

全方位カメラを用いた個人注目映像の生成

Generation of Individual-focused Images with an Omni-directional Camera

白須 遼[†]
Ryo Shirasu

加藤 ジェーン[†]
Jien Kato

石井 健一郎[†]
Kenichiro Ishii

1 はじめに

今日、監視カメラや会議の場などにおいて、複数の人物の挙措をそれぞれ映像として記録したいという需要が存在している。例として石川ら [9] が行っている幼稚園児の1日ダイジェストの自動生成が挙げられる。この研究で生成されたダイジェストに対する保護者のアンケート結果において、「園児の表情が分かるようにカメラを設置して欲しい」との意見が得られた。このことから実際に対象人物の目線で映像を記録することに需要があることが伺える。このような場合、通常の室内設置型カメラを撮影に用いると、カメラ下部が死角となること、撮影角度の関係から顔が記録できない人物の存在など、種々の問題が発生する。この対策として、人物毎にカメラを用意することがまず考えられるが、カメラの設置・管理コスト及び記録映像のデータ量の増加といった新たな問題が浮上するため、現実的ではない。そこで本研究では映像の撮影に全方位カメラを採用することでこの問題に対処する。

全方位カメラは360°全方位を撮影できるという特徴を持ち、ロボットの周囲環境認識 [1] や、複合現実感実現のための実世界の3次元モデル構築 [2]、複数方向の歩容画像を特徴とした個人認証 [3] 等に利用され、応用研究が活発である。またその一方、全方位カメラ利用の際に発生する代表的な問題として、カメラのキャリブレーション、全方位を1枚の画像に収めることによる映像の解像度劣化、パノラマ展開における歪みの発生が挙げられる。どの問題に対しても解決手法の研究は進んでおり [4][5][6]、上記の全方位カメラを利用した応用研究の基盤を支えている。以上のように、全方位映像は従来のカメラの映像よりも情報が多く特にその視野が360°不変であるという特徴を持つことから、全方位カメラは様々な研究のアプローチにおいて利用されている。

この特徴を利用すれば、例えば、全方位カメラを部屋の中心に設置することで部屋にいる人物全員の動きを記録した映像を撮影することが可能である。その結果、部屋の中央に全方位カメラを1台設置するだけで死角を生ずることなく室内の人物全員の映像を記録できるようになり、カメラの設置・管理コストの削減が期待できる。さらに、全方位カメラで撮影される映像は室内設置型カメラの俯瞰視点と異なり、対象人物と同じ目線で記録されることから、人物の顔、服装及び挙措が、俯瞰視点の映像よりも



図1 個人注目映像の生成

鮮明に記録することができる。

ここで、全方位カメラで撮影される映像はカメラの周囲を記録した、極めて横長の映像となる。このような映像は個人の動作を観察する場合撮影範囲が冗長であり、視聴し難いものとなる。そこで本研究では全方位カメラで撮影した映像を図1のように映像中の各人物に注目した領域で切り出すことを提案する。本論文ではこの映像を個人注目映像と呼称する。また、映像を人物毎に切り出すことによるメリットとして複数の人物が混在した全方位映像を記録するのではなく、その全方位映像に存在する各人物についての映像を別々に記録・管理することが可能になること、つまり、管理しやすい映像をデータ管理者に提供できることも挙げられる。

全方位映像からの個人注目映像の切り出し処理において、目標となるのは切り出した後の個人注目映像が自然で見やすい映像になることである。映像中で人物は自由に動き回り、時折短時間で緩急の変化が強い動きを行い、そして極端にカメラから離れる、または近づく場合がある。これらのことを考慮した映像の切り出しを行うことで視聴者が視聴しやすい個人注目映像を生成することを図る。

