

修士論文

複数のPTZカメラを用いた保育施設における
複数ユーザの要求に応じた映像提示システム

大野 純佳

2007年3月9日

奈良先端科学技術大学院大学
情報科学研究科 情報システム学専攻

本論文は奈良先端科学技術大学院大学情報科学研究科に
修士(工学) 授与の要件として提出した修士論文である。

大野 純佳

審査委員：

横矢 直和 教授 (主指導教員)

小笠原 司 教授 (副指導教員)

複数のPTZカメラを用いた保育施設における 複数ユーザの要求に応じた映像提示システム*

大野 純佳

内容梗概

近年の計算機の高性能化や通信の高速化に伴い，遠隔地からの映像配信が容易になり，自宅等から保育施設にいる子供の様子を見ることが出来るライブカメラシステムへの関心が高まりつつある．このようなシステムではユーザの要求に応じた映像を提供する必要があり，さらに複数人の同時アクセスにも対応する必要がある．既存システムではPTZ(Pan,Tilt,Zoom)カメラを操作してユーザは自分の見たい方向を高い解像度で見ることが出来るが，複数人が同時にアクセスしている場合には操作権を待つ必要があった．監視システムの分野では，複数人を複数のカメラで追跡するようカメラを制御する試みも成されているが，どの人物の映し方も同じ設定となっており，複数ユーザそれぞれの要求が考慮されたカメラ制御は行われていない．

本研究では，複数人が同時にアクセスした場合でも，各ユーザの要求に合った映像を提示することが可能なシステムの実現を目標とする．そのために，複数台のPTZカメラを用い，そのカメラ台数よりも多いアクセスがあった場合にもユーザの要求を満たすようにカメラを自動制御する手法を提案する．本研究では事前に行ったアンケートの結果を基にユーザの要求とその要求の満足度を表す評価関数を定義する．そして評価関数により求めたユーザの満足度が最大となるカメラをユーザに割り当て，カメラ毎に，同じカメラを割り当てられたユーザの満足度の総和を最大とするようにカメラパラメータを最適化することでPTZカメラを制

* 奈良先端科学技術大学院大学 情報科学研究科 情報システム学専攻 修士論文, NAIST-IS-MT0551029, 2007年 3月 9日.

御する．シミュレーション実験により，ユーザに提示する映像を定量的に評価することでカメラ制御手法の妥当性を示す．さらに本手法を適用したライブカメラシステムを構築し，実環境での被験者実験により本システムの有効性を検証する．

キーワード

複数 PTZ カメラ，複数ユーザ，ユーザの要求，カメラの自動制御，保育施設

Video presentation system satisfying multiple users' requests using multiple PTZ cameras in child-care facility*

Sumika Ohno

Abstract

With high-speed network and high performance PCs, live broadcasting camera systems which allow users to look at states of children in a child-care facility have become popular. Such a system has to present videos satisfying users' requests to users and it is necessary to cope with simultaneous accesses of multiple users. In conventional systems, a user controls a PTZ camera and can look at the direction of user's request with high resolution. However, a user is constrained to wait for the control when multiple users access to the system simultaneously. In the field of visual surveillance, multi-camera control to track multiple persons is investigated. Cameras are usually controlled based on the same strategy. Therefore systems do not satisfy multiple requests of users.

This study aims at realizing a system which can show live videos which satisfy requests of users when multiple users access to the system simultaneously. To realize the system, this paper proposes a method for controlling multiple PTZ cameras automatically for satisfying requests of users when the number of users is larger than the number of PTZ cameras. An evaluation function representing user's satisfaction is defined based on a result of a questionnaire carried out in advance. The proposed method assigns a camera to each user that maximizes the

* Master's Thesis, Department of Information Systems, Graduate School of Information Science, Nara Institute of Science and Technology, NAIST-IS-MT0551029, March 9, 2007.

evaluation function. Then, the method controls multiple PTZ cameras optimizing camera parameters so as to maximize the sum of evaluation values of users' satisfaction. In the experiment of camera control, the validity of the proposed method has been quantitatively evaluated using videos generated from a simulated environment. Furthermore, a live broadcasting camera system based on the proposed camera control method has been developed and the availability of the system has been examined with a subjective experiment in a real environment.

Keywords:

multiple PTZ cameras, multiple users, users' requests, automatic camera control, child-care facility

目次

1. はじめに	1
1.1 保育施設ライブカメラサービスに関する背景	1
1.2 自動カメラ制御による映像提示の関連研究	3
1.2.1 監視の分野における映像提示	3
1.2.2 番組制作における自動カメラワークによる映像提示	6
1.2.3 複数ユーザの要求に応じた講義の映像提示	9
1.3 本研究の位置付けと方針	9
2. 複数の PTZ カメラを用いた複数ユーザの 要求に応じた映像提示	11
2.1 ユーザの満足度を表す評価関数の定義	11
2.1.1 事前アンケートによる調査	11
2.1.2 評価関数	14
2.2 評価関数に基づくカメラの割り当てと制御	18
2.2.1 概要	18
2.2.2 各ユーザへのカメラの割り当て	19
2.2.3 評価値の最大化によるカメラ制御	19
2.3 シミュレーションによるカメラ制御実験	20
2.3.1 実験条件	20
2.3.2 提示画像の定量的評価	21
3. 複数の PTZ カメラによるライブカメラシステム	29
3.1 システムの設計方針と概要	29
3.2 機器構成	32
3.3 観測人物の位置推定	32
3.4 ユーザの要求入力インターフェース	34
3.5 ユーザの要求に応じた映像の配信と受信	36

4. ライブカメラシステムの被験者実験	38
4.1 実験環境と条件	38
4.2 実験結果	43
4.3 被験者による主観評価	46
4.4 考察	49
5. まとめと今後の課題	51
謝辞	53
参考文献	54
付録	57

目 次

1	PTZ カメラを用いた既存のライブカメラサービス [1]	2
2	浮田らのシステムにおける対象観測結果 [12]	5
3	知的ロボットカメラ (Copyright 2001 NHK)[19]	7
4	デジタルカメラワーク [17]	8
5	デジタルカメラワークによる生成画像 [17]	8
6	アンケートで提示した画像	13
7	イマジナリーライン	16
8	イマジナリーラインによるカメラ間の遷移係数の決定	17
9	カメラ割り当て・制御の処理の流れ	18
10	シミュレーション実験での実験環境	20
11	各ユーザへの提示画像における指定人物の画像中心からの距離	23
12	人物の映る大きさの例	24
13	実験 I-10 でのユーザ H への提示画像	25
14	満足度の変化とカメラの割り当て変更	27
15	システムの処理の流れ	30
16	プロトタイプシステムの構成	31
17	ユーザのインタフェース	34
18	カメラ映像の配信・受信	36
19	実際の環境での実験環境	39
20	室内の様子 (実験前)	39
21	実験毎に行うアンケート用紙	40
22	事後アンケート用紙 1 枚目	41
23	事後アンケート用紙 2 枚目	42
24	各ユーザへの提示画像 (A ~ C)	44
25	各ユーザへの提示画像 (D ~ E)	45
26	カメラ 2 台の実験時におけるカメラ 1 の画像	50

表 目 次

1	サービス利用時間に関するアンケート結果	12
2	ある時間にある人数がアクセスする可能性 (%)	12
3	サービス使用中の各シーンを見る時間の割合	12
4	実験パターン	22
5	各ユーザへの提示画像上で指定人物の画像中で占める割合	24
6	各ユーザへの提示画像における映る人数	25
7	システムが算出した各ユーザの満足度	25
8	各ユーザへの提示画像上で指定人物の占める割合 (実験 II-15)	27
9	PC のスペック	31
10	各ユーザの子供	43
11	システムが算出した各ユーザの満足度の平均	46
12	提示映像に対する評価結果	47
13	システムの評価結果 (評価)	47
14	システムの評価結果 (コメント)	48
15	全実験パターン	57
16	実験 IV における各ユーザの希望人数	57
17	実験 III,IV-5 の評価結果	58
18	実験 III,IV-10 の評価結果	58
19	実験 I,II-15 の評価結果	59
20	実験 III,IV-15 の評価結果	59
21	実験 I,II-35 の評価結果	60

1. はじめに

近年の計算機の高性能化や通信の高速化に伴い，遠隔地からの映像配信を利用したサービスや研究が盛んに行われている．保育施設におけるライブカメラサービスもその一つであり，保育室に設置されたカメラの映像をインターネットを通じて自宅等から見るができる．これは保育施設に預けている子供の様子を家事の合間や仕事に出ている間に見ることや，家が離れている祖父母が普段なかなか見ることのできない孫の顔を見ることができる等の理由から人気のサービスとなっている [1, 2, 3]．この様なライブカメラサービスではユーザ毎に『どの子供を見る』等の異なる要求があり，その要求に応じた映像を提供する必要がある．さらに複数のユーザが使用するサービスなので，複数ユーザが同時にサービスを使用する可能性があり，限られたカメラ台数で複数人の同時アクセスにも対応する必要がある．

そこで本研究では，高解像度に情景を捉えられる PTZ カメラを複数台使用し，そのカメラ台数よりも多いアクセスがあった場合でも複数ユーザそれぞれの要求に応じた映像を提示するシステムの実現を目標とする．以下本章では，保育施設におけるライブカメラシステムに関する背景と自動カメラ制御による映像提示の関連研究を概観し，本研究の位置付けと方針について述べる．

1.1 保育施設ライブカメラサービスに関する背景

インターネットコム株式会社・株式会社インフォプラントが300名の保育園・幼稚園に通う子供をもつ親に対して行ったアンケート調査 [4] では，インターネットを通じて保育施設の様子をパソコンや携帯電話で見られるサービスを利用したいという意見は76%と高かった．この結果からも，保育施設に預けた子供の様子が心配で自分の目で確かめたいと思う親が多いことが分かる．同アンケート調査において，園児の様子をインターネットを通じて公開することについては，アクセス制限をかける等の条件付きの意見は多いものの「賛成」とする意見が75%となっていた．プライバシーの問題やセキュリティなど，慎重に考慮すべき問題はあるが，安心して子供を預けたいと願う保護者と保育施設をつなぐ一つの手段と



図 1 PTZ カメラを用いた既存のライブカメラサービス [1]

して、このようなサービスへの期待は高く、今後もますます普及することが考えられる。

保育施設での子供の様子をインターネットを通してパソコンや携帯電話から見ることのできる既存のライブカメラサービスとしては、PTZ(Pan, Tilt, Zoom)カメラを使用したシステムが一般的である [1, 2]。これは、図 1 のシステムのようにカメラの操作機能が付いており、ユーザは自由にカメラを操作して自分の見たい方向を高い解像度で見ることができる。高解像度に場面を捉えられる PTZ カメラを使用していることで、子供の表情まで読み取ることができるので、子供がどのような状態にあるかを把握しやすい。しかし、複数のユーザが同時にアクセスしている場合には、操作権の取得を待つ必要があり、操作権のないユーザにとっては自分の子供が映らないといった無意味な映像が提示される可能性がある。

一方、全方位カメラを使用したライブカメラサービス [3] も運用されている。これは鉛直下向きの曲面ミラーと鉛直上向きのカメラから構成される全方位カメラ [5] により周囲 360 度を一度に撮影可能となっており、全方位画像からユーザの見たい方向の画像を切り出して透視投影変換することにより映像を提供する。

そのため、複数のユーザが同時にアクセスしていても、複数のユーザはそれぞれの見たい方向を同時に見ることができる。しかし曲面ミラーによる全方位カメラは1台のカメラで周囲360度の情景を映すために、空間分解能が一般的なカメラよりも低く、全方位画像から生成した特定方向の平面透視投影画像は通常のカメラに比べ低解像度な画像となってしまいうために、子供の表情まで見たいという要求には応えられない。

1.2 自動カメラ制御による映像提示の関連研究

ある空間に集まる人々の様子をカメラで捉えて第三者にその映像を提示する場面は、保育施設におけるライブカメラサービス以外にも数多く考えられる。例えば、美術館や空港等の監視、会議や講義への遠隔地からの参加やスポーツシーンにおける映像の提示、また様々なテレビ番組もこのような例に挙げられる。このようなそれぞれの状況において、人の動きに合わせて場面を自動で撮影して映像を提示しようという研究は盛んに行われている。そこで本節では、それらの研究を保育施設のライブカメラサービスに適用可能かどうかを検証するために、監視の分野における映像提示、スポーツシーンやテレビ番組制作における自動カメラワークによる映像提示、そして、複数人の要求に応じた映像提示の関連研究について概観する。

1.2.1 監視の分野における映像提示

監視の分野では監視領域に侵入した人物の詳しい情報を得るために、侵入者の注視（追跡）が必要となる。そのため、監視領域への侵入者の追跡を行い、侵入者の映像をより詳細に提示しようとする研究が従来から行われている [6, 7, 10, 8, 9, 11, 12]。

同一人物の継続的追跡

同一人物の継続的な追跡は文献 [6, 7, 10] において実現されている。Taoら [6] は固定のステレオカメラを利用し、人物の三次元位置を正確に推定することで広範囲に複数台配置されたカメラ間において、人物同定を正確に行い、広範囲を歩

く同一人物を捉え続ける．また中澤ら [7] は複数の固定カメラのうち対象を観測可能なカメラにおいて，対象と 1 対 1 に対応付けられたエージェントを生成し，各対象の追跡を行う．異なるカメラ間においてエージェント同士が対象情報を交換することで，追跡を行っていたカメラ画像に対象が映らなくなった場合にも，他のカメラで対象を観測することができる．また，各カメラにおいて観測可能な対象が存在すればエージェントを生成するので，カメラ台数よりも多い複数対象を観測することが可能である．複数のアクティブカメラを使用したシステムとして楊ら [10] は指定人物のリアルタイム追跡を行っている．各カメラ画像中のオプティカルフローに基づいて検出した移動人物の動きに合わせてカメラの視線の向きを変化させ，移動人物が遮蔽された場合でも複数のカメラからの情報を元に人物領域を予測して再び現れた時にも追跡を続けることができる．これらの研究では，カメラの撮影可能範囲内の人物を継続して捉えることが可能である．しかし，ズーム制御が行われないため，対象を観測する解像度が限られてしまう．さらに，文献 [7] において，もし PTZ カメラを使用したとしても，各カメラにおいて複数のエージェントがそれぞれの追跡対象を注視するために，一つのカメラに対して同時に相反するカメラ制御を要求する可能性がある．

侵入者の詳細な映像提示

被写体をより詳細に映し出すため，複数の PTZ カメラで複数人を追跡する試みも成されている [8, 9, 11, 12]．Faisal ら [8] や Cash ら [9] は，駅などの広範囲を歩く複数人の一人一人を，カメラをズームインして高解像度に捉えるために効率の良いカメラ制御の手法を提案し，シミュレーション環境において検証実験がされている．しかし実環境では，人物の動きは人それぞれ異なる上，一人一人についても動く速度などに変化がある．そのため実際には，このような変化のある人物の動きにカメラ制御を合わせることが必要である．実環境で複数の PTZ カメラを用いたシステムとしては浮田らの人物追跡システム [11, 12] が挙げられる．文献 [11] ではカメラと 1 対 1 に対応付いたエージェントにより，同じ対象を追跡するエージェントを生成し，カメラ間で対象の情報を交換して，継続的に，また高解像度に対象を捉えることができる．しかしカメラとエージェント，エージェントと対象は共に 1 対 1 に対応付いているため，カメラ台数よりも対象数が多い

