

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

大野 純佳[†] 石川 智也[†] 山澤 一誠[†] 横矢 直和[†]

[†] 奈良先端科学技術大学院大学 情報科学研究科

〒630-0192 奈良県生駒市高山町 8916-5

E-mail: †{sumika-o,tomoya-i,yamazawa,yokoya}@is.naist.jp

あらまし 近年、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サービスが普及している。このようなサービスでは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の取得を待つ必要があった。そこで本研究では、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を目標とする。そのために、本研究では事前に行ったアンケート調査を基に評価関数を設定し、その評価関数によりカメラをユーザに割り当て、各カメラに割り当てられたユーザの評価値を最大とするように最適化を行うことでPTZカメラを制御する。本報告では、シミュレーションによる定量的な評価と実環境での実験結果について述べる。

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 satisfying multiple users' requirements using multiple PTZ cameras in child-care facility

Sumika OHNO[†], Tomoya ISHIKAWA[†], Kazumasa YAMAZAWA[†], and Naokazu YOKOYA[†]

[†] Graduate School of Information Science, Nara Institute of Science and Technology

8916-5 Takayama, Ikoma, Nara, 630-0192 Japan

E-mail: †{sumika-o,tomoya-i,yamazawa,yokoya}@is.naist.jp

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. In conventional systems, a user controls a PTZ camera and can look at the direction of user's requirement with high resolution. However, a user is constrained to wait for the control when multiple users access to the system simultaneously. In this study, we aim at realizing a system which can show live videos which satisfy requests of users when multiple users access to the system simultaneously. We define an evaluation function which means user's satisfaction based on a result of a questionnaire. Our method assigns each user to a camera that maximizes the evaluation function. Then, the method controls multiple PTZ cameras by optimization of camera parameters to maximize the sum of the evaluation values of users' satisfaction. In an experiment of camera control, we show the validity of our method by quantitatively evaluating the video generated from a simulated environment and an experiment in a real environment.

Key words multiple PTZ cameras, multiple users, users' requirements, automatic camera control, child-care facility

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ことが出来る。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会えない孫の顔を見ることが出来るという理由から人気のサービスとなっている [1], [2]。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[3]でも、この様なライブカメラサービスを利用したいという意見は76%と高く、今後もますます普及することが考えられる。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、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ことを目標とする。

既存のライブカメラサービス[1],[2]では、PTZ(Pan, Tilt, Zoom)カメラを使用しており、ユーザが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。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る。一方、全方位カメラを使用したライブカメラサービス[4]では複数ユーザがアクセスしていても、ユーザそれぞれの見たい方向を同時に見ることができる。しかし全方位画像から生成したある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であり、子供の表情まで見たいという要求には応えられない。

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試みが成されている[5]~[7]。Faisalら[5]やCash[6]らは、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、仮想環境において検証実験がされている。また、実環境で複数のPTZカメラを使用した浮田ら[7]の人物追跡システムでは、カメラに接続されたPC間で対象情報を交換することで複数の対象を高解像度で捉えられる。しかし、これらの研究では、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。

複数ユーザそれぞれの要求に応じた映像を提供するシステムには、宮崎ら[8]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く。そして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。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は意味のない映像が提示されてしまう。

そこで本研究では、複数のPTZカメラを用いて複数ユーザの要求に応じた映像を提示する手法として、ユーザに提示される画像に対するユーザの満足度を表す評価関数を設定する。この評価関数はユーザの要求について事前に調べたアンケート結果に基づいて設定した。そして評価関数により求めた各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、各カメラに割り当てられたユーザの評価値の総和を最大とするように最適化を行うことでPTZカメラを制御する。

以下、2節では評価関数の定義について述べ、3節では評価

表1 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%

関数に基づいたカメラ制御について述べる。4節でシミュレーション実験と実環境における実験により本手法を適用したカメラ制御の結果を示し、5節でまとめと今後の課題を述べる。

2. 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 アンケートによる調査

アンケートは約3歳~12歳の子供を持つ保護者92名に対して行った。アンケートでは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画像を見てもらい、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つかを質問した。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表1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「見たい人物」と「見たい人数」の2点とした。

2.2 評価関数

評価関数はユーザの要求に合わせて、ユーザの「見たい人物」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。

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

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「イマジナリーライン」(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。