以上のことを踏まえて、本研究では複数の人物が存在する1つの全方位映像から対応した個人注目映像を生成するプロトタイプシステムを実現し、その評価を行う。

2 研究のアプローチ

本章では、見やすい個人注目映像を生成するために本研究が取るアプローチについて述べる。

2.1 映像に対する処理の内容

本研究では、複数の人物の動きを撮影した全方位映像から見やすい個人注目映像を各人物について生成することを目的とする。

[†] 名古屋大学大学院情報科学研究科

そのために全方位カメラの映像を入力として、以下の処理を実行するプロトタイプシステムを制作する。

1. 人物の検出・追跡
2. 人物の情報取得
3. 個人注目映像の切り出し境界の決定

プロトタイプシステムでは複数の人物が存在する室内を撮影した全方位映像とその背景画像を入力とし、検出された人物1人1人の個人注目映像を予め決定された縦横比、画面サイズで出力する。問題の簡略化のため、プロトタイプシステムは室内の映像を対象とする。さらに、背景の急な変化及び人物の入退室がなく、3人以上の人物の重なりによるオクルージョンは発生しないという制約のある撮影環境を想定して制作する。個人注目映像の生成では、まず人物の検出・追跡、人物同士のオクルージョン発生への対処が第一の課題であるが、これらに対処する手法の実現は既に研究が進んでいる分野である [7][8]。そこで本研究では簡便かつ既知の手法、具体的には検出に背景差分法、追跡に人物の追跡には色分布を用いたフレーム間特徴マッチングをそれぞれ採用することでこの課題に対処する。個人注目映像の切り出し境界決定の基準となる情報として、各人物について全フレームにおける外接矩形、重心及びその移動ベクトル及びオクルージョン発生の有無を用いる。各情報は人物検出の不完全さ及びオクルージョンの発生によりノイズ、ギャップが生じる。ここで各情報は時間的に連続であることが望ましいため、情報の平滑化及び補間を同時に行う。以上の情報を用いた切り出し境界の決定においては、あらかじめ定めた見やすい映像の定義に沿った映像が生成されるような手法を提案・実装する。そして提案手法とベースラインについて比較実験を行った後、提案手法の長所と短所の評価、改善策についての考察を行う。