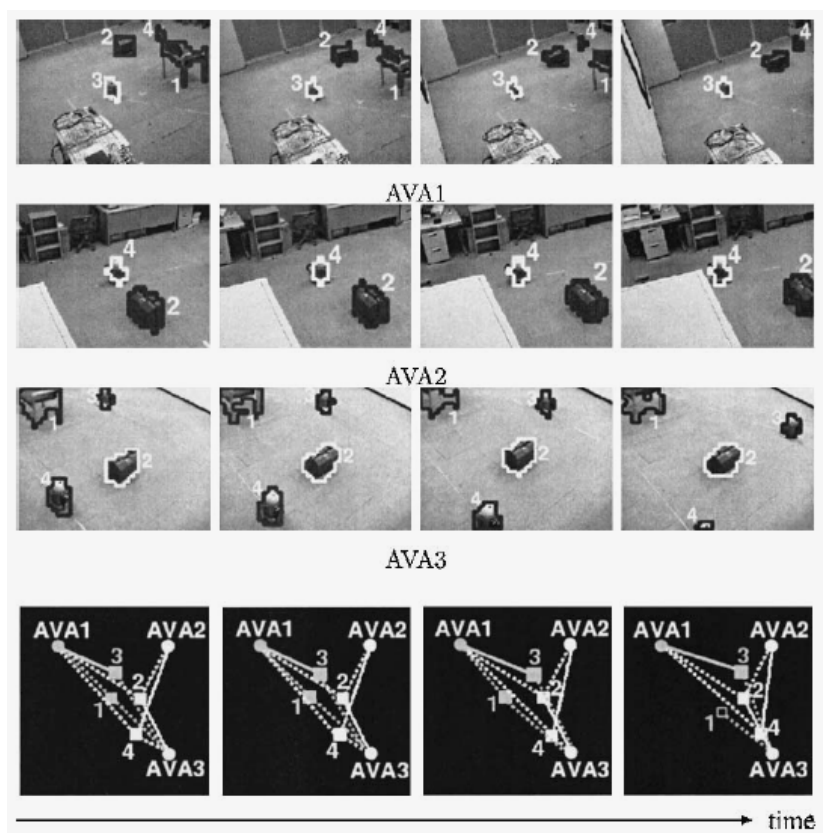


図 2 浮田らのシステムにおける対象観測結果 [12]

場合は追跡不可能になってしまう．文献 [12] において，カメラ台数よりも多い対象の追跡を，エージェントの追跡対象以外の対象が画像上に映っていた際にその情報を交換するという手法で実現している．このシステムによる対象の観測結果を図 2 に示す．結果画像中の AVA は各カメラにおけるエージェントを表し，観測対象 1,2,3,4 を 3 台のカメラで捉えている．このシステムでは対象重要度を設けて，重要度の低い対象は継続的な注視が保障されていなくても，一時的に対象情報を収集するだけで十分であるとしており，図 2 における観測対象 1 のように画像端に捉えられることも多い．しかし，保育施設のライブカメラサービスのようにユーザ毎に見る対象が異なる場合，ただ対象が画像内に映っていれば画像端でも良いというものでは，ユーザは満足できない．そして，複数のユーザへの映像

提示を想定したシステムではないため、どの人物の映し方も同じ設定となっており、複数ユーザそれぞれの要求が考慮されたカメラ制御は行われない。

以上のように監視の分野では、監視領域に侵入する全ての人物を追跡することが必要であり、なるべく多くの対象を捉えることを目的としている。しかし、保育施設のライブカメラサービスでは、このような状況と異なり、サービスにアクセスしているユーザの见たい方向・範囲を映すことが必要である。このためには、シーンを捉える複数台のカメラを自動で制御し、それぞれのユーザの要求に合ったカメラの映像のみを提示することが必要である。

1.2.2 番組制作における自動カメラワークによる映像提示

カメラのパンニングやチルティング、ズームングを自動で行い、シーンを撮影して映像を生成する研究として、テレビ番組のような高品質な映像の提示を目標とした研究が行われている [16, 17, 15, 18, 13, 14]。

PTZ カメラによる自動カメラワーク

人間が行うような自然なカメラワークで、効率よく映像を作成することを支援する知的ロボットカメラが研究されている [16, 19]。NHK 放送技術研究所では、スタジオやスポーツ中継における放送カメラマンのカメラワークを分析し、それに基づいたカメラ制御のアルゴリズムを実装し、スタジオでの 2 人のトークを色認識によって自動撮影することを実現している (図 3)。尾関ら [17] はテレビ番組でよく用いられるショットの分類と、注目対象物と注目すべき状態の分類を行い、適したカメラワークを選択することで机上作業シーンを自動撮影している。また、井口らはアメリカンフットボールの試合映像をリアルタイムで解析して複数台の PTZ カメラを自動制御し、状況に適したカメラへ切り替えることで映像を生成している [15]。これは試合におけるプレイのパターンにより、あらかじめ選手の状況に適した画角のカメラを準備しておく。そして選手群の分布と動きを解析してどのプレイかを判断し、プレイに合った画角が設定されたカメラへ切り替える。プレイのパターンが決まっており、人物の動きが予測できるスポーツ映像を生成する手法としては有効であると考えられる。Yong ら [18] は、各カメラに接続される PC からのカメラの状態情報により、講義の状況に応じたカメラ

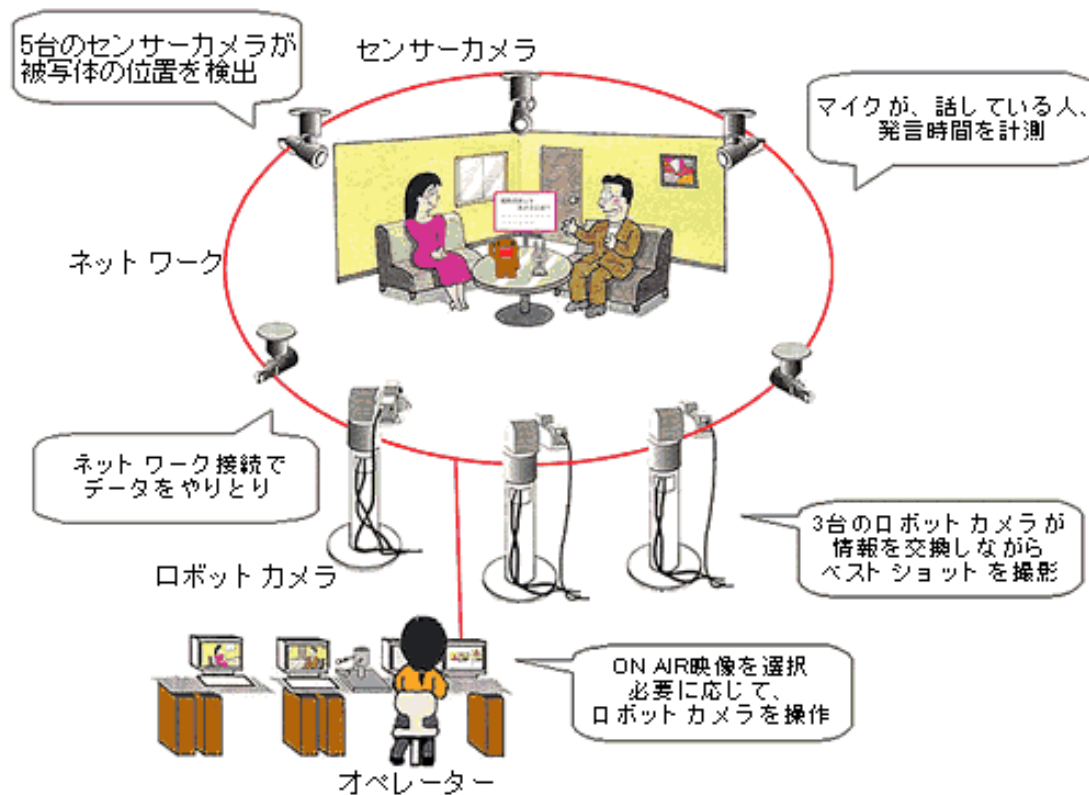


図 3 知的ロボットカメラ (Copyright 2001 NHK)[19]

制御の遷移確率を利用して次のカメラワークを選択することで講義の自動撮影を行っている。しかし、これらの自動撮影の研究はいずれも、すべての視聴者に対して同じストリームを生成するものであり、ユーザによってシーンの捉え方の要求が異なるといったことは考慮されていない。

固定カメラによるデジタルカメラワーク

大西ら [13] や熊野ら [14] は固定カメラによって撮影した映像の一部分を切り取って拡大提示することで映像の生成を行い擬似的にカメラを動かして撮影したかのように見せかけるデジタルカメラワークによりスポーツシーンの映像生成を行っている。デジタルカメラワークでは、切り取る領域を移動させることでパンニングやチルティング、拡大率を変化させることでズームングの効果を作

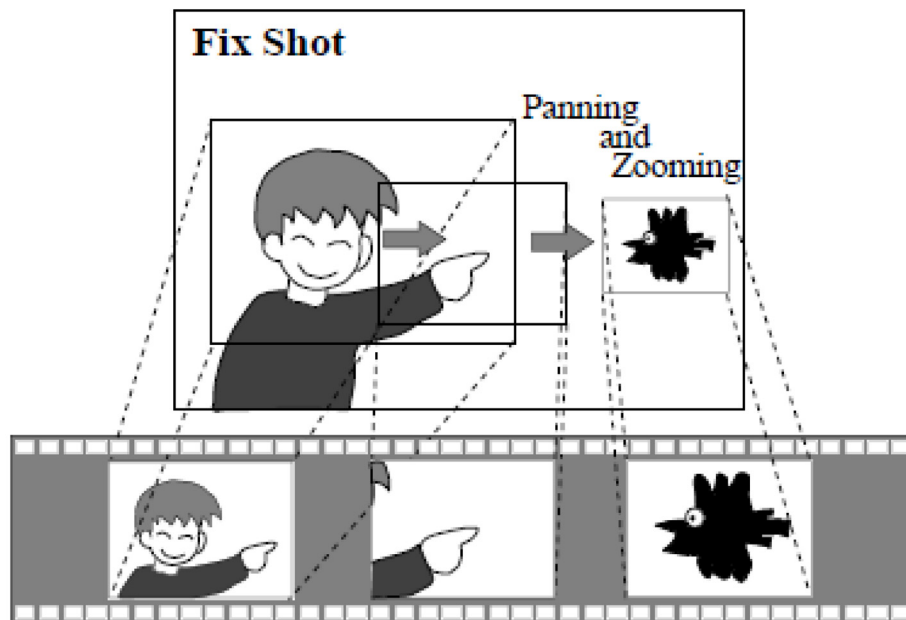


図 4 デジタルカメラワーク [17]



図 5 デジタルカメラワークによる生成画像 [17]

り出すことができる(図4, 図5)。そのため, 全方位カメラを用いたライブカメラサービス [3] のように複数ユーザそれぞれの要求に応じた映像を提示可能である。しかし, 固定カメラに映る映像を切り取って拡大提示するために解像度は低くなる。UXGA(1600 × 1200pixels) の高解像度の広角カメラを使用して90度の範囲を映してデジタルカメラワークを行う場合と, 水平画角4.4度~48.8度のVGA(640 × 480pixels) のPTZカメラにより光学ズームを行う場合を比較する。デジタルカメラワークでは光学ズームは行わないので, ズーミングをしても空間分解能は変わらず, UXGAのカメラで90度を捉えるときの空間分解能では, このPTZカメラでワイド端から2倍に光学ズームしたときの効果を出すことすらできない。また, PTZカメラの解像度も今後上げられるであろう。このことより, 子供の表情を捉えるためにはデジタルカメラワークではなく, PTZカメラを使用したカメラ制御により映像を提示することが適していると考えられる。

1.2.3 複数ユーザの要求に応じた講義の映像提示

PTZカメラを用いて複数ユーザそれぞれの要求に応じた映像を提示するシステムには, 宮崎ら [20] の複数ユーザに対する講義の実時間映像化の研究がある。この研究では講義状況を5つに分類し, それぞれに各ユーザの希望するカメラワークを事前に入力しておき, 講義状況に応じて多数決によりカメラワークを選択する。そして, 各ユーザの要求に近いカメラ制御を行うカメラの映像を, それぞれのユーザに提示する。多数決からはずれたユーザの要求は考慮されないが, 講義という状況下なので多数決から外れてもある程度意味のある映像になると考えられる。しかし, 保育施設におけるライブカメラサービスではユーザの子供が映像に映っていることが重要であり, 多数決から外れたユーザにとっては意味のない映像が提示されてしまうことが考えられる。

1.3 本研究の位置付けと方針

保育施設におけるライブカメラサービスでは, サービスを使用しているユーザの見た方向の映像を提示する必要がある。しかし, 例えば見たい子供が異なる

等の理由から，見たい方向がユーザによって異なる．また，ライブカメラサービスを使用するユーザは子供の置かれている状況や体調を把握したいという思いがあり，高解像度のカメラで子供の様子を捉えることも必要である．そのため，ライブカメラサービスに使用するカメラとしては，各ユーザの見たい方向を映すように操作可能な PTZ カメラが適していると考えられる．しかし，このようなサービスは複数のユーザが使用するものであり，限られたカメラ台数に対して複数ユーザが同時にアクセスする可能性が高い．このような場合には，後からアクセスしたユーザは PTZ カメラの操作権を待つ必要がある．

そこで本研究では，複数台の PTZ カメラを使用し，そのカメラ台数よりも多いアクセスがあった場合でも複数ユーザそれぞれの要求に応じた映像を提示するシステムの実現を目標とする．1.2 節で挙げた関連研究ではユーザによってシーンの捉え方の要求が異なるということは考慮されないものがほとんどであった．複数ユーザそれぞれの要求に応じた映像提示を行う宮崎らの手法 [20] は，多数決でカメラワークを決定するので，講義という状況においては有効な手法であるが，ユーザによっては子供が映らない可能性のあるこの手法は目標とするライブカメラシステムには適さない．

本研究ではユーザに提示される映像に対して，ユーザの満足度を表す評価関数を定義し，この評価関数の最大化によりカメラ制御を行う．また評価関数は，ユーザがどのような映像を見たいと思っているのかをアンケートにより調査し，この結果から定義する．複数人が同時にアクセスすることについては，複数ユーザの満足度を高くするカメラの割り当てと制御を行うことで，カメラ台数よりも多いユーザの要求に応じた映像提示を実現する．そして，提案手法を適用したライブカメラシステムを実現する．さらに，被験者実験により，そのシステムの有効性を検証する．以下，2 章では提案する映像提示の手法について述べ，3 章では提案手法を適用したシステムの構築について述べる．4 章では実環境において実験を行い，その結果と被験者による主観評価からシステムの有効性について述べる．そして，最後に 5 章で本論文のまとめと今後の課題について述べる．

2. 複数のPTZカメラを用いた複数ユーザの 要求に応じた映像提示

本章では、保育施設におけるライブカメラサービスのための、複数ユーザの要求に応じた映像を提示する手法について述べる。そして、シミュレーション実験により、手法に合ったカメラ制御ができているかを確認する。

2.1 ユーザの満足度を表す評価関数の定義

複数のPTZカメラの制御をユーザの要求に合わせて行うために、各カメラの画像に対する各ユーザの満足度を表す評価関数を定義する。そのために、ユーザはライブカメラサービスでどのような映像を見たいと思っているのかをアンケートにより調査した。そして、このアンケート結果に基づいて評価関数を定義する。

2.1.1 事前アンケートによる調査

アンケートは、2006年4月に開催されたNAISTのオープンキャンパスに参加した約3歳～12歳の子供を持つ保護者92名に対して行った。まず、同時に何人がアクセスするかを考察するために、サービスを利用するとした時の、利用頻度、利用時間、利用時間帯を調査した。その結果を表1に示す。文部科学省の定める幼稚園における一クラスの定員制限は35名であり、1回のアクセス時間の平均が13.3分であった。所与の時間間隔で発生する離散的な事象を数える特定の確率変数を持つ離散確率分布であるポワソン分布を利用して、ある時間にある人数がアクセスする確率を算出した結果を表2に示す。この表は、1時間から5時間の各時間において、何人が同時にアクセスする可能性があるかを示している。1時間の間隔で考えると15人以上が同時にアクセスする確率は1%を下回っている。しかし、サービスの利用時間帯が開園時間から14時の5時間に分布していることから、2時間以上の間隔で考えると、10人以上が同時にアクセスする可能性が1%を下回って低いことが分かる。このことから、本研究では、10人程度までの同時アクセスに対応できるシステムの実現を目標とする。