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。

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

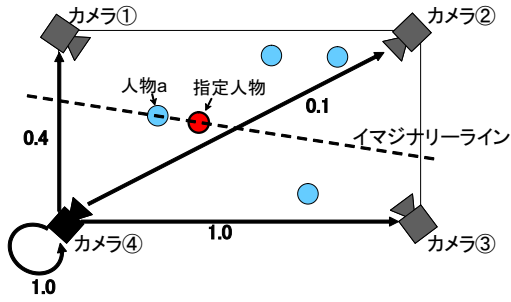


図1 イマジナリーラインによるカメラ間の遷移係数の決定

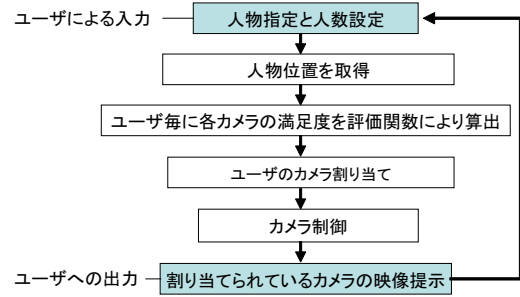


図2 カメラ割り当て・制御の処理の流れ

k_1, k_2, k_3 は各項目への重みである。 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 の見たい人物がどのくらい離れているかを表し、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。

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

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、 F_{ui} は最大値となる。

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。カメラを切り替える際にこの線を越えた位置から撮影すると人の位置関係や動きの方向性が逆になり、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 [9]。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 1 に示す。図 1 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③と④はイマジナリーラインを越えない位置にあるので、遷移係数は高くする。カメラ①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は対角のカメラなのでほとんど切り替わらないくらい低く設定する。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。

3. 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、2.2 項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 (3.2 項) と、各カメラのパラメータ (パン・チルト・ズーム値) の決定 (3.3 項) 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。

3.1 概要

図 2 に本提案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。全ユーザのカメラ割り当て後、カメラに割り当てられた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。

3.2 各ユーザへのカメラの割り当て

各ユーザのカメラの割り当ては、2.2 項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。

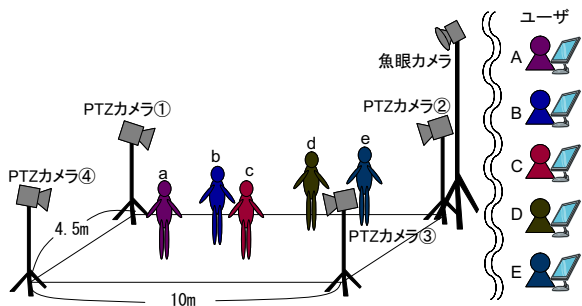


図3 実験環境

このようにしてカメラを全てのユーザに割り当てる。また、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ために、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。

3.3 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）の最適化を行い、カメラパラメータを決定する。カメラパラメータは2.2項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、ユーザが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、誰かが割り当てられるのを待つ。

4. 実験

提案手法の有効性を確認するためにシミュレーション環境と実環境で実験を行った。シミュレーション環境での実験では、提案手法に合ったカメラ制御ができているかを確認することを目的として、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。実環境における実験では実際のカメラの映像からユーザに提示される画像を示し、ユーザの要求に応じた映像が提示されているかを考察する。

実験は図3に示すようにシミュレーション環境・実環境とも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、カメラ台数より多い5名を撮影対象として、それぞれを見ている5名のユーザがいることとしている。また、ユーザの要求は表2のように設定し、図3中のユーザA,B,C,D,Eはそれぞれ人物a,b,c,d,eを指定して見ているものとする。人物位置情報の取得はPTZカメラとは別に魚眼カメラを使用し、追跡は従来手法[10]のCamShiftによる追跡法を利用している。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、その直線が床に交差する点を人物の位置として計算する。シミュレーション実験では、実環境での実験時に取得した撮影対象人物の移動経路データとカメラの制御データを利用して状況を再現した。なお、PTZカメラにはSONY EVI-D30を使用した。実験時の式(1)、(2)における各項の重み係数は経験的に W_1, W_2 は

