の値が変わってしまい、肌色情報の抽出結果に差が現れたと考えられる。

また、照明環境の違いにより検証を行った。照明環境の違いにより抽出できる肌色情報に差がでることを確認した。とくに、暗い (蛍光灯ひとつ点灯) 環境であれば、安定して肌色情報を抽出することが難しいことが示唆された。これは、学習を行う環境が暗くなってしまったため、設定しておいたしきい値では肌色情報を抽出できなかったことが原因だと考えられる。

6.5 比較検証の考察

本システムでは、学習開始時に顔領域内の肌色情報と全体の肌色情報、背景の肌色情報を取得することにより、照明環境や使用するカメラの等の学習環境の変化に対応しているため、肌色抽出結果が変わることはさほど問題ではないと考えられる。そして、顔領域内の肌色情報の占める割合が検出に大きく関するため、顔領域内の肌色情報を安定して取得することができれば、安定した検出結果を得ることができる。

肌色情報を安定して取得するためには、肌色情報のしきい値を変化させるなどの対策が必要である。学習開始時に顔領域内の肌色情報を抽出する際に、顔領域内の全ピクセルから H (色相) の平均値を取得する。この平均値から、その学習環境における肌色抽出のしきい値を設定する等の対策を施す必要がある。

学習者がいるときといないときで肌色情報が変化したことについて

学習者がいるときには肌色として抽出されなかった部分が、学習者がいなくなった際に肌色として抽出されたことを確認した。表 6.4 に記載してある Logicool 1.3-MP WebCam C200 を利用した検証では、学習者がいなくなったときに 2088px もの肌色情報を抽出し、学習者の存在も正しく検出することができないという結果になってしまった。これは、カメラが自動でホワイトバランスを調整した際に、各画素の H (色相) の値が変わってしまったことが原因であると考えられる。

この結果から、現在普及している Web カメラにどのような特性があるのかを実際に調べる必要がある。そして、価格や、画素数、画角、解像度等の違いから検出結果にどのような影響がでるのかを調べ、その影響に対して適切な対応をとる必要がある。

6.5.2 装飾品の有無

眼鏡やマスクの有無によって顔検出と全体の肌色情報による検出に影響がでることを確認した。眼鏡を着用した際は、顔検出を安定して行うことができ、さらに全体の肌色情報からも問題なく検出することができた。しかし、マスクを装着した際には、顔検出ができず、全

6.5 比較検証の考察

体の肌色情報による検出にも影響がでる可能性があることがわかった。

顔検出ができなかった原因として、学習者の口と鼻がマスクで隠れてしまったことが原因であると考えられる。本研究は顔検出をパターンマッチングにより行っているため、これらがマスクで隠れてしまったことによって、学習パターンと同様のパターンを抽出することができなかったと考えられる。また、全体の肌色情報による検出においても、本システムでは背景の肌色情報有無を考慮し、顔領域内の肌色情報と背景の肌色情報を算出した上で検出を行っている。しかし、本システムでは、顔検出を行った上で、その領域内の肌色情報を求めているため、マスク装着時に顔領域内の肌色情報取得することができないという問題がある。

この対策として、本システムでは顔検出を行った上で顔領域内の肌色情報を取得しているが、決められた枠内に顔を入れて撮影した際に、その枠内における肌色情報を同時に取得する方法が考えられる。この方法を用いることで、顔検出ができる状態でもできない状態でも、顔領域内の肌色情報を取得することができる。また、学習者が必ずこの指定された領域内に顔を入れて撮影を行うとは言い切れない。さらに、マスクを装着することで取得できる顔領域内の肌色情報が減少してしまう。そのため、肌色情報からだけでなく、服の色も検出の対象する等の対策を施す必要がある。

第7章

おわりに

本研究では、学習者がコンピュータの前において学習を行っているか把握するために、インターネットを利用した e-Learning において学習者をリアルタイムで検出するシステムを構築した。また、本研究で構築したシステムを用いて、学習者の存在を正しく検出することができるか評価を行った。評価の結果として、学習者がコンピュータの前にいる状態と、コンピュータの前にはない状態は正しく検出することができた。しかし、学習者が大きく姿勢を変えた時に誤検出が発生した。