表 1 サービス利用時間に関するアンケート結果

アンケート項目	アンケート結果
サービスを利用する頻度	平均1.7回/1日
サービス1回の利用時間	平均13.3分
サービスの利用時間帯	開園時間～14時

表 2 ある時間にある人数がアクセスする可能性 (%)

	1人	2人	3人	4人	5人	6人	7人	8人	9人	10人	11人	12人	13人	14人	15人
1時間	0.3	1.3	3.3	6.4	10.0	12.9	14.3	13.9	12.0	9.3	6.6	4.2	2.5	1.4	0.7
2時間	8.0	15.6	20.1	19.5	15.1	9.8	5.4	2.6	1.1	0.4					
3時間	19.5	25.2	21.7	14.0	7.3	3.1	1.2	0.4							
4時間	27.9	27.0	17.5	8.5	3.3	1.1	0.3								
5時間	32.8	25.5	13.2	5.1	1.6	0.4									

表 3 サービス使用中の各シーンを見る時間の割合

自分の子供だけ	2人 (自分の子含)	5～6人 (自分の子含)	部屋全体 (自分の子含)	他人の子供	園の外 (子供無)
21.74%	17.56%	25.25%	24.16%	6.16%	5.11%



(a) 自分の子供だけ



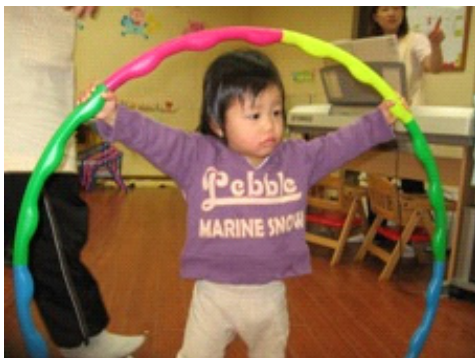
(b) 自分の子供と友達の2人



(c) 自分の子供と5~6人の友達



(d) 部屋全体



(e) 他人の子供



(f) 園の外の方(子供以外)

図6 アンケートで提示した画像

また，ライブカメラサービスを使用しているとして，異なる人数を含めた図6に示す各画像6枚を例として見てもらい，サービスにアクセス中，その各画像のような映し方の映像をどのくらいの時間見ると思うかを質問した．そしてアクセス時間に占めるそれぞれのシーンを見る時間の割合を求めた結果を表3に示す．自分の子供が映っていないシーンについては全体の1割程度と値が低く，自分の子供を映す必要があることが分かった．また，何人を含めて映すかについては答えにバラつきが見られ，ユーザにより，その時の子供の状況等から見たい人数が異なるということが分かった．以上の考察から，本研究ではユーザの要求を「見たい子供」と「見たい人数」の2点とした．

2.1.2 評価関数

評価関数はユーザの要求に合わせて，ユーザの「見たい子供」と「見たい人数」を映すことで評価値が高くなるように設定する．そこで，あるカメラ i についてあるユーザ u の満足度を表す評価関数 S_{ui} は，カメラ i のパン角を θ_i ，チルト角を ϕ_i ，ズーム値を Z_i とし，ユーザ u の見たい人物の環境中での三次元位置を P_u として次のように定義する．なお本手法では，保育室内にいる子供の位置は既知としている．

$$S_{ui}(\theta_i, \phi_i, Z_i, P_u) = (W_1 \times G_{ui} + W_2 \times F_{ui}) \times Trans_{ui} \quad (1)$$

G_{ui} はカメラ i の画像がユーザ u の見たい人物の要求をどの程度満たしているかを表す関数であり， F_{ui} はカメラ i の画像がユーザ u の見たい人数の要求をどの程度満たしているかを表す関数である．また， W_1, W_2 は各項への重みである．そして， $Trans_{ui}$ はカメラが切り替わったときの違和感を軽減するために「イメージナリーライン」(後述)を考慮した撮影技法を適用して見やすい映像とするための係数である．以下に，評価関数 S_{ui} について詳述する．

(1) 見たい人物の要求に対する評価

式 (1) 中の見たい人物の要求に対する関数 G_{ui} は次のように定義する．

$$G_{ui}(\theta_i, \phi_i, Z_i, P_u) = k_1 \times \left(\frac{d_{ui}^2 - (w/2)^2}{(w/2)^2} \right) + k_2 \times D_{ui} + k_3 \times O_{ui} \quad (2)$$

k_1, k_2, k_3 は各項への重みである． d_{ui} は画像座標上において画像中心からユーザ u の見たい人物までの距離， w は画像の幅である．これは人物の三次元位置をカメラ i の画像上に投影して計算する． k_1 の項はユーザ u が見たい人物の画像内での位置を評価しており，ユーザ u の見たい人物が画像中心に近いほど値が高くなるように設定した．また， k_2 の項の D_{ui} はカメラ i とユーザ u の見たい人物との距離を評価し，この距離が近い方が値が高くなる． D_{ui} を以下のように定義する．

$$D_{ui} = 1.0 - \frac{\text{カメラ } i \text{ と人物とのユークリッド距離}}{\text{部屋の対角線の長さ}} \quad (3)$$

k_3 の項の O_{ui} はユーザ u の指定する見たい人物がオクルージョンとなっているかを考慮したもので，他の人物によってオクルージョンとなったときに値が低くなる． O_{ui} は以下のように設定する．

$$O_{ui} = 1.0 - \left(\frac{\text{画像上で指定人物と他人の重なる幅}}{\text{画像上に投影した指定人物の幅}} \right)^2 \quad (4)$$

(2) 見たい人数の要求に対する評価

式 (1) 中の見たい人数の要求に対する関数 F_{ui} は以下のように定義する．

$$F_{ui}(\theta_i, \phi_i, Z_i, P_u) = e^{-\log^2 \frac{Z_{ui}}{Z_i}} \quad (5)$$

Z_{ui} はユーザ u のカメラ i における理想のズーム値であり，ユーザ u の希望人数が画像内に納まるように画角を決定したときの焦点距離である．この画角はユーザ u の見たい人物のいる方向を理想の方向とし，見たい人物の周りにはいる人物のうち近い順に希望人数に含めて，この希望人数が画像内に入るように決定する． Z_{ui} と現在のカメラのズーム値 Z_i が等しくなるとき， F_{ui} は最大値となる．

(3) イマジナリーラインによる評価

見やすい映像とするための係数 $Trans_{ui}$ は「イマジナリーライン」を考慮した撮影技法を適用した遷移係数である．「イマジナリーライン」とは主要な二人の人物を結んだ線のことである（図 7）．図 7 中のカメラ 1 からカメラ A への切り替えのように，カメラを切り替える際にこの線を越えた位置のカメラに切り替えると人の位置関係や動きの方向が逆になり，見ている人に映っている人物の位置関係を分からなくさせてしまう．そのため，イマジナリーラインを越えた撮影はなるべくしない方が良いという撮影技法がよく知られている [21] ．

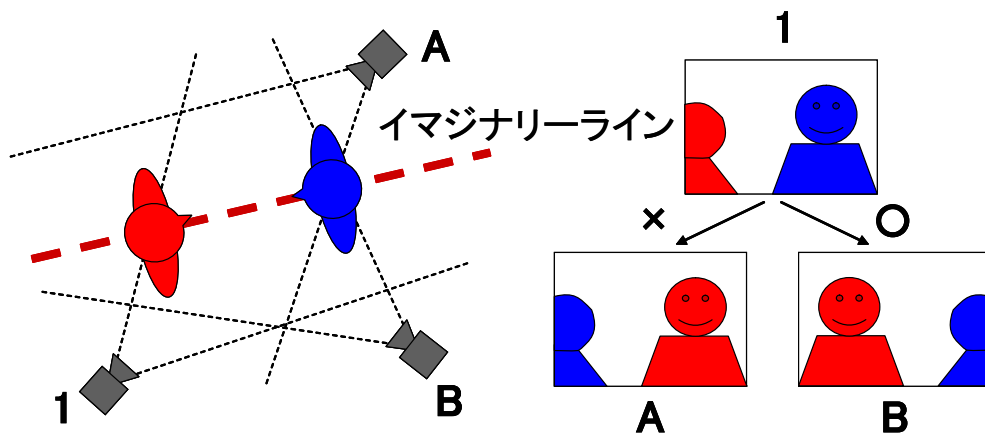
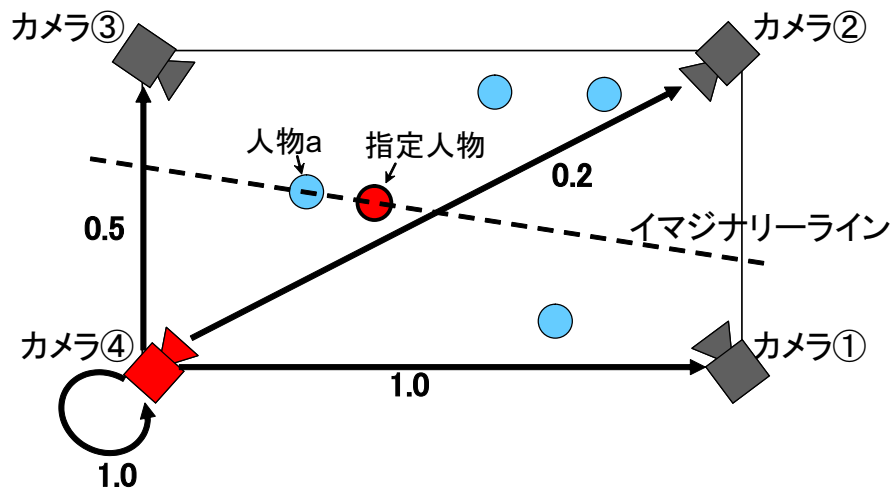
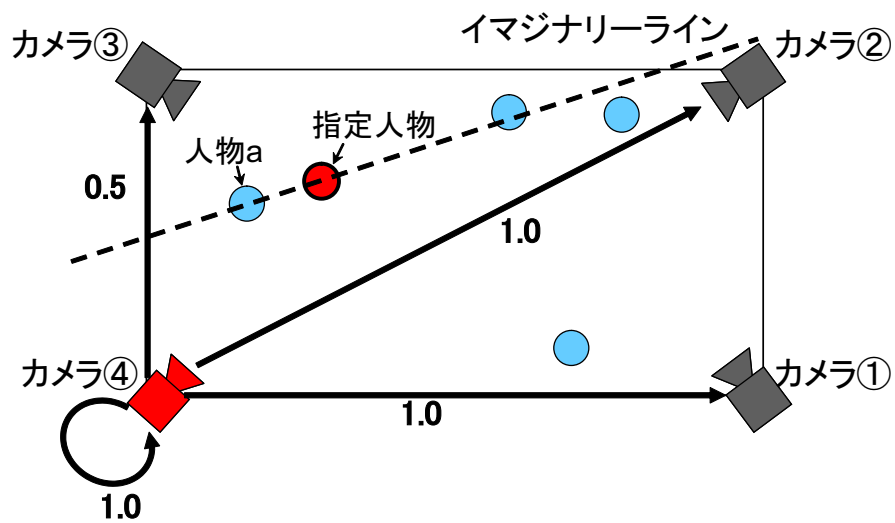


図 7 イマジナリーライン

そこで、あるカメラの映像を見ているユーザの、次に切り替わるカメラへの遷移係数を評価関数に影響させることで、ユーザの映像の見辛さを軽減する。本研究では、ユーザ u の指定した見たい人物と、その人物に最も近い人物を結ぶ線をユーザ u にとってのイマジナリーラインと考え、カメラ間の遷移係数はこのイマジナリーラインを越えるか越えないかで決定する。遷移係数の決定の例を、環境を上から見た図 8 に示す。図 8(a) は、ユーザにカメラ④の映像を提示していたとき、指定人物と最も近い人物 a とを結ぶ線をイマジナリーラインとして遷移係数を決定する様子を表す。カメラ①と④はイマジナリーラインを越えない位置にあるので、遷移係数は高く設定する。カメラ③はイマジナリーラインを越えているので遷移係数は低く設定し、カメラ②はイマジナリーラインを越えて対角のカメラなので、人物の位置や方向は真逆になるため、ほとんど切り替わらないよう低く設定する。また、図 8(b) のような場合では、カメラ②は対角にあるカメラであるが、イマジナリーラインを越えないため、高い値を設定する。遷移係数は、新規割り当ての場合はどのカメラも 1.0 であり、図中の数値は経験的に決定した。



(a)



(b)

図 8 イマジナリーラインによるカメラ間の遷移係数の決定

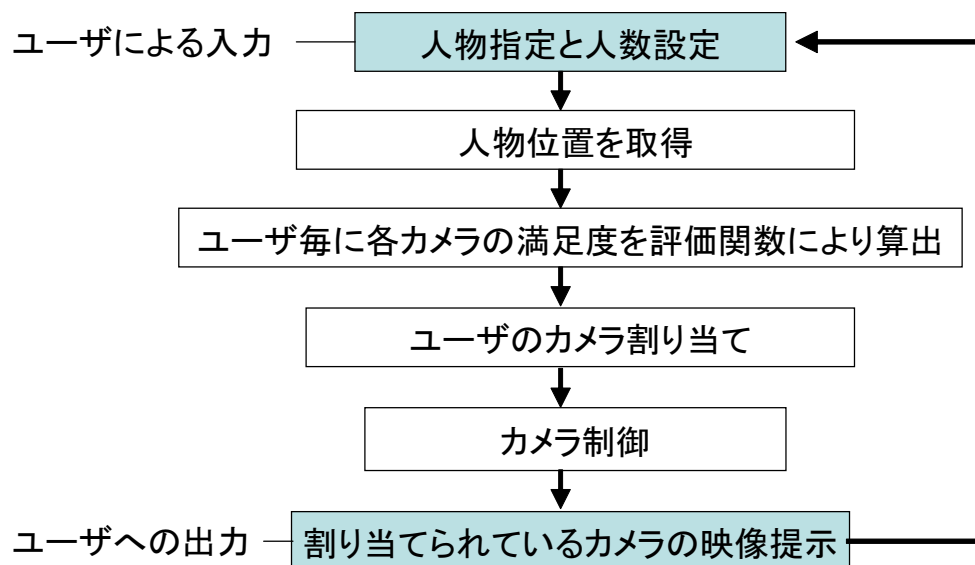


図 9 カメラ割り当て・制御の処理の流れ

2.2 評価関数に基づくカメラの割り当てと制御

本研究では、部屋に複数台の PTZ カメラを設置した状況で、室内に存在する人物を複数のユーザが見ることを想定する。また、ユーザはシステムに対する要求として、「見たい人物」と「見たい人数」を指定する。以上の状況においてユーザの要求に基づき、2.1.2 節で定義した評価関数を用い、カメラ制御を各ユーザに対するカメラの割り当て（2.2.2 節）と、各カメラのパラメータ（パン・チルト・ズーム値）の決定（2.2.3 節）の 2 段階の処理によって行う。これらの具体的な手法について以下で説明する。

2.2.1 概要

図 9 に本手法におけるカメラ制御の処理の流れを示す。本手法ではユーザにより人物指定と希望人数の指定が行われ、室内に存在する全人物の環境中の三次元位置は何らかの方法で取得できるものとする。現在のカメラの状態から、ユーザ毎に全カメラ画像の満足度を評価関数により算出し、満足度の高いカメラを各

ユーザに割り当てる．全ユーザへのカメラ割り当て後，カメラ毎にユーザの満足度を最大とするようにカメラパラメータを最適化してカメラ制御を行う．この処理を繰り返し，ユーザには割り当てられているカメラの映像を提示する．