次節では、本論文で提案する自然で見やすい映像の定義について述べる。

2.2 自然で見やすい映像の定義

この節では本論文で提案する自然で見やすい映像の定義を述べる。定義は「自然で見やすい映像とは人物の周りの背景に注目した映像である」という考えに基づいたものである。

以下で述べる条件で構成される定義に基づいて、4章において個人注目映像の切り出し境界決定手法を提案する。

定義では以下の3つの条件を映像に要求する。

条件1

人物が常に映像の中心に位置することはなく、人物の動き、その速さに応じて映像中の人物の位置が変化する

条件2

人物に対するズームイン・アウトは予め決められたズーム率を基に段階的に実行される

条件3

人物にオクルージョンが発生した場合、人物が映像の中心に配置され、同時にズームアウトが行われる

上記3つの条件を採用した理由はいずれも映像が人物の周りの背景に注目したものになるようにするためである。条件1を

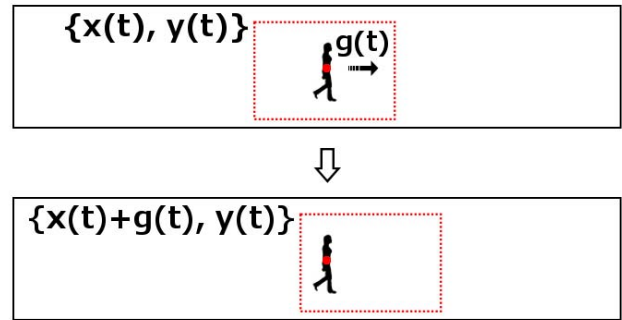


図2 見やすい映像の定義:条件1

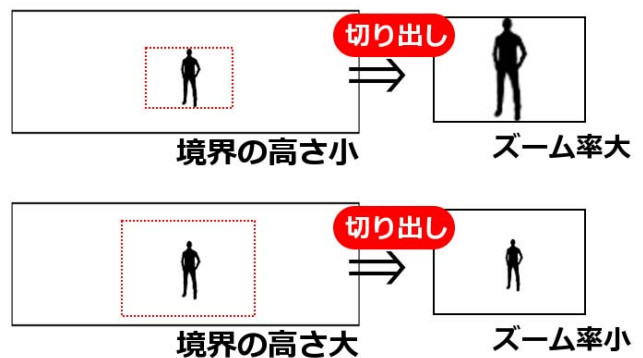


図3 切り出し境界の高さと映像のズーム率の関係

採用すると、例えば、人物が速く移動している時、図2のように映像中で人物を進行方向とは逆に位置させることで人物が何に対して移動しているか、また人物に対して起こるであろうことが予測できる映像となる。また条件2を採用すると、人物が常に映像の大部分を占めることを避けることができる。その結果、人物の周りの状況を把握できるような映像となる。また、ズームイン・アウトの回数を抑え、映像の激しい動きを抑える効果もある。予め決められたズーム率を基にズームイン・アウトすることは、個人注目映像の切り出しにおいて予め段階化された切り出し境界の高さを定義し、それに基づいて個人注目映像の切り出しを行うことで等価な操作を実現する。予め決められたズーム率が段階化された境界の高さに相当し、その境界の高さに合わせた映像の切り出しがズームイン・アウトに相当する。ここで、図3で示すように切り出し境界の高さが小さいほど映像のズーム率は高くなり、大きいほどズーム率は低くなる。最後に条件3を採用すると、人物にオクルージョンが発生した際、図4に示すように人物に対して何が起きているか、特に人物が何に隠れているかを把握できるような映像となる。

3 人物の追跡

本章では、各人物の情報取得のために個人注目映像生成プロトタイプシステムが行う処理の内容について述べる。

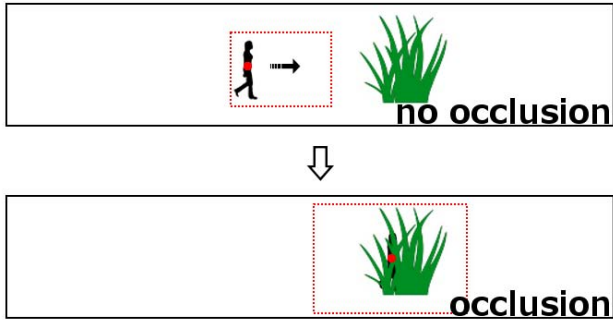


図4 見やすい映像の定義:条件3

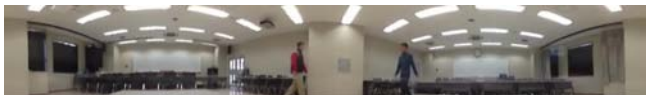


図5 入力画像



図6 背景差分法によって得られる前景画像

3.1 背景差分法による人物の検出

プロトタイプシステムでははじめに、入力映像の各フレームについて現在のフレーム画像と背景画像から、背景差分法により映像中で人物が占める領域の検出を行う。以下で、入力画像と背景画像から差分画像を計算し、結果に二値化及びノイズ除去を実行する処理手順をその結果とともに述べる。

画像の各画素の座標位置を x, y 、各フレームの時間を t とし、式 (1) のように入力画像 $I_i(x, y, t)$ と背景画像 $B_i(x, y, t)$ から背景差分画像 $S_i(x, y, t)$ を求める。ただし、 $i = r, g, b$ で各画素の RGB 成分を表す。

$$S_i(x, y, t) = I_i(x, y, t) - B_i(x, y, t) \quad (1)$$

得られた背景差分画像を Wharton の手法 [10] を用いて濃淡画像 $G(x, y, t)$ に変換した後、大津の判別分析法 [11] を使って求めたしきい値 T により式 (2) で二値画像 $N(x, y, t)$ を生成する。

$$N(x, y, t) = \begin{cases} 1 & G(x, y, t) \geq T \text{ のとき} \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

この結果 $N(x, y, t)$ では 1 の画素が白色、0 の画素が黒色の画素として二値化される。以上の処理により得られた二値画像には小さな粒状のノイズが存在するため、画像のオープニング・クローズング演算 [12] を行いノイズを除去し、前景画像を得る。ここまでの処理の結果を図5、図6に示す。