誤検出は、学習者の顔がカメラに入っていなかったとき、学習者がおおきくのけぞってしまったとき、学習者の顔がカメラに大きく近づいたときに発生した。この結果に対して、肌色のしきい値を設定しなおすことや、肌色だけでなく服の色も検出の対象に入れること、映像から判断するだけでなく、検出できない時間を計測し、その時間に応じて検出結果を変化させる等の対策を行い、誤検出を防止する必要がある。

また本研究では、処理の最後で動体検出により学習者の存在をより詳しく検出しているため、学習者がコンピュータの前にはないときに、背景で学習者以外の人を通った場合に“学習者がいる”と判定されてしまう可能性がある。そのため、今後の課題としてこれら背景で人が通るような動的に変化する環境においても学習者がいる状態といない状態を正しく検出できるようにする必要がある。

今後の展望として、学習者の存在を検出するだけでなく、寝ている状態等のより詳しい学習者の状態を検出することが求められる。さらに、コンテンツや LMS と連携し、学習者の状況によって再生状況を変化させること、学習ログとして学習者がいる状態といない状態を撮影し、画像として LMS に保存する機能の実装が求められる。

謝辞

本研究および本論文に関して、多大なるご指導、適切なお助言を頂きました高知工科大学情報システム工学科、妻鳥貴彦講師に心より御礼申し上げます。

ご多忙な中、本研究の副査をお引き受け頂き適切なお助言を頂いた同学科、岡田守教授に心より御礼申し上げます。

同じくご多忙な中、本研究の副査をお引き受け頂き適切なお助言を頂いた同学科、吉田真一講師に心より御礼申し上げます。

同研究室の修士2年の畠山博和氏、藤原健太郎氏、山崎雄大氏にはご自身の研究もある中で、本研究および本論文に関して数々のご助言、ご指導をいただきましたことを心より感謝いたします。

同じく同研究室の修士1年の清水雅也氏、森拓也氏には就職活動でお忙しい中、本研究のシステムの実装について、梗概の書き方や本論文の書き方について、適切なお助言、ご指導をいただきましたことを心より感謝いたします。

同じく同研究室の学部4年生の中澤大樹氏、濱野純平氏、細川恭平氏、前田晃宏氏、松井勇貴氏、には本研究に関して、多くのご助言をいただきましたことを心より感謝いたします。

同じく同研究室の学部3年のみなさんには、本研究の評価等にお付き合いいただきましたことを心より感謝致します。

本研究で、多くの人に迷惑をかけながら様々なことを学ばせていただきました。今後大学院に進学し研究活動を行う中で、本研究で学んだことを活かして今後の修士論文に精進していきます。

参考文献

- [1] 総務省, “大学設置基準”, <http://law.e-gov.go.jp/cgi-bin/idxsearch.cgi>
- [2] 小野 邦彦, 後藤 幸功, 半田 純子, 本間 千恵子, 遠藤 孝治, 鈴木 克明, “サイバー大学の e ラーニングに関する質保証の取組”, 日本教育工学会 第 25 回全国大会, pp501-502, 2009.
- [3] 高度情報通信ネットワーク社会推進戦略本部 (I T 戦略本部) , “IT 新改革戦略”, <http://www.kantei.go.jp/jp/singi/it2/kettei/060119honbun.pdf>
- [4] 文部科学省, “「遠隔授業」の大学設置基準における取り扱い等について”, http://www.mext.go.jp/b_menu/shingi/12/daigaku/toushin/971202.htm
- [5] 総務省, “短期大学設置基準”, <http://law.e-gov.go.jp/cgi-bin/idxsearch.cgi>
- [6] 総務省, “高等専門学校設置基準”, <http://law.e-gov.go.jp/cgi-bin/idxsearch.cgi>
- [7] 総務省, “大学通信教育設置基準”, <http://law.e-gov.go.jp/cgi-bin/idxsearch.cgi>
- [8] Sherrah, J. & Gong, S (2001). Skin Color Analysis. CvOnline. <http://homepages.inf.ed.ac.uk/cgi/rbf/CVONLINE/entries.pl?TAG288>, 2009 年 2 月 5 日