2.2.2 各ユーザへのカメラの割り当て

各ユーザへのカメラの割り当ては，2.1.2 節で定義した評価関数により，現在のカメラパラメータの状態ですべてのユーザ毎に各カメラの満足度を求め，満足度が最大となるカメラをユーザに割り当てる．初期状態においてカメラ①の満足度が最大であるとき，このユーザにはカメラ①を割り当てる．そして指定人物の移動などにより各カメラの満足度が変化し，カメラ②の満足度が最大となったとき，先に割り当てられていたカメラ①の満足度との差を比較し，閾値以上であればカメラの割り当てを変更する．閾値を用いることにより，カメラの切り替わりが細かくなって見辛い映像となることを抑える．さらに，カメラの割り当てを変更してから3秒間は他のカメラに変更しない．このようにして全てのユーザにカメラを割り当てる．

2.2.3 評価値の最大化によるカメラ制御

カメラを全てのユーザに割り当てた後，カメラ毎に，割り当てられているユーザの満足度の総和を最大とするようにカメラパラメータ（パン・チルト・ズーム値）を決定する．カメラパラメータは2.1.2 節で示した評価関数をパン・チルト・ズームの各パラメータで偏微分した以下の式を用いて，最大勾配法により最適解を探索することで決定する．この際，最大勾配法の初期値として直前のカメラのパン・チルト・ズーム値を利用する．

$$\text{パン値の最適化：} \quad \frac{\partial S_{ui}}{\partial \theta_i} \quad (6)$$

$$\text{チルト値の最適化：} \quad \frac{\partial S_{ui}}{\partial \phi_i} \quad (7)$$

$$\text{ズーム値の最適化：} \quad \frac{\partial S_{ui}}{\partial Z_i} \quad (8)$$

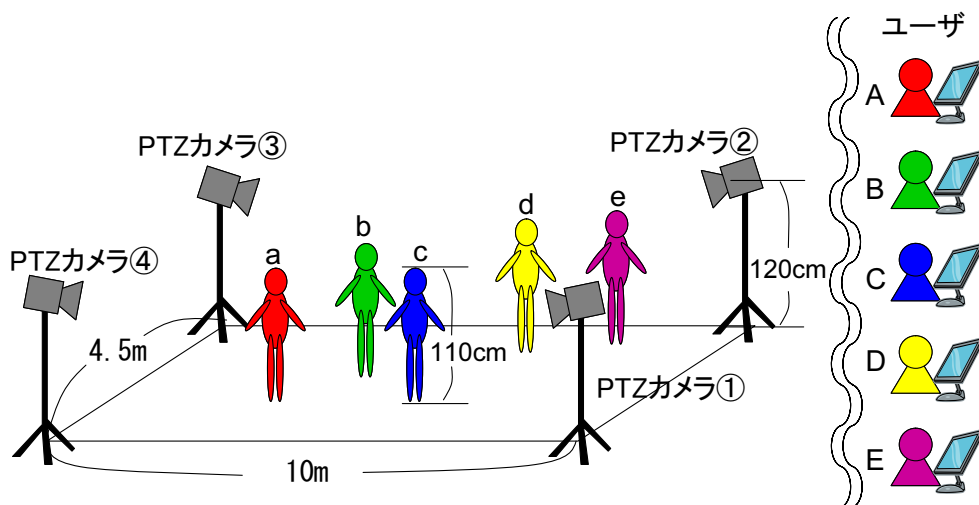


図 10 シミュレーション実験での実験環境

これにより求めたパン・チルト・ズームの最適解をカメラの制御パラメータとし、カメラ制御を行う。また、どのユーザにも割り当てられなかったカメラは、そのカメラに最も近い人物を指定するユーザにとっての理想のカメラパラメータを適用して、誰かに割り当てられるのを待つ。

2.3 シミュレーションによるカメラ制御実験

シミュレーション環境での実験により、提案手法に合ったカメラ制御ができていないかを確認する。また、限られたカメラ台数で何人のユーザに対応できるのかをユーザ数を変化させて考察する。各ユーザに提示される画像中の人物の映り方を、人物のCGモデルを用いて定量的に評価する。

2.3.1 実験条件

実験は図 10 に示すように 10m × 4.5m の空間の四隅に 1 台ずつ PTZ カメラ (計 4 台) を配置し、その空間内には複数の人物がおり、それぞれを見ているユーザが

いることとする。PTZカメラは初め、部屋の各隅から部屋の内向き 45 度の方向に設置し、制御パラメータの初期値は

- パン値：0 度
- チルト値：0 度
- ズーム値：最大水平画角 (48.8 度) における焦点距離

のように設定した。室内の人物の数とユーザ数は表 4 のように変更して実験を行った。実験はユーザの希望人数が全員 1 人である場合 (I) と、ばらつきのある場合 (II) の 2 パターンずつ行った。希望人数にばらつきのある実験 II では、ユーザ A の希望人数が 1 人であり、他はアルファベット順に 1 人ずつ希望人数を増やしている。以下では、各実験のことを「ユーザ 5 人の実験 I」を「実験 I-5」のように表記する。ユーザの指定する人物はあらかじめ与えているものとし、例えば図 10 に示すように、ユーザ A,B,C,D,E はそれぞれに対応する人物 a,b,c,d,e を指定して見ているものとする。人物の CG モデルは、子供の身長に合わせて身長 110cm とし、カメラの高さは床から 120cm の位置にあることとする。また、人物の位置は 3.3 節において述べる手法を用いて、実環境の実験で子供が遊んでいる時に取得したデータを利用した。子供が 5 人以上の場合は、取得時間の異なるデータを組み合わせて、人数分の位置データとした。なお、実験時の式 (1), (2) における各項の重み係数はカメラ台数に関係なく、経験的に W_1, W_2 は 0.6, 0.4, $Trans_{ui}$ は 2.1 節 2.1.2 項で説明したように子供の位置に関して値が変化する係数となっており、図 8 に示す数値を利用した。 k_1, k_2, k_3 は 0.7, 0.15, 0.15 とした。

2.3.2 提示画像の定量的評価

各ユーザに提示される映像にどのように人物が映っているかを定量的に評価した。各ユーザが指定する人物の、画像上での画像中心からの距離、画像上で占める割合と、それぞれの画像に何人が映っているかを評価した。本システムは 10 名のユーザにまで対応できることを目標とするので、ユーザ数 10 名で行った実験結果を中心に評価を行う。その他の実験結果の詳細については付録を参照されたい。

表 4 実験パターン

実験 ユーザ数(人)	I	II
5	全ユーザ 希望人数 一人	希望人数 ばらつき 有り
10		
15		

(1) 提示画像上での指定人物の位置とカメラの割り当て変更

実験 I-10 において、各ユーザに提示される画像中に指定人物が画像中でどの位置に映っているかを評価する。各ユーザの指定人物が画像中心からどのくらい離れているかの割合を図 11 に示す。図 11 は画像中心と人物の画像への投影点間の距離を算出した結果であり、グラフ中の縦軸は 0 に近いほど画像中心に近く、 $w/2$ に近いほど画像端に映っていることを表す。また、その時割り当てられているカメラを、図下に示すように番号に対応する色でグラフ中表示する。

約 4 秒まではカメラ③が割り当てられているユーザは E のみであり、画像中心に捉えられていることから、指定人物を画像中心に捉えるようにカメラ制御されていることが分かる。また、ユーザ B, D のグラフを見てみると、二人のユーザはほとんど同じカメラが割り当てられている。この二人のグラフはほぼ同じ形状をしており、人物 b, d の二人が左右対称に映っていると分かり、カメラ制御が同じカメラを割り当てられた全ユーザの満足度を高くするように行われていることが分かる。全ユーザの割り当てられるカメラが変更されたときのデータの変化を見ると、前に割り当てられていたカメラよりも画像中心に指定人物を捉えることができていることが読み取れる。このことより、満足度が高くなるようなカメラが割り当て直されていると考えられる。しかしほとんどのユーザについて、指定人物が映らない時間が 3 秒程度ある。この理由として、全員が一人希望であり、複数人に割り当てられたカメラは、その全てのユーザの指定人物を映そうとしながら、それぞれの人物については一人を捉えるようなズーム制御をしようとするために、カメラパラメータの最適化の際に局所解に陥っていることが考えられる。

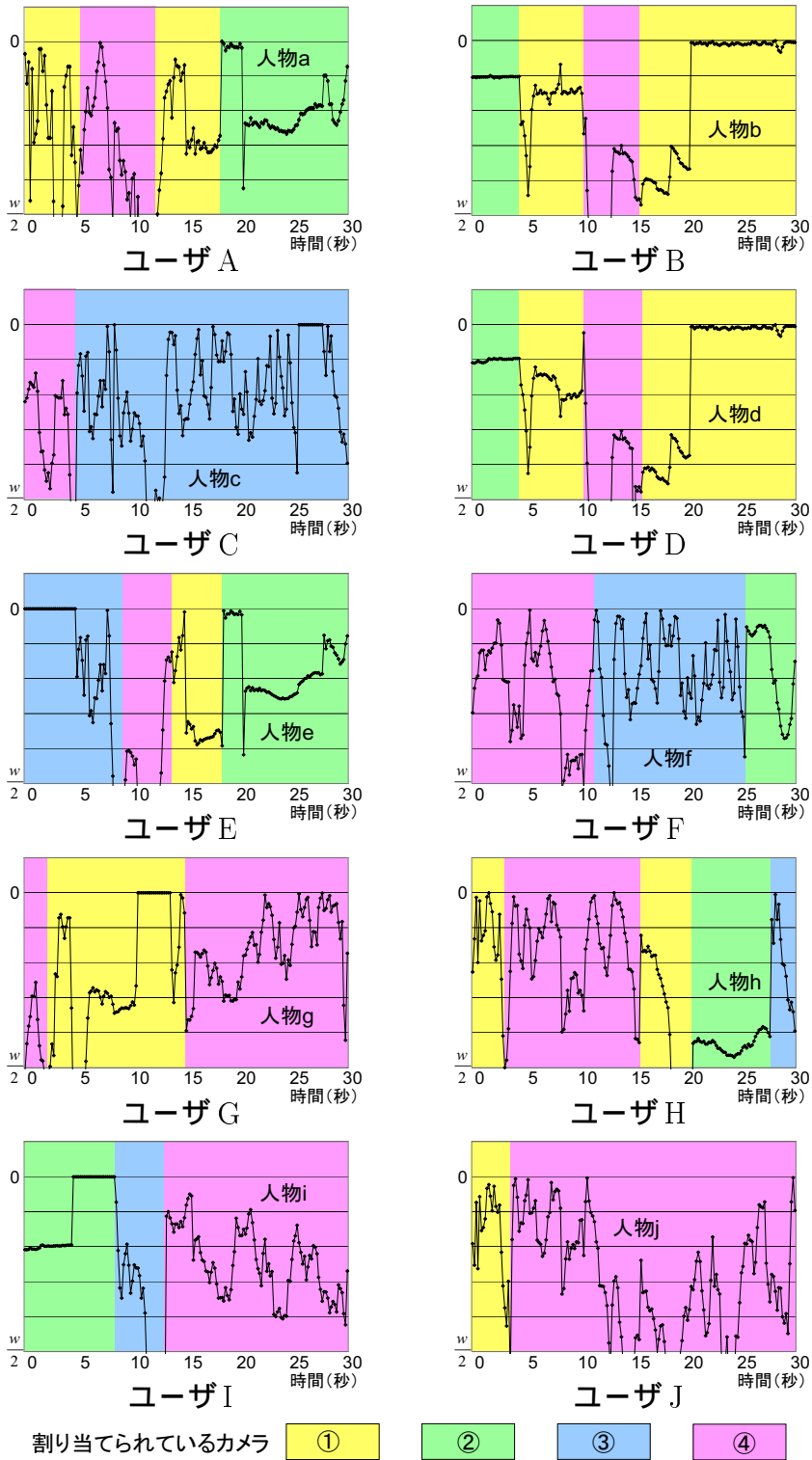


図 11 各ユーザへの提示画像における指定人物の画像中心からの距離

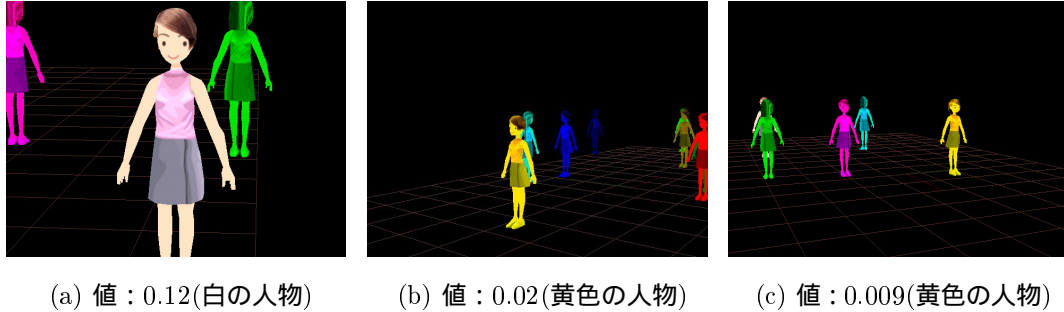


図 12 人物の映る大きさの例

表 5 各ユーザへの提示画像上で指定人物の画像中で占める割合

		A	B	C	D	E	F	G	H	I	J
5 人	I	0.084	0.067	0.076	0.063	0.063					
	II	0.107	0.099	0.036	0.008	0.009					
10 人	I	0.034	0.057	0.050	0.051	0.046	0.047	0.065	0.048	0.039	0.045
	II	0.028	0.024	0.027	0.021	0.023	0.021	0.022	0.023	0.008	0.009

(2) 提示画像上で指定人物の占める割合

次に、5人で行った実験と、10人で行った実験のそれぞれについて、各ユーザに提示される画像中で指定人物の占める割合の約2分間の平均を表5に示す。

提示画像中で指定人物の占める割合は、ユーザに提示される画像に映る指定人物領域の画素数と全画素数(640×480pixels)から求めている。このグラフにおける人物の映る大きさの例を図12に示す。図12(a)の正面を向いている人物が画像上で占める割合は0.12であり、図12(b)の手前の黄色い人物が画像上で占める割合は0.02、図12(c)の一番右に映っている人物の画像上で占める割合は0.009である。実験IIでは、希望人数はユーザのアルファベット順に1人から1人ずつ増えている。表5実験IIより、希望人数に合わせて人物を映す大きさが変わっていることが分かる。ただし、実験II-10では、ユーザAからHまで人物を捉える大きさに大きな変化が見られない。これは、同じカメラを割り当てられるユーザが多数存在し、それぞれの指定する人物同士が離れていることが原因であると考

表 6 各ユーザへの提示画像における映る人数

		A	B	C	D	E	F	G	H	I	J
5人	I	2.2	2.1	2.3	2.3	2.5					
	II	1.6	2.7	3.3	3.7	4.1					
10人	I	5.1	4.0	4.7	5.6	4.5	4.4	4.7	5.3	4.4	4.2
	II	5.7	6.0	6.0	6.1	6.1	6.8	7.7	7.3	6.5	7.6

表 7 システムが算出した各ユーザの満足度

		A	B	C	D	E	F	G	H	I	J
5人	I	0.824	0.818	0.841	0.785	0.788					
	II	0.906	0.895	0.852	0.857	0.894					
10人	I	0.573	0.697	0.726	0.685	0.653	0.676	0.732	0.681	0.618	0.657
	II	0.561	0.687	0.735	0.751	0.783	0.820	0.817	0.743	0.802	0.802



図 13 実験 I-10 でのユーザ H への提示画像

えられる。しかし、図 12(b) に示す手前の黄色い人物のような大きさに映っており、5 人を希望するユーザに関して言えば、実験 II-5 における 5 人希望のユーザの指定人物よりも、大きく捉えることができている。これは室内の人数が増えることで人物同士が近くなり、狭い画角でも希望の人数を含めることができるためである。