次に、この前景画像から各フレームに存在する人物の数と各人物の大まかな位置を、前景画像を膨張させ各人物の領域を1つの連結成分にすることで求める。この処理によって、背景差分法によって得られる前景画像中の各人物の領域から穴及び凹部を



図7 膨張処理によって得られる各人物を表す連結成分

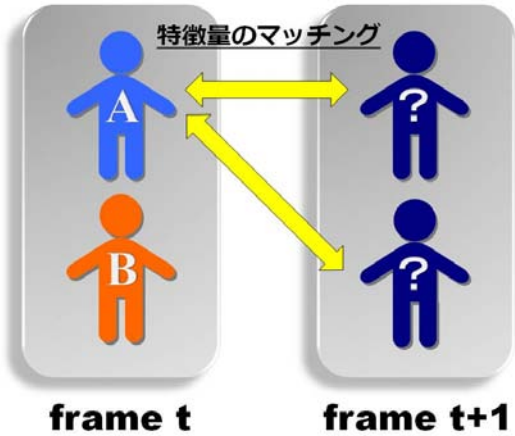


図8 フレーム間特徴マッチングによる人物追跡

除去し、それぞれを1つの連結成分とする。膨張処理によって得られた二値画像内の連結成分の数と位置をそのフレーム t 中に存在する人物の数とその大まかな位置とする。ここまでの処理の結果を図7に示す。

3.2 特徴量マッチングによる人物追跡

フレーム間での各人物の特徴量マッチングを最初のフレームから最後のフレームまで行い、各人物の追跡を行う。図8で示すように現在のフレーム t と次フレーム $t+1$ の間で、全人物の特徴量マッチングを行い、フレーム t 内のある人物についてフレーム $t+1$ 内で最も特徴の類似度が高い人物の位置をフレーム $t+1$ におけるその人物の位置とする。人物の特徴として用いたのは人物についての色相分布、彩度分布、RGB それぞれのチャンネルの分布の合計5つである。人物のマッチングにこれらの特徴を採用した理由は、全方位カメラの映像で人物を見たときにその人物を一意に判別できる最適な特徴は人物の服装、つまり人物の色分布と考えられるからである。以下で、上記の5つの特徴から人物間の特徴の類似度を算出する過程を述べる。

フレーム間での分布の類似度の比較手法は複数パターンの入力映像で最も正確に人物を追跡できたことから、経験的に Bhattacharyya マッチング [13] を採用する。ある特徴分布を表現した2つのヒストグラム H_1, H_2 について、その Bhattacharyya 距離を $d(H_1, H_2)$ とする。 d は

$$0.0 \leq d \leq 1.0$$

の範囲の値を取り、低い値で特徴がよく一致している事を示し、高い値で特徴が一致していない事を示す。Bhattacharyya 距離を用いた人物の類似度導出のために人物の5つの特徴分布を5つのヒストグラムで表現する。人物 i の色相分布、彩度分布、

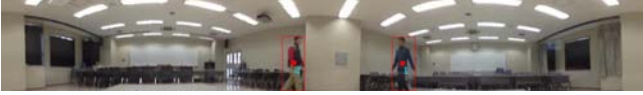


図9 人物の外接矩形及び重心の計算結果

RGB それぞれのチャンネルの分布をヒストグラムで表したものを H_h^i , H_s^i , H_r^i , H_g^i , H_b^i とし、各ヒストグラムを特徴とした人物 i , j 間の Bhattacharyya 距離を d_h^{ij} , d_s^{ij} , d_r^{ij} , d_g^{ij} , d_b^{ij} とする。ここで、人物 i , j の特徴の類似度 $D(i, j)$ を以下のように定める。

$$D(i, j) = 5.0 - (d_h^{ij} + d_s^{ij} + d_r^{ij} + d_g^{ij} + d_b^{ij}) \quad (3)$$

$$D(i, j) \text{ は} \quad 0.0 \leq D(i, j) \leq 5.0$$

の範囲の値を取り、高い値で人物がよく類似している事を示し、低い値で人物があまり類似していない事を示す。 $D(i, j)$ を人物の特徴量マッチングにおける人物の類似度として用いる。