表2 ユーザの要求の設定

ユーザ	指定人物	希望人数
A	a	1人
B	b	2人
C	c	3人
D	d	4人
E	e	5人

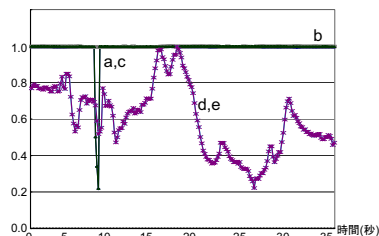


図4 各ユーザへの提示画像における指定人物の画像中心からの距離

0.5, $Trans_{ui}$ は図1に示す数値を利用し, k_1, k_2 は0.3, k_3 は0.4とした。

4.1 シミュレーション実験による評価

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。

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4に示す。図4における値は以下の式で計算する。

$$1.0 - \frac{\text{画像中心と人物の画像への投影点間の距離}}{\text{画像の幅}/2} \quad (6)$$

これは値が1.0に近いほど画像中心に近く、0.0に近いほど画像の端に映っていることを表す。人物a,b,cについてはほとんど1.0に近い値となり、指定人物をほぼ画像中心に捉えられていることが分かる。人物d,eについては0.8以下の値が多いが、このとき同じカメラを、ユーザD,Eの二人に割り当てており、この二人ともを捉えるようにカメラ制御が行われているからであると考えられる。このユーザD,Eは多い人数を希望しているので、他のユーザよりもズームアウトしたカメラ制御となり、二人の指定人物を同じカメラで捉えられている。また、二人の要求に応じるようにカメラパラメータを最適化して制御しているので、ちょうど左右対象に指定人物の二人が映るようになっていることがd,eの値が同じであることから分かる。また、人物a,cの数値が一時的に落ちているのはユーザAにユーザCと同じカメラを割り当てたためであるが、その後すぐに数値が上がっている。これらのことより、2.2項で示した見たい人物の要求に対する式(2)の k_1 の項に合ったカメラ制御ができていると考えられる。

次に、提示画像上で指定人物が占める割合を図5に示す。これはユーザに提示される画像に映る指定人物の画素数の総和を全画素数(640 × 480pixels)で割ることで求めている。このグラフにおける人物の映る大きさの例を図6に示す。図6(a)の正面を向っている人物が画像上で占める割合の値は0.12であり、図6(b)の一番右に移っている人物の画像上で占める割合の

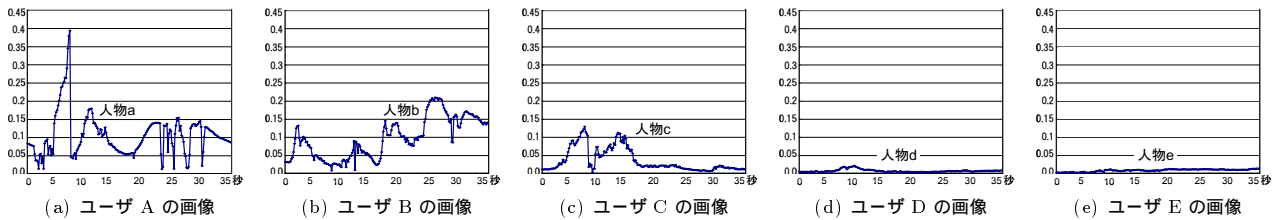


図5 各ユーザへの提示画像における指定人物の占める割合

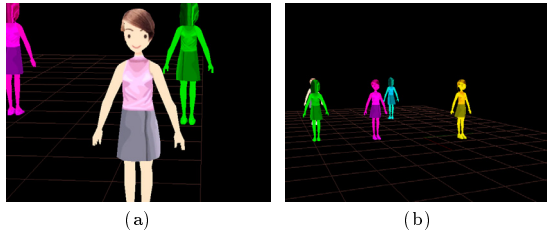


図6 人物の映る大きさの例

値は 0.009 である．希望人数が少ないほど映る大きさが大きくなっていることが分かる．希望人数が少ないほど，なるべく高い解像度で子供を見たいとユーザが希望していると考えられ，その希望にあったズーム値の決定をしていることが，この結果から読み取れる．

各ユーザに提示される画像に，どの人物が映っているかを図7に示す．それぞれ指定人物を常に捉えることができています．また，希望人数に合わせて映りこむ人数が変わっていることが分かる．カメラ台数よりも多い人数であるが，カメラ画像における人物の映り方からユーザの満足度を算出して，各ユーザに合ったカメラを割り当てていることで，ユーザの要求に応じた映像が提示できていると考えられる．しかし，ユーザ D, E はそれぞれ 4 人と 5 人を希望しているが，グラフの 30 秒以降では 3 人しか映っていない．これはその 3 人の近くに他の二人がいなかったためにこのような結果となっている．このとき，周りにいない他の二人を映そうとして，カメラは最大までズームアウトして映しているため二人の映り方は小さいままであった．また，一人を希望しているユーザ A の画像で数フレーム 5 人が映っているところが見られる．これは人物 a をアップで映しているものの，後ろに他の人物が映りこんでしまったためであり，図5の人物 a の画像上で占める割合を見ても分かるように，ユーザ A には a の人物が大きく映るようにカメラ制御された映像が提示されているので，適切なカメラ制御であると考えられる．