(3) 提示画像に映る人数

次に、(2) と同じ時間において、ユーザへの提示画像に映っている人数の平均を表 6 に、その時のシステムが算出する各ユーザの満足度の平均を表 7 に示す。

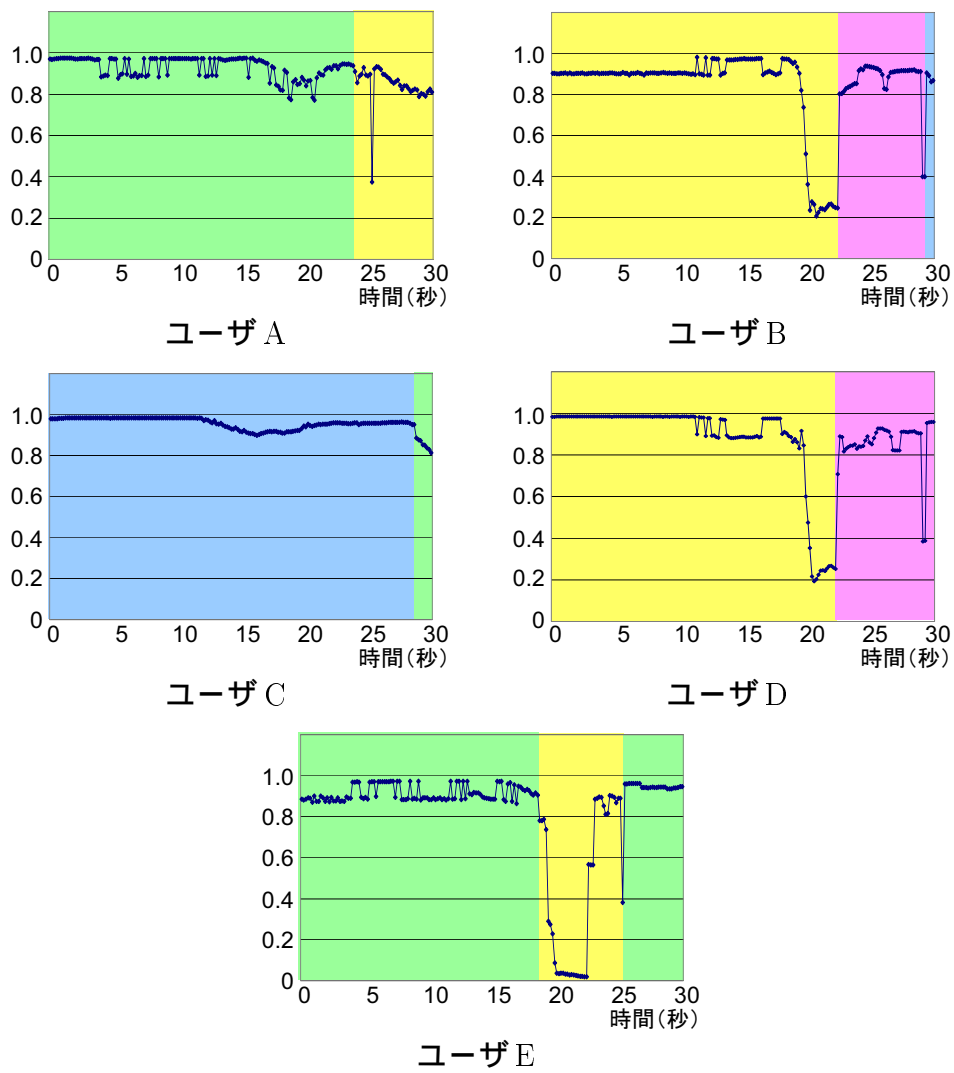
実験 II-5 では、希望人数に応じたカメラ制御ができていることがわかる。しかし、10 人のユーザでの実験では、I、II とともに希望人数の少ないユーザは希望よりも多く映っている。同じカメラが割り当てられる複数人の満足度を高くするようにカメラ制御を行うため、希望人数が 1 人や 2 人のユーザも、他のユーザと同じように他の人物を映すように制御されることが 1 つの原因である。図 13 は、実験 I-10 における黄色の人物を指定するユーザ H への提示画像である。画像に映る人数は希望よりも多いが、この図のように比較的大きく指定人物を捉えることができている。映る人数が希望より多くなることは、指定人物のうしろの人物が映り込むことも原因の 1 つである。このときは指定人物を大きく捉えるように制御できているので、カメラの制御としては正しい制御ができていると考えられる。しかし、これらの理由から、希望人数の少ないユーザについては、満足度があまり上がっていないことが表 7 より分かる。

(4) 満足度の変化とカメラ割り当ての変更

実験 I-5 におけるユーザの満足度の変化とカメラの割り当て変更の関係を表すグラフを図 14 に示す。ユーザ A と E、B と D のグラフを見ると、同じカメラが割り当てられ、二人の満足度が同時に低くなったとき、カメラの割り当てが変更され、満足度を高くするように処理されていることが分かる。他のユーザが割り当てられるカメラと同じカメラを割り当てられたときは、満足度が落ちるところも見られる。

(5) ユーザ 15 名での実験との比較

次に、実験 II-15 の画像中で指定人物が占める割合の結果を表 8 に示す。表 8 よ



割り当てられているカメラ ① ② ③ ④

図 14 満足度の変化とカメラの割り当て変更

表 8 各ユーザへの提示画像上で指定人物の占める割合 (実験 II-15)

	A	B	C	D	E	F	G	H
15人(Ⅱ)	0.030	0.013	0.026	0.022	0.020	0.020	0.020	0.015
	I	J	K	L	M	N	O	
15人(Ⅱ)	0.013	0.018	0.008	0.025	0.023	0.015	0.014	

り、人物の映る大きさは希望人数にあまり関係が見られない。ユーザ数が 15 人まで増えると、同じカメラを多くのユーザに割り当てるために、各ユーザの希望人数の要求に応じて映すことが難しくなっていることが分かる。

これらの結果より、ユーザが 10 人まで増えた場合、希望人数の少ないユーザの要求を大きくは満たせないものの、10 人までのユーザ数に対しては、要求に応じたカメラ制御が可能であると考えられる。

3. 複数の PTZ カメラによるライブカメラシステム

本章では、複数の PTZ カメラを使用し、各ユーザに適したカメラの映像を提示するライブカメラシステムについて述べる。

3.1 システムの設計方針と概要

本研究では、保育室にいる子供の位置情報から、ユーザの要求に応じたカメラ制御をしているカメラの映像を提示するシステムの実現を目的とする。このような目的を実現するためには、

1. 部屋にいる子供の位置情報の取得
2. ユーザの要求の入力
3. 1, 2 の情報からの各カメラの制御パラメータの決定とカメラ制御
4. カメラ画像の配信と受信

の処理を行い、図 15 のようなシステムを構築する必要がある。本研究では、そのプロトタイプシステムとして図 16 に示す構成のシステムを実現した。本システムでは、室内の子供の位置情報を魚眼カメラの画像より取得した。そして、ユーザの入力した希望人数の要求と合わせて、サーバにおいて 2 章で提案した手法を適用し、ユーザにカメラを割り当てて各カメラの制御パラメータを決定した。その決定したカメラの制御データを PTZ カメラに繋がる各 PC に送信し、制御データに従ってカメラを制御した。PTZ カメラに繋がる各 PC は、そのカメラ画像を配信した。ユーザ側 PC では、割り当てられたカメラ番号の情報から、そのカメラの画像を受信した。各 PC 間のデータの送受信はネットワークを介して行った。図 15 におけるサーバの役割は、プロトタイプシステムでは PTZ カメラに繋がる一台の PC で行った。保育室の音声は、マイクを PTZ カメラに繋がる一台の PC に接続して、Windows Media エンコーダを利用して遠隔地へ配信した。

さらに、ライブカメラサービスにアクセスするユーザは PC を使い慣れているとは限らないので、ユーザが簡単に操作可能なインタフェースを実装した。

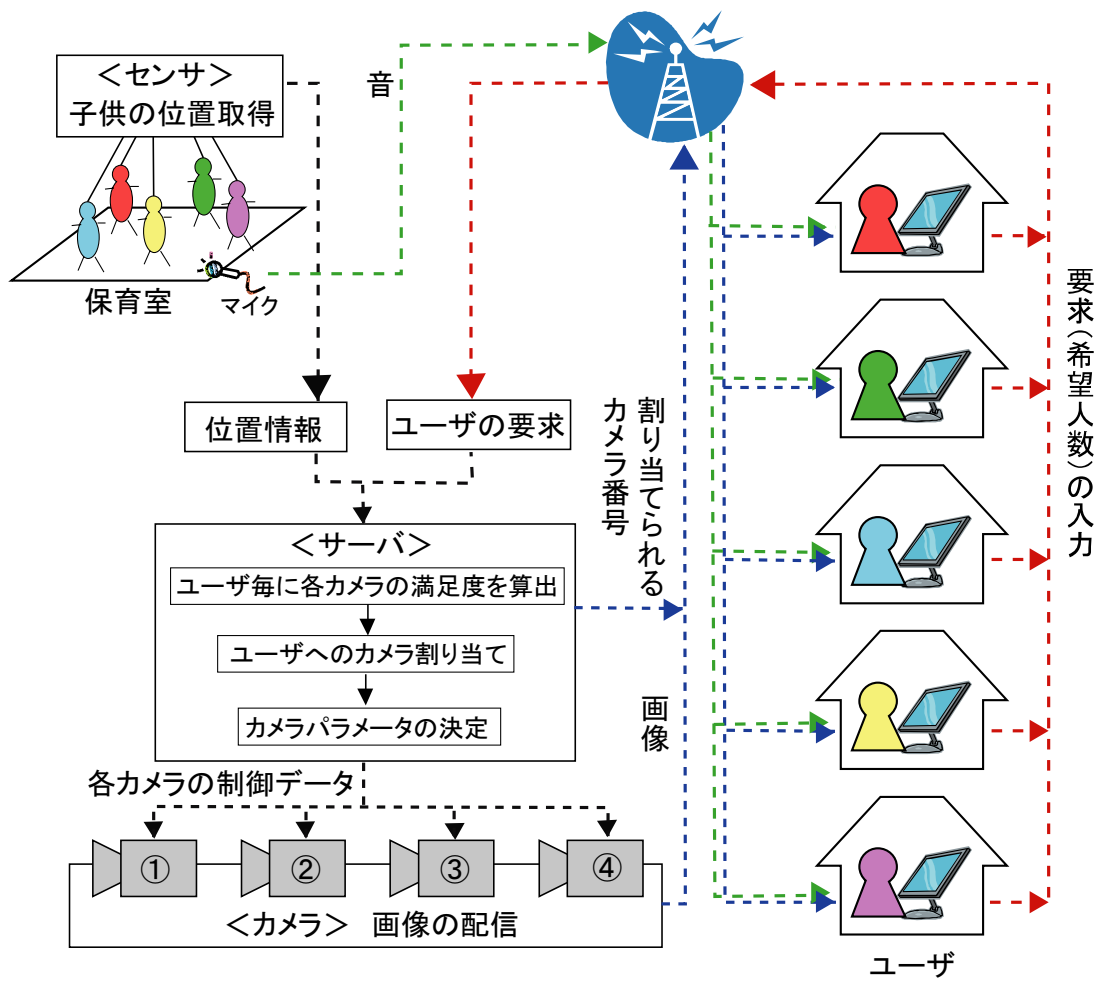


図 15 システムの処理の流れ

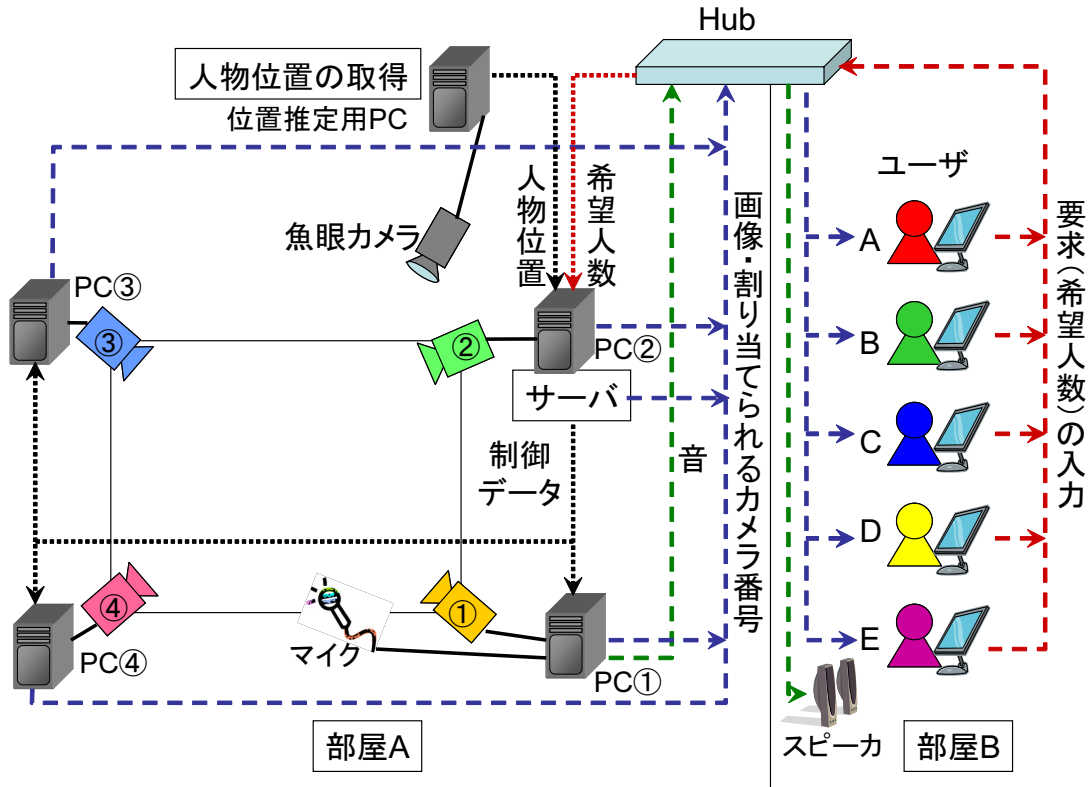


図 16 プロトタイプシステムの構成

表 9 PC のスペック

PC	CPU	メモリ
PC①	CoreDuo 1.83GHz	1.0GB
PC②	PentiumD 3.2GHz	2.0GB
PC③	Pentium4 3.0GHz	1.0GB
PC④	Pentium4 3.2GHz	2.0GB
位置推定用PC	Athlon64X 2.2GHz	2.0GB
ユーザ用PC 5台	PentiumM 2.0GHz	1.0GB

3.2 機器構成

プロトタイプシステムにおける機器構成について述べる．図 16 中の各 PC の仕様を表 9 に示す．全ての PC は Gigabit Ethernet に接続した．魚眼カメラには POINT GREY RESERCH 社製の Flea(解像度：1024 × 768pixels) を使用した．PTZ カメラは以下の仕様の SONY 社製 EVI-D30 を使用した．

- 解像度：640 × 480pixels
- パン範囲：±100 度
- チルト範囲：±25 度
- 水平画角：4.4 度 ~ 48.8 度

3.3 観測人物の位置推定

本研究におけるプロトタイプシステムでは，室内にいる人物の位置推定を PTZ カメラとは別に魚眼カメラを使用して行う．カメラ画像上での人物の追跡は CamShift アルゴリズム [22] を利用する．CamShift アルゴリズムは次に示すような手順で追跡を行う手法である．

1. 基準フレーム（初期フレーム）において追跡したい領域を矩形で指示．
2. 以下の方法で追跡用画像を作成．
 - (a) 基準フレームの原画像を HSV 表色系に変換する．
このとき HSV の値域は $[0, 255]$ とする．
 - (b) 1. で指示した領域中において H（色相）のヒストグラムを求める．
 - (c) 最も頻度の高い H の値を 255 とし，他の H についても頻度の割合に応じて，以下の式により値を正規化する．
 L_h : H の値が h である画素に対応する追跡用画像の仮の値

F_h : H の値が h である画素の指示領域中における頻度
 h : 値域 $[0,255]$. H の値

$$L_h = F_h / \text{Max}(F) * 255 \quad (9)$$

(d) さらに H の頻度を表す追跡用画像を出力するために , S と V に最低の閾値 (Th_s, Th_v) を固定値として与える .

O_h : H の値が h である画素に対応する追跡用画像の値

$$O_h = \begin{cases} L_h & S > Th_s \text{ and } V > Th_v \\ 0 & \text{otherwise} \end{cases} \quad (10)$$

3. (9) (10) 式に従って値域 $[0,255]$ の追跡用画像を生成 .
4. 全てのフレームにおいて追跡用画像を生成 .
5. 追跡用画像から meanshift 法 (値域が $[0,255]$ の輝度の山を探索する手法) を用いて探索 .
6. フレームごとの矩形領域情報を得る .

本システムでは , 観測人物に色の異なる衣類を身につけてもらい , 基準フレームにおける矩形領域の指示は , 観測人物の衣類を選択する . この手法により得られた画像上での人物の追跡結果の座標から , カメラ中心とカメラ画像上の座標を通る世界座標系における空間直線を求め , その直線が床から 70cm の位置に交差する点を子供の位置として計算する .

このような方法による位置推定では , 魚眼カメラから離れるほど推定位置と実際に子供のいる位置との間に誤差が生じる . この位置推定の精度は , カメラから 2m 離れた場所にいる子供の位置は約 10cm の誤差 , カメラから 10m 離れた場所にいる子供の位置は , 約 40cm の誤差であった . また , 「床から 70cm の位置を子供の位置とする」としたことにより , 子供が座ったり寝転んだりしたことで推定位置に誤りが生じる . カメラから約 7.5m 離れた位置で寝転んでいる子供の推定位置は , 実際の位置より , 約 1m カメラ寄りの位置であった .



図 17 ユーザのインタフェース

3.4 ユーザの要求入力インタフェース

ユーザの使用するシステムのインタフェースは、ユーザがPCを使い慣れているとは限らないため、操作が簡単であることが必要である。そのため、できる限り簡単に要求を入力できるインタフェースとした。

まず、ユーザの要求は事前に行ったアンケートの結果より「見たい子供」と「見たい人数」である。このアンケートの中では、自分の子供が映っていないシーンについて、全体の1割程度と低かったので「見たい子供」は自分の子供であると考えられる。そこで、本システムではユーザがアクセスした時にユーザと子供の対応が取れるものとする。そして「見たい人数」の入力はスライダーを操作するだけの簡単な入力とした。これはスライダーを操作するだけの簡単な入力とする。ユーザへのインタフェースを図17に示す。図17中の右側にあるスライダーにより「見たい人数」を入力し、上に上げることで希望人数が一人に近づき、

自分の子供を大きく映して欲しいという要求を入力したこととなる。周りの子供を含めて映したいときには、下に下げることによって希望人数を増やすことができる。システムにアクセスしたときの初期状態におけるユーザの希望人数は、なるべく大きく子供を捉え、かつ映る可能性が高くなるように、2人としている。また、図17に示すように、カメラが映そうとしている人数と、入力した希望人数との差を表示する。これは、要求通りにカメラ制御が行われないときに発生すると考えられるユーザのストレスを軽減するための工夫である。なぜなら、同じカメラを割り当てられる全てのユーザの満足度を最大とするようにカメラ制御を行うため、ユーザが希望人数の変更を行っても、その通りの制御が行われるとは限らないためである。また、割り当てられているカメラ映像の解像度はVGA(640×480pixels)で提示される。

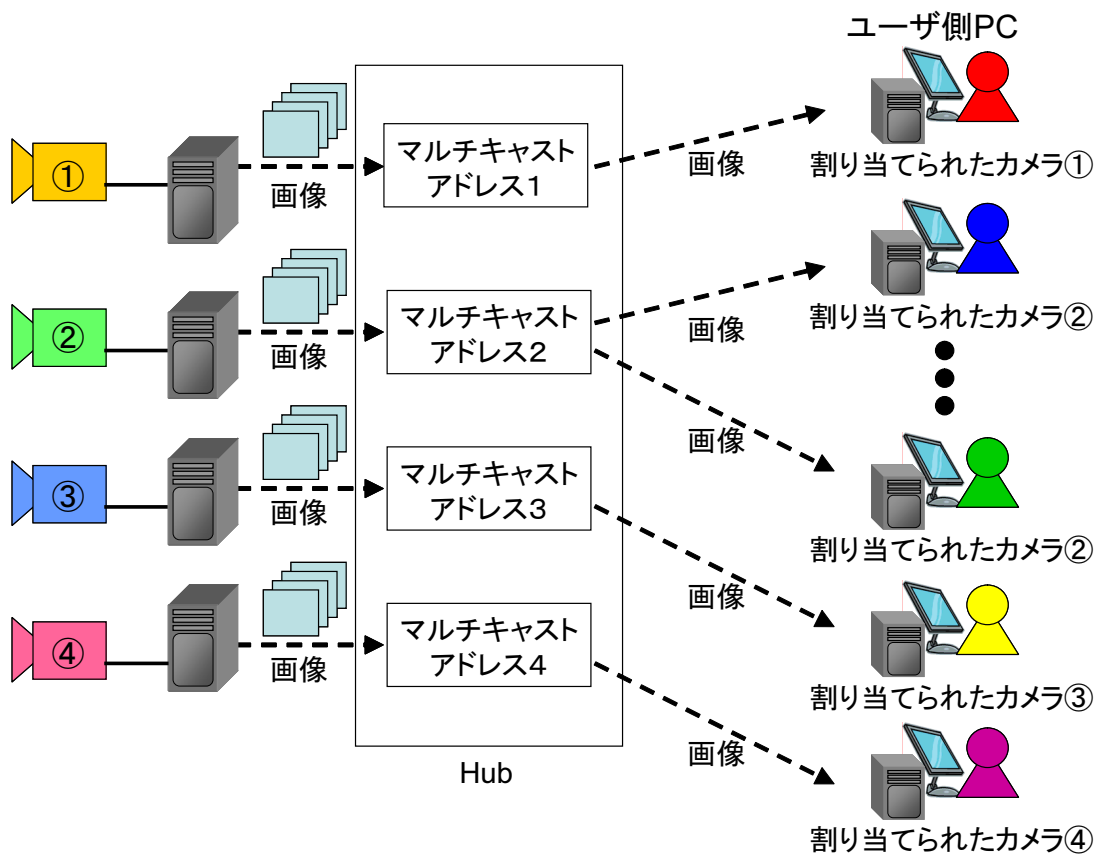


図 18 カメラ映像の配信・受信

3.5 ユーザの要求に応じた映像の配信と受信

ユーザ側では、割り当てられているカメラの映像のみを提示する。これは全てのPTZカメラから映像をリアルタイムに配信し、ユーザ側のPCで割り当てられるカメラの映像を取得することで実現した。この際、各PTZカメラから得られた動画像をマルチキャストプロトコルを用いて伝送することにより、カメラ台数以上の複数のユーザが配信される動画像を受信しても、ネットワークの負荷は増えることはない。カメラ1台あたりの画像を送信するためのビットレートを $N[\text{bps}]$ とし、その画像を受信するユーザ数を S 人とする、ユニキャストを利用した場合は、データ送信の処理負荷は1台のカメラについて $NS[\text{bps}]$ となる。これに

対し、マルチキャストを用いることで、処理負荷は1台のカメラについて $N[\text{bps}]$ のみとなる。映像の配信と受信の様子を図 18 に示す。ユーザは図のように、それぞれ割り当てられたカメラ番号のカメラ画像を受信する。各カメラを接続された PC は、図のようにカメラの画像データをパケットに分割して、対応するマルチキャストアドレスに配信し続ける。カメラ番号とマルチキャストアドレスは対応付けられており、カメラ①ならばマルチキャストアドレス 1 に、カメラ②ならばマルチキャストアドレス 2 に、パケットを送信する。そして、Hub において同じカメラを割り当てられるユーザ数分のパケットを複製する。ユーザ側 PC には、サーバから割り当てられたカメラの番号が送られており、割り当てられたカメラ番号に対応するマルチキャストアドレスからそのカメラの画像データのパケットを受信する。カメラの切り替わりは、割り当てられたカメラ番号に変更があったときに、ユーザ側 PC で受信するマルチキャストアドレスを変更することで行う。

4. ライブカメラシステムの被験者実験

プロトタイプシステムを用いて実際の親子5組による評価実験を行った。本実験は、ユーザに提示される映像の満足度、システムの有用性を被験者により主観的に評価してもらうことを目的とする。そして、実験結果と幼稚園児を持つ実際の母親による主観評価の結果から考察を行い、本システムが有効であるかを検討する。

4.1 実験環境と条件

実験環境を図19に、魚眼カメラから捉えた実験室内の様子(実験前)を図20に示す。部屋の大きさ・PTZカメラの配置、また式(1)、(2)における各項の重み係数はシミュレーション実験と同様である。室内にいる子供の位置推定に用いる魚眼カメラは、部屋の一隅に約2.7mの高さで設置した。子供は幼稚園の年中・年長に通う5人であり、身長は110cm前後であった。実験中、子供達には位置推定のため、図20中で見られるような色の異なる服を着てもらい、1人の大人が付き添っている状態で、室内で自由に遊んでもらった。各子供の母親には、別室においてそれぞれのPCから本システムを使用して、子供のいる部屋の様子を見てもらった。子供の服の色と母親の利用するPCが対応するようにシステムの設定はあらかじめ行っている。希望人数の入力は、子供の人数のみで入力してもらい、付き添っている大人は人数に含めないよう説明した。実験は次のように3パターン行い、一回の実験時間は15分とした。

- 1回目 カメラ2台(図19中のカメラ①④使用)
- 2回目 カメラ3台(図19中のカメラ①②④使用)
- 3回目 カメラ4台(図19中のカメラ①～④使用)

各実験後には提示された映像の評価アンケートを行い、最後に、システムの評価アンケートを行った。それぞれのアンケート用紙を図21、22、23に示す。

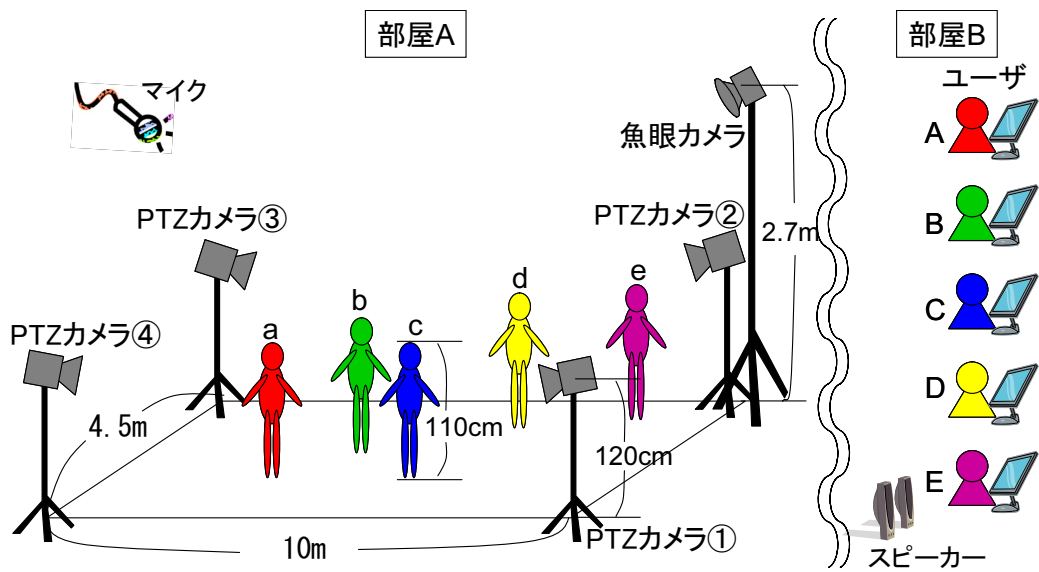


図 19 実際の環境での実験環境



図 20 室内の様子 (実験前)

ライブカメラシステムに関するアンケート

(ご氏名：)
(PCナンバー：)

次のそれぞれの質問に対して、5段階評価で当てはまる数字に○を付けて下さい。

■ 各実験についてのアンケート調査にご協力よろしくお願い致します。

1. お子様を常時映すことができていましたか？

1回目 (5：大満足 4：満足 3：普通 2：不満足 1：かなり不満足)
2回目 (5：大満足 4：満足 3：普通 2：不満足 1：かなり不満足)
3回目 (5：大満足 4：満足 3：普通 2：不満足 1：かなり不満足)

2. カメラは見たい方向に向けられていましたか？

1回目 (5：大満足 4：満足 3：普通 2：不満足 1：かなり不満足)
2回目 (5：大満足 4：満足 3：普通 2：不満足 1：かなり不満足)
3回目 (5：大満足 4：満足 3：普通 2：不満足 1：かなり不満足)

3. 希望通りの大きさでお子様を映すことができていましたか？

1回目 (5：大満足 4：満足 3：普通 2：不満足 1：かなり不満足)
2回目 (5：大満足 4：満足 3：普通 2：不満足 1：かなり不満足)
3回目 (5：大満足 4：満足 3：普通 2：不満足 1：かなり不満足)

4. お子様の様子を把握できましたか？

1回目 (5：良く把握できた 4：把握できた 3：普通
2：あまり把握できなかった 1：全く把握できなかった)
2回目 (5：良く把握できた 4：把握できた 3：普通
2：あまり把握できなかった 1：全く把握できなかった)
3回目 (5：良く把握できた 4：把握できた 3：普通
2：あまり把握できなかった 1：全く把握できなかった)

図 21 実験毎に行うアンケート用紙

■ 実験お疲れ様でした。

最後に、システムの全体的なアンケート調査にご協力よろしくお願い致します。
また、その評価を選んだ理由をなるべくご記入下さい。

1. 映す人数の希望の変更を行う操作は簡単でしたか？
(5：簡単！ 4：まあ簡単 3：普通 2：難しい 1：かなり難しい)

理由： _____

2. 本システムは同じカメラの映像を見る全ユーザの希望を満たすようにカメラ制御を行うため希望したズーム制御がそのまま反映されるとは限りません。そこで、ご自身が希望された映す人数と、カメラが映そうとしている人数との差を表示する機能を付けていましたが、これにより思い通りにカメラ制御が行われない時のストレスを軽減することができましたか？

(5：軽減できた 4：まあ軽減できた 3：分からない
2：あまり軽減できなかった 1：全く軽減できなかった)

理由： _____

3. 提示された映像における画質には満足できましたか？
(5：大満足 4：満足 3：普通 2：不満足 1：かなり不満足)

理由： _____

4. 映像は見やすいものでしたか？
(5：とても良い 4：良い 3：普通 2：悪い 1：非常に悪い)

理由： _____

5. 本システムでは、ご自身のお子様を中心に見るようになっていましたが、見たい子供を変更できる機能は必要ですか？カメラはクラスごとに付いているものだとします。

(5：必要！ 4：あれば嬉しい 3：どちらとも言えない
2：必要ない 1：全く必要ない)

理由： _____

図 22 事後アンケート用紙 1 枚目

6. 本システムでは、映す人数の希望を変更することによって、お子様を捉える大きさを変更するようなズーム制御をおこなっていました。この方法について満足されましたか？

また、他の方法でズーム制御を変更できれば良いと思われた方法があればご記入願います。

(5：大満足 4：満足 3：普通 2：不満足 1：かなり不満足)

理由：

ズーム制御の良いと思われる他の方法：

7. お子様に通っている保育施設にライブカメラサービスを導入してほしいと思いますか？

(5：是非導入してほしい 4：あれば嬉しい
3：どちらでも良い 2：あまり必要ない 1：導入しないでほしい)