4.2 実環境での実験

実環境で，本手法によるカメラ制御の実験を行い，カメラから得られた画像からユーザの要求に応じた画像となっているかを考察する．本実験では，大人 5 名が観測される人物として，保育施設における子供の遊ぶ様子を模擬するように室内を動いた．また，カメラの制御周期は，カメラ制御用 PC (Pentium D 940) を使用したところ，カメラ割り当てや最適化処理および魚眼カメラによる位置推定にかかる処理時間から，5～6 フレーム/秒となっていた．実験における各ユーザに提示される画像を図8に示す．これは約 7 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．図8では，説明のために各ユーザの指定する人物を楕円で示す．

「見たい人物」の要求に関しては，どのユーザの提示画像からも，画像内にユーザの指定人物を捉えられていることが分かるので，要求に応じた映像が提示できていると考えられる．

カメラ割り当てについてユーザ A, C の映像から考察する．ユーザ A, C は 3584～3602 フレーム辺りで，カメラ①とカメラ④に割り当てられている．そして 3613 フレームでユーザ A のカメラ割り当てがカメラ④に変わっている．このときカメラ④は人物 a と人物 c を捉えるように制御されるが，人物 c は画像の端の方に映っている．そのため 3620 フレームでは，ちょうど人物 c を捉えていたカメラ①にユーザ C の割り当てが変更されている．以上のことから，指定人物を捉えるようにカメラ割り当ての処理が効果的に行われていることが分かる．

「見たい人数」の要求に関しては，希望人数一人のユーザ A の画像と，希望人数 4,5 人のユーザ D, E の画像を見比べると，ユーザ A の画像ではユーザ D, E に比べて一人だけを映すようにズームインしていることが分かるので，希望人数に合わせたカメラ制御となっていることが確認できる．しかし，ユーザ B の画像を見ると 3602～3613 フレームでは両端に大きなスペースがあり，ユーザ B に適したカメラパラメータとなっていなかった．これは，ユーザ B の指定人物が走っていたため，人物位置推定の精度が低下したことによる影響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，複数のユーザが同時にアクセスした時でも，複数台の PTZ カメラを制御して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．本研究では，このようなサービスにおけるユーザの見たい映像をアンケートにより調査し，ユーザの要求を「見たい人物」と「見たい人数」の 2 点として，提示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．そしてその評価関数が最大となるようにカメラをユーザに割り当て，各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・ズーム値の最適化を行ってカメラの制御を行った．また実験により，ユーザの要求が考慮されていることを確認した．今後の課題として，さらに人数が多い場合など様々な状況において実験を行い，評価関数におけるパラメータを決定することや，システムを構築し，実際の保育施設での実験を通した保護者による主観評価を行うことが挙げられる．

文 献

[1] <http://www.kidscamera.net/index.html>:キッズカメラ.NET,

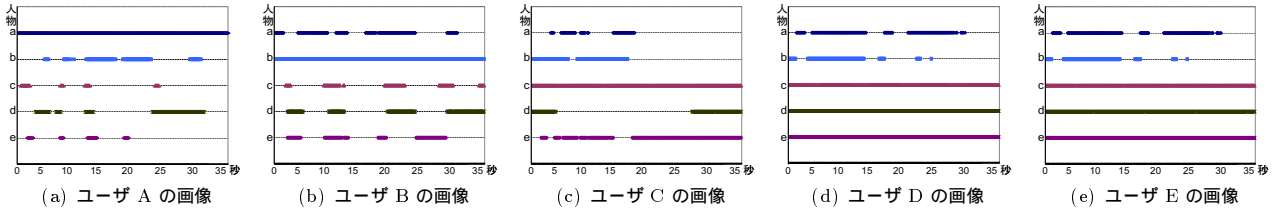


図 7 ユーザ提示画像に映る人物



図 8 各ユーザへの提示画像

株式会社インターネットオープンカレッジ.

[2] <http://livekids.jp/>: ライブキッズ, イルガラーヂュ株式会社.
 [3] <http://www.kidscamera.net/kouka.html>: 保育園ライブカメラに関するアンケート調査結果, 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.
 [4] <http://www.ntt-neo.com/news/2002/020726.html>: みーんな元気だよ, 株式会社 NTT ネオメイトみやこ.
 [5] 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.
 [6] Cash J. Costello, Christopher P. Diehl, Amit Banerjee, 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.
 [7] 浮田宗伯, 松山隆司: "能動視覚エージェント群による複数対象の実時間協調追跡", *情報処理学会 CVIM 研究会論文誌*, Vol.43, No. SIG11, pp.64-79, 2002.
 [8] 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", *電子情報通信学会論文誌*, Vol.J82-D-II, No.10, pp.1598-1605, 1999.
 [9] 純丘曜彰: "エンターテインメント映画の文法 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.
 [10] Gray R. Bradski: "Computer Vision Face Tracking For Use in a Perceptual User Interface", *Proc. of Intel Technology Journal*, Q2, 1998.