理由：

8. お子様の通う保育施設にライブカメラサービスが導入されるとして、本システムは有効であると思われますか？

(5：有効！ 4：まあまあ有効 3：どちらとも言えない
2：あまり有効でない 1：有効でない)


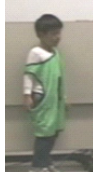



理由：

■ その他ご意見などありましたら、ご記入よろしくお願ひ致します。

本日は本当にご協力ありがとうございました。

図 23 事後アンケート用紙 2 枚目

表 10 各ユーザの子供

ユーザ	A	B	C	D	E
子供					

4.2 実験結果

3 回目の実験中，ある時間において各ユーザに提示された画像を図 24，25 に示す．これは約 30 秒間の各ユーザに提示される映像から，6 フレームを抜き出した画像である．画像左下にはユーザの入力した希望人数を，画像右下にはユーザに割り当てられたカメラ番号を表示してある．各ユーザの子供は表 10 に示す．

・ユーザ A と D の結果画像

ユーザ D は希望人数が少ないが，はじめは他のユーザも同じカメラを割り当てられているために，あまりズームインされていない．22 秒において同じカメラが割り当てられていたユーザ A が他のカメラを割り当てられたので，ユーザ D の子供を中心に捉えることができている．

・ユーザ C の結果画像

カメラ④がユーザ B に割り当てられる 26 秒まで，子供に近いカメラでユーザ C の子供を中心に大きく捉えることができている．

・ユーザ B と E の結果画像

17 秒において，同じカメラが二人のユーザに割り当てられている．ユーザ E は希望人数が 1 人であるが，この時ほとんどズームインされずに小さく映っている．これは，ユーザ B の希望人数が 4 人であり，この時 3 人は近くにいるが，4 人目の子供が離れているために，その子供を映そうとしているためであると考えられる．22 秒では，ユーザ B は希望人数を 1 人に変更し，17 秒の時に比べてズームインされている．しかし，二人の子供が離れていくため，これ以上はズームインされず，26 秒ではユーザ B に割り当てられるカメラが変更される．

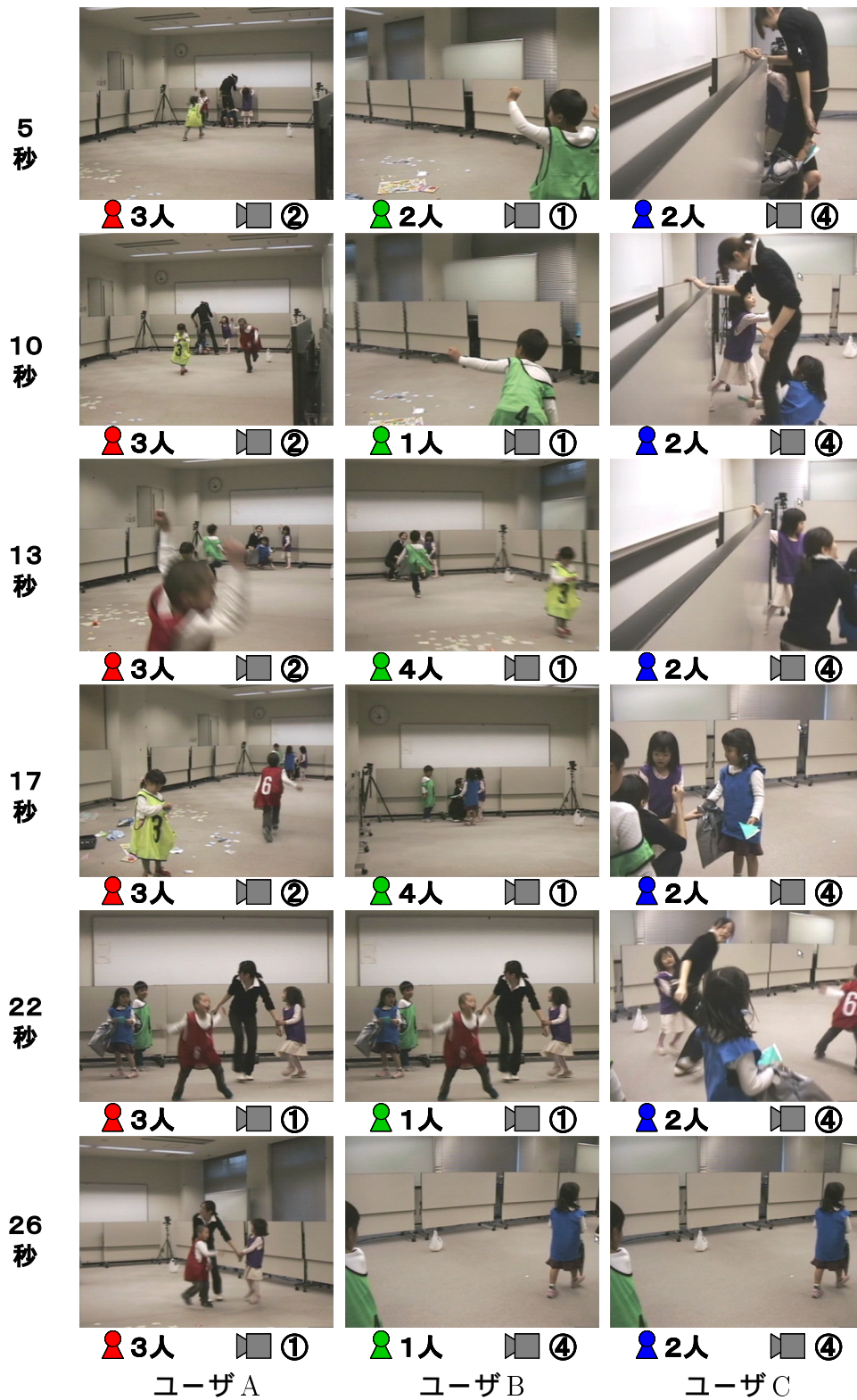


図 24 各ユーザへの提示画像 (A ~ C)



図 25 各ユーザへの提示画像 (D ~ E)

表 11 システムが算出した各ユーザの満足度の平均

	ユーザA	ユーザB	ユーザC	ユーザD	ユーザE
1回目	0.891	0.945	0.913	0.897	0.848
2回目	0.824	0.771	0.793	0.769	0.763
3回目	0.866	0.846	0.847	0.823	0.813

4.3 被験者による主観評価

実験は1回目がPTZカメラ2台，2回目がPTZカメラ3台，3回目がPTZカメラ4台で行った．システムが算出した各ユーザの満足度の実験時間中の平均の結果と，全てのアンケート結果を次のように提示する．

- 表 11 システムの算出した実験中の各ユーザの満足度の平均
- 表 12 提示映像に対する満足度を訊ねた評価アンケートの結果
- 表 13 システムの評価を行う事後アンケートの結果
- 表 14 事後アンケートに記入された主なコメントの結果

アンケート項目は全て，数値が高いほど質問に対する肯定的な回答である．

(1) 提示画像に対するユーザの満足度

表 12 から，カメラ2台では低い評価となり，ユーザの望むようなカメラ制御は行われず，カメラ3台以上では，平均3ポイント以上の高評価であったことから，ユーザの要求に応じた映像が提示されていたことが分かる．しかし，この結果に反して，システムが算出するユーザの満足度はカメラ2台で実験を行った時に最も高かったことが表 11 より分かる．これについては4.4節で考察を行う．

(2) システムの評価

システムの評価としては，平均3ポイント以上の高評価であり，特に操作方法や画質，映像の見やすさについては高い評価が得られた．しかし，同じカメラを割り当てられる全ユーザの満足度を高くするようにカメラ制御が行われるため，希望人数の入力をして大きく反映されないこともあり，この時にストレスを感

表 12 提示映像に対する評価結果

質問内容	実験	A	B	C	D	E	評価(平均)
1. 子供を常時映すことはできていたか？	1回目	2	1	3	3	4	2.6
	2回目	3	3	3	3	3	3.0
	3回目	4	5	3	3	4	3.8
2. カメラは見たい方向に向けられていたか？	1回目	1	1	2	3	2	1.8
	2回目	3	2	3	4	3	3.0
	3回目	4	4	4	4	4	4.0
3. 希望通りの大きさで子供を映すことはできていたか？	1回目	2	1	3	3	2	2.2
	2回目	3	3	4	4	3	3.4
	3回目	4	3	4	4	4	3.8
4. 子供の状況を把握できたか？	1回目	3	2	3	2	4	2.8
	2回目	4	3	4	4	4	3.8
	3回目	5	5	4	3	5	4.4

表 13 システムの評価結果 (評価)

質問内容	評価(平均)
1. 映す人数の希望の変更を行う操作は簡単か？	3.8
2. カメラが制御している人数と、希望人数との差を表示することでストレスの軽減はできたか？	3.2
3. 画質に満足できたか？	4.2
4. 映像は見やすかったか？	3.8
5. 子供を変更できる機能は必要か？	3.2
6. 人数の希望でズームの変更を行う方法に満足できたか？	3.2
7. 子供の通う保育施設にライブカメラサービスを導入して欲しいか？	2.8
8. ライブカメラサービスとして本システムは有効か？	3.8

表 14 システムの評価結果 (コメント)

質問	コメント
1	<ul style="list-style-type: none"> ・操作自体は簡単で使いやすいが、望み通りのズームにまだ遠い ・人数変更を行ったが、映像に変化が見られなかった
2	<ul style="list-style-type: none"> ・軽減できている気もするが、結局は思い通りの操作ができていないと同じ ・自分の子供を中心に見たいので、思い通りにズームができないとストレスを感じる ・ユーザが複数アクセスしていると知ることで、待つ心構えができた
3	<ul style="list-style-type: none"> ・十分に認識できる画質であった ・子供の顔が良く見れた ・思っていたよりクリアだった
4	<ul style="list-style-type: none"> ・子供の動きに追いつけない時もあったが、充分であった ・子供が動くのもスムーズに捉えられていて、臨場感があった ・切り替えの動きが少し遅かった
5	<ul style="list-style-type: none"> ・自分の子供のみがはぐれている時、他の子供が何をしていて、どうして参加できないのかを知るためには必要 ・自分の子供だけ見れたら良い ・兄弟姉妹がいる場合は必要 ・他の子とトラブルがあった時にその子の様子を見たい
6	<ul style="list-style-type: none"> ・大きさが変えられなかった ・期待以上に子供の様子が捉えられたので満足 ・全体の様子が見れたのも良かった ・ユーザ主導でズーム等を行える方が良い ・簡単だった
7	<ul style="list-style-type: none"> ・他の友達とどのように関わっているのかを知りたい ・あると、自分の子供が気になってしまい、子供の欠点や幼稚園の不満が高まってストレスが溜まる ・仕事に集中しにくい
8	<ul style="list-style-type: none"> ・子供を探すのが大変なので、有効である ・自分の子供が心配なので、自分の子供を中心に追いかけるのが嬉しい ・どの程度フォローされるのか分からないので、どちらとも言えない
その他の意見	<ul style="list-style-type: none"> ・カメラは子供の届かない位置に設置すべき ・音声はあった方が良いが、なくても良いのでは ・1回目の実験で、人物群が中心にいなかったのが気になった

じることが質問 2, 6 におけるコメントから読み取れる。また、ライブカメラサービスの必要性については、サービスがあると子供の様子が気になって時間を奪われるという意見から、平均 2.8 ポイントという評価であった。しかし、システムの有効性については平均 3.8 ポイントの高評価となり、自分の子供を中心にカメラ制御が行われることが良かったようであるとコメントの結果からも分かる。

4.4 考察

(1) 位置推定の誤差による問題

カメラ 2 台の 1 回目の実験では、提示される映像への評価は低かったものの、システムが算出したユーザの満足度が非常に高い。システムとしては満足度の高くなる制御を行っていたにもかかわらず、実際に提示される映像は満足できないものであったことが分かる。これは、事後アンケートのコメントにおける「その他の意見」にあるように人物群が画像の端に映っていたことが理由と考えられる。この原因は、魚眼カメラで検出した位置から三次元位置に変換を行うときの条件であった「床から 70cm の位置を子供の位置とする」としたことである。カメラ 2 台での実験の際、図 26 に示すように、子供達が床に座り込んで遊んでいたために、実際の位置よりも魚眼カメラ側に寄った位置を子供の位置として決定していたことで、実際に子供のいる位置よりも魚眼カメラ側を映し、子供達は画像端に映っていたのだと考えられる。また、子供達が動き回らずに固まって遊んでいたため、位置推定の誤差がなければ、画像内に全子供を納めることが可能であり、システムが算出したユーザの満足度は高くなったと考えられる。子供の位置推定については、子供の動きが早すぎると、基準フレーム（初期フレーム）で指示した子供の服の色の領域追跡がはずれてしまうことが起こり、その際には子供のいない位置を子供の位置としてカメラ制御が行われた。提案手法では保育室内での子供の位置は既知としているが、システムが算出したユーザの満足度と、実際のユーザの満足度の差を小さくするために、センサなどを用いて人物位置をより正確に取得できる手法を適用するべきである。

(2) ズーム制御の問題

同じカメラを割り当てられるユーザが複数人いる場合には、各ユーザの希望人



図 26 カメラ 2 台の実験時におけるカメラ 1 の画像

数が少数であっても、あまりズームされないことがある。シミュレーション実験での結果でも見られたが、同じカメラを割り当てられたユーザの指定人物同士が離れていると、希望人数が少なくてもあまりズームインされないことが原因であると考えられる。これについては、誰にも割り当てられないカメラがある場合には、そのカメラを利用できないか検討して割り当ての修正を行うことで軽減できると考えられる。

(3) システムの有効性

コメントから、自分の子供を見たいという意見が多く見られる。事後アンケートにおける質問 5 の子供を変更できる機能の必要性については、自分の子供と他の子供との関わりを見たいため必要とする意見か、自分の子供だけ見られたら良いので必要ないとする意見となっており、自分の子供を心配する様子が分かる。保育施設でたくさんの子供の中から自分の子供を探すことは大変であり、本システムは、アンケートの結果から、ユーザの子供を中心にカメラ制御を行う点で保育施設のライブカメラサービスに有効であると考えられる。

5. まとめと今後の課題

保育施設におけるライブカメラサービスで、複数台の PTZ カメラを制御してカメラ台数よりも多くのユーザが同時にアクセスした時でも、各ユーザの要求に応じた映像を提示する手法を提案した。そして、複数のユーザが同時に保育室の様子を見ることが可能なプロトタイプシステムを実現した。

まず、保育施設におけるライブカメラサービスを利用するユーザは、どういった映像を見たいと考えているのかをアンケートにより調査し、ユーザの要求を「見たい子供」と「見たい人数」の2点として、カメラ画像がこの要求をどの程度満たしているかを表す評価関数を定義した。そして評価関数により算出した満足度が最大となるカメラをユーザに割り当て、同じカメラに割り当てられたユーザの満足度の総和を最大とするようにパン・チルト・ズーム値を最適化することで、カメラの制御を行う手法を提案した。シミュレーション実験により、ユーザが増え、見たい人数が少ないユーザと多いユーザに同じカメラが割り当てられた場合には、見たい人数の少ないユーザの要求を大きくは満たせないものの、カメラ4台に対してユーザ数10名程度まで対応可能な手法であることを確認した。

次に、提案手法を適用した以下の処理を行うシステムを実現した。

1. 部屋にいる子供の位置情報の取得
2. ユーザの要求の入力
3. 1, 2 の情報からの各カメラの制御パラメータの決定とカメラ制御
4. カメラ画像の配信と受信

本システムを用いた被験者実験により、ズーム制御には改良の余地があるものの、ユーザの子供を中心にカメラ制御を行う点で保育施設のライブカメラサービスに有効であることを確認した。

今後の課題として、ユーザへのカメラ割り当て後に割り当て修正の再検討を行うこと、ユーザが簡単に使用可能なシステムとするように Web ブラウザ上でのシステムの構築がある。また、被験者のアンケート結果では、子供の指定を変更できる機能を必要とする意見もあった。これについては、全方位カメラを保育室

に設置し，ユーザは見たい方向を見ながら，画像上で見たい子供をマウスクリックにより指定できる機能を導入することが考えられる．この全方位カメラの設置により，アクセスするユーザが多くなりすぎて一定値以上満足度が上がらないようならば，後からアクセスしたユーザに全方位画像を見て待ってもらうことが可能になると考えられる．さらに，プロトタイプシステムでは子供の識別と位置推定を，魚眼カメラを用いて画像処理により行ったが，これは実際に使用するシステムとしては精度が悪く，非現実的である．将来的には，アクティブ型RFIDなどのセンサ機器を利用して各子供の位置を取得する方法が考えられる．そして，実際の保育施設での検証実験を行うことが挙げられる．

謝辞

本研究の全過程を通して，懇切なる御指導，御鞭撻を頂いた視覚情報メディア講座横矢直和教授に心より感謝申し上げます．また，本研究の遂行にあたり，有益な御助言，御鞭撻を頂いたロボティクス講座小笠原司教授，ならびに本研究を進めるにあたり多くの御指導，御助言を頂いた視覚情報メディア講座山澤一誠助教授に厚く御礼申し上げます．さらに，物心両面において温かい御支援を頂いた視覚情報メディア講座神原誠之助手，的確な御助言を頂いた視覚情報メディア講座佐藤智和助手に深く感謝致します．また，研究室での生活を支えて頂いた視覚情報メディア講座事務補佐員守屋智代女史に心より感謝申し上げます．そして，本研究の全過程を通して無数のご助言，ご指導を賜った視覚情報メディア講座石川智也氏に厚く御礼申し上げます．さらに，忙しい中，本研究の実験をサポートして頂いた山中一樹氏，粟津優作氏，高橋典宏氏をはじめ，視覚情報メディア講座の諸氏には，研究活動において常に温かい御助言を頂き，深く感謝致します．最後に，実験に参加して頂いたNAIST KIDS プラスのメンバーの方々に心より感謝申し上げます．

参考文献

- [1] 株式会社インターネットオープンカレッジ: “キッズカメラ.NET”, <http://www.kidscamera.net/index.html>.
- [2] イルガラージュ株式会社: “ライブキッズ”, <http://livekids.jp/>.
- [3] 株式会社 NTT ネオメイトみやこ: “みーんな元気だよ”, <http://www.ntt-neo.com/news/2002/020726.html>.
- [4] インターネットコム株式会社, 株式会社インフォプラント: “保育園ライブカメラに関するアンケート調査結果”, <http://www.kidscamera.net/kouka.html>.
- [5] 山澤一誠, 八木康史, 谷内田正彦: “移動ロボットのナビゲーションのための全方位視覚センサ HyperOmniVision の提案”, 電子情報通信学会論文誌, Vol. J79-D-II, No. 5,, pp. 698–707, 1996.
- [6] Tao Zhao, Manoj Aggarwal, Rakesh Kumar and Harpreet Sawhney: “Real-time Wide Area Multi-Camera Stereo Tracking,” Proc. of IEEE Conf. on Computer Vision and Pattern Recognition, pp. 976–983, 2005.
- [7] 中澤篤志, 日浦慎作, 加藤博一, 井口征士: “分散視覚エージェントを用いた複数人物追跡システム”, 情報処理学会論文誌, Vol. 42, No. 11, pp. 2699–2710, 2001.
- [8] Faisal Z. Qureshi and Demetri Terzopoulos: “Surveillance Camera Scheduling:A Virtual Vision Approach,” Proc. of ACM international workshop on Video surveillance and sensor networks, pp. 131–140, 2005.
- [9] Cash J. Costello, Christopher P. Diehl, Amit Banerjee and Hesky Fisher: “Scheduling an Active Camera to Observe People,” Proc. of ACM 2nd international workshop on Video surveillance and sensor networks, pp. 39–45, 2004.

- [10] 楊長輝, 白井良明, 三浦純, 先山卓朗: “複数のアクティブカメラを用いた指定人物のリアルタイム追跡”, 画像の認識・理解シンポジウム (MIRU2004) 講演論文集, Vol. I, pp. 559–564, 2004.
- [11] Norimichi Ukita and Takashi Matsuyama: “Real-Time Cooperative Multi-Target Tracking by Communicating Active Vision Agents,” Proc. of International Conference on Pattern Recognition (ICPR’02), Vol. 2, p. 20014, 2002.
- [12] Norimichi Ukita: “Real-time Cooperative Multi-target Tracking by Dense Communication among Active Vision Agents,” Proc. of International Conference on Intelligent Agent Technology, pp. 664–671, 2005.
- [13] 大西正輝, 泉正夫, 福永邦雄: “デジタルカメラワークを用いた自動映像生成”, 画像の認識・理解シンポジウム (MIRU2000) 講演論文集, Vol. I, pp. 331–336, 2000.
- [14] “ボールと選手に着目したデジタルカメラワークの実現法 - デジタルシューティングによるサッカー解説映像生成システムに向けて-”, .
- [15] 井口泰典, 土居元紀, 眞鍋佳嗣, 千原國宏: “スポーツ映像放送のための実時間映像解析によるマルチカメラの自動制御と自動スイッチング”, 映像情報メディア学会誌, Vol. 56, No. 2, pp. 271–279, 2002.
- [16] 加藤大一郎, 津田貴生, 井上誠喜: “知的ロボットカメラ”, NHK 技研 R & D, No. 64, pp. 10–17, 2000.
- [17] 尾関基行, 中村裕一, 大田友一: “机上作業シーンの自動撮影のためのカメラワーク”, 電子情報通信学会論文誌, Vol. J86-D-II, No. 11, pp. 1606–1617, 2003.
- [18] Yong Rui, Anoop Gupta, Jonathan Grudin and Liwei He: “Automating lecture capture and broadcast: technology and videography,” .

- [19] NHK 放送技術研究所: “知的ロボットカメラによる自動番組制作”,
<http://www.nhk.or.jp/strl/open2001/tenji/id20/index.html>.
- [20] 宮崎英明, 亀田能成, 美濃導彦: “複数のカメラを用いた複数ユーザに対する講義の実時間映像化法”, 電子情報通信学会論文誌, Vol.J82-D-II, No.10,
pp. 1598–1605, 1999.
- [21] 純丘曜彰: エンターテイメント映画の文法 ヒットを約束する脚本からカメラワークまで, フィルムアート社, 2005.
- [22] Gray R.Bradschi: “Computer Vision Face Tracking For Use in a Perceptual User Interface,” Proc. of Intel Technology Journal, Q2, 1998.

表 15 全実験パターン

実験 ユーザ数(人)	室内の子供の数:ユーザ数		室内の子供の数:35人	
	I	II	III	IV
5	全ユーザ 希望人数 一人	希望人数 ばらつき 有り	全ユーザ 希望人数 一人	希望人数 ばらつき 有り
10				
15				
35				

表 16 実験 IV における各ユーザの希望人数

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
ユーザ5人	1	5	5	20	35										
ユーザ10人	1	1	5	5	10	10	20	20	35	35					
ユーザ15人	1	1	5	5	5	10	10	10	20	20	30	30	35	35	35

付録

シミュレーション実験における結果の詳細を載せる。実験は本文中 2.3.1 節において述べた条件で行っている。また、表 15 に示すように、本文中の実験パターンに加えて、室内に 35 人の子供がいるとき、ユーザの希望人数が全員一人である場合 (III) と、希望人数にばらつきのある場合 (IV) の実験結果も示す。実験 IV における各ユーザの希望人数は表 16 に示す。

それぞれの実験の各ユーザへの提示画像から、画像上で指定人物が占める割合 (i)、画像上に映る人数 (ii)、指定人物を捉えることのできていた割合 (iii)、各ユーザの満足度 (iv) の約 2 分間の平均を表 17 から表 21 に示す。また、実験 I-5,II-5,I-10,II-10 の結果については、本文中で結果を示しているので省略する。

表 17 の実験 IV-5 の結果 (i) では、本文中に示した実験 II-5 の時の結果よりも指定人物が画像上で占める割合が小さくなっている。これは室内に存在する子供の人数が増えたために、オクルージョンが生じていることが原因である。結果 (ii)

表 17 実験 III,IV-5 の評価結果

	III				IV			
	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)
A	0.059	8.8	0.85	0.75	0.058	10.9	0.97	0.84
B	0.043	7.7	0.96	0.80	0.037	9.2	0.99	0.83
C	0.048	7.5	0.94	0.81	0.027	12.4	0.97	0.79
D	0.037	9.7	0.87	0.78	0.018	18.8	1.0	0.85
E	0.043	8.8	0.89	0.78	0.018	19.3	1.0	0.90

表 18 実験 III,IV-10 の評価結果

	III				IV			
	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)
A	0.025	13.0	0.79	0.62	0.036	16.4	0.90	0.64
B	0.041	9.5	0.83	0.70	0.041	12.1	0.97	0.76
C	0.033	10.6	0.89	0.68	0.032	12.5	0.97	0.77
D	0.020	11.4	0.64	0.59	0.030	14.3	0.97	0.75
E	0.025	10.8	0.88	0.68	0.025	14.7	0.97	0.75
F	0.046	8.8	0.82	0.75	0.020	15.5	0.97	0.76
G	0.044	10.7	0.82	0.68	0.014	21.7	0.93	0.76
H	0.037	10.8	0.85	0.62	0.015	19.6	0.94	0.71
I	0.048	13.7	0.86	0.67	0.009	21.7	0.97	0.77
J	0.034	11.9	0.90	0.69	0.008	21.8	0.91	0.71

を見ても、画像内に映りこむ人数が増えている。表 18 の実験 IV-10 の結果 (i) を見ると、室内に 35 人の子供がいる場合にも、ユーザ 10 人に対して希望人数に合わせた大きさで各ユーザの子供を捉えることができていることが分かる。そして、ユーザ数が 15 人 35 人と増えると、希望人数には関係なくなっていることが分かる。これらの結果より、室内の子供が増えた場合にも、ユーザ数 10 名程度まで対応可能であると考えられる。

表 19 実験 I,II-15 の評価結果

	I				II			
	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)
A	0.035	7.5	0.73	0.53	0.030	8.9	0.89	0.55
B	0.038	8.6	0.92	0.61	0.013	10.0	0.93	0.51
C	0.052	5.3	0.91	0.69	0.026	7.0	0.84	0.66
D	0.048	7.9	0.90	0.64	0.022	9.4	0.90	0.58
E	0.035	6.9	0.84	0.60	0.020	7.2	0.94	0.75
F	0.037	6.0	0.88	0.58	0.020	9.0	0.94	0.70
G	0.039	6.5	0.81	0.57	0.020	9.7	0.79	0.70
H	0.039	6.5	0.90	0.61	0.015	7.7	0.97	0.77
I	0.034	8.1	0.92	0.63	0.013	10.4	0.98	0.80
J	0.047	7.0	0.86	0.68	0.018	9.5	0.95	0.76
K	0.033	7.1	1.0	0.75	0.008	10.8	1.0	0.84
L	0.047	8.9	0.94	0.61	0.025	10.9	0.98	0.80
M	0.073	8.1	1.0	0.77	0.023	11.3	1.0	0.87
N	0.039	8.0	1.0	0.73	0.015	10.9	1.0	0.77
O	0.063	7.6	1.0	0.80	0.014	10.9	1.0	0.87

表 20 実験 III,IV-15 の評価結果

	III				IV			
	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)
A	0.021	17.3	0.89	0.56	0.018	20.6	0.80	0.44
B	0.036	13.9	0.93	0.62	0.036	14.7	0.96	0.67
C	0.037	12.5	0.84	0.58	0.022	14.8	0.92	0.63
D	0.024	17.8	0.81	0.53	0.019	19.9	0.95	0.58
E	0.039	12.2	0.92	0.64	0.030	14.2	0.87	0.67
F	0.036	11.1	0.85	0.57	0.014	20.2	0.88	0.65
G	0.047	11.9	0.87	0.65	0.020	19.1	0.86	0.64
H	0.026	14.1	0.89	0.57	0.020	18.6	0.95	0.68
I	0.031	16.2	0.88	0.55	0.009	20.8	0.97	0.68
J	0.029	14.0	0.84	0.56	0.016	19.6	0.93	0.74
K	0.017	16.2	1.0	0.71	0.004	21.8	1.0	0.85
L	0.053	16.4	0.99	0.72	0.021	21.8	1.0	0.92
M	0.050	18.6	1.0	0.71	0.015	23.5	1.0	0.83
N	0.019	18.7	0.89	0.56	0.009	24.6	1.0	0.74
O	0.026	16.2	0.99	0.73	0.005	21.8	1.0	0.84

表 21 実験 I,II-35 の評価結果

	I				II			
	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)
A	0.015	25.9	0.83	0.37	0.012	23.8	0.94	0.44
B	0.007	25.8	0.80	0.33	0.011	21.3	0.75	0.35
C	0.017	19.8	0.84	0.46	0.009	20.5	0.68	0.40
D	0.010	24.4	0.89	0.40	0.010	23.6	0.83	0.37
E	0.013	19.4	0.71	0.36	0.012	21.5	0.75	0.45
F	0.013	19.4	0.76	0.37	0.009	20.3	0.71	0.45
G	0.013	21.7	0.83	0.45	0.014	22.3	0.82	0.55
H	0.014	21.8	0.95	0.42	0.013	20.8	0.98	0.58
I	0.010	23.8	0.81	0.36	0.006	25.2	0.83	0.47
J	0.013	21.8	0.88	0.44	0.011	21.6	0.86	0.55
K	0.006	26.8	1.0	0.47	0.006	25.0	1.0	0.61
L	0.025	25.7	0.98	0.55	0.017	22.4	1.0	0.79
M	0.026	24.9	0.99	0.47	0.028	22.4	1.0	0.82
N	0.011	26.2	1.0	0.51	0.009	25.2	1.0	0.73
O	0.014	26.1	1.0	0.52	0.011	25.1	1.0	0.69
P	0.016	25.0	0.91	0.49	0.009	26.2	0.92	0.58
Q	0.007	27.8	0.97	0.38	0.006	26.2	0.97	0.59
R	0.021	25.7	0.93	0.43	0.007	23.1	0.71	0.53
S	0.018	27.9	0.96	0.46	0.017	23.7	1.0	0.78
T	0.009	27.1	0.84	0.42	0.009	24.8	0.90	0.65
U	0.016	22.9	0.94	0.48	0.009	19.0	0.94	0.65
V	0.013	21.4	0.86	0.40	0.015	21.1	0.89	0.68
W	0.005	26.9	0.95	0.46	0.005	27.1	1.0	0.68
X	0.020	26.3	0.99	0.48	0.014	22.8	0.99	0.75
Y	0.018	25.4	1.0	0.56	0.018	26.4	1.0	0.80
Z	0.014	24.2	0.83	0.40	0.008	25.6	0.85	0.65
AA	0.006	26.7	0.96	0.37	0.005	27.1	0.97	0.73
BB	0.013	21.8	0.78	0.33	0.008	24.2	0.67	0.54
CC	0.009	25.0	0.85	0.33	0.008	24.6	0.84	0.64
DD	0.006	25.3	0.96	0.37	0.005	24.1	0.95	0.67
EE	0.018	22.9	0.73	0.39	0.016	20.0	0.76	0.64
FF	0.015	23.7	0.88	0.39	0.007	25.0	0.84	0.66
GG	0.011	22.2	0.95	0.44	0.007	23.0	0.85	0.65
HH	0.005	26.7	0.89	0.25	0.005	22.9	0.95	0.60
II	0.010	23.0	0.96	0.42	0.008	21.9	0.95	0.69