3.3 追跡情報の取得

映像中の全フレームにおける各人物の追跡が完了した後、各フレームにおける各人物の外接矩形と重心を求める。本論文では人物 i のフレーム t における外接矩形 $X_i(t)$ 及び重心 $G_i(t)$ を以下のように記述する。

- $X_i(t) = \{x_X(t), y_X(t), w_X(t), h_X(t)\}$
 - $x_X(t)$: 外接矩形の左上角の x 座標
 - $y_X(t)$: 外接矩形の左上角の y 座標
 - $w_X(t)$: 外接矩形の幅
 - $h_X(t)$: 外接矩形の高さ
- $G_i(t) = \{x_G(t), y_G(t)\}$
 - $x_G(t)$: 重心の x 座標
 - $y_G(t)$: 重心の y 座標

人物の外接矩形及び重心は、図6により示される、背景差分法によって得られた差分画像中の人物領域について外接矩形と重心をそれぞれ計算することで求める。あるフレームでの各人物の外接矩形と重心の計算結果を図9に示す。図中、人物の重心は赤点で示されている。同時に、人物の外接矩形及び重心にノイズ軽減のための平滑化を行う。各人物について、3フレーム以上連続で外接矩形及び重心が求められているフレーム t の各値をフレーム $t-1$ とフレーム $t+1$ での平均値で更新する。

次に各人物の重心の移動ベクトルを求める。本論文では人物 i のフレーム t における重心の移動ベクトル $\vec{g}_i(t)$ を以下のように記述する。

- $\vec{g}_i(t) = \{x_g(t), y_g(t)\}$
 - $x_g(t)$: 重心の移動ベクトルの x 成分
 - $y_g(t)$: 重心の移動ベクトルの y 成分

重心の移動ベクトルは現在のフレーム t とフレーム $t-\tau$ の重心を用いて式(4)で計算する。

$$\vec{g}_i(t) = \begin{cases} x_g(t) = x_G(t) - x_G(t-\tau) \\ y_g(t) = y_G(t) - y_G(t-\tau) \end{cases} \quad (4)$$

重心の移動ベクトルの計算に何フレーム前の重心を用いるかの値 τ はプロトタイプシステムのパラメータとして与えられる。

次にオクルージョンの発生により生じる外接矩形及び重心のギャップの補間を行う。オクルージョン発生前のフレームを t_s 、発生後のフレームを t_e 、オクルージョン発生フレーム数を T とする。 T フレーム間について式(5)と式(6)を用いて外接矩形及び重心の線形補間を行う。ここで、 $(u = 1, \dots, T)$ である。

$$\begin{aligned} x_X(t_s + u) &= x_X(t_s) + u \cdot \frac{x_X(t_e) - x_X(t_s)}{t_e - t_s} \\ y_X(t_s + u) &= y_X(t_s) + u \cdot \frac{y_X(t_e) - y_X(t_s)}{t_e - t_s} \end{aligned} \quad (5)$$

$$\begin{aligned} w_X(t_s + u) &= w_X(t_s) + u \cdot \frac{w_X(t_e) - w_X(t_s)}{t_e - t_s} \\ h_X(t_s + u) &= h_X(t_s) + u \cdot \frac{h_X(t_e) - h_X(t_s)}{t_e - t_s} \end{aligned}$$

$$\begin{aligned} x_G(t_s + u) &= x_G(t_s) + u \cdot \frac{x_G(t_e) - x_G(t_s)}{t_e - t_s} \\ y_G(t_s + u) &= y_G(t_s) + u \cdot \frac{y_G(t_e) - y_G(t_s)}{t_e - t_s} \end{aligned} \quad (6)$$

人物にオクルージョンが発生していること及びその開始フレームと終了フレームは切り出し境界決定のための情報として記録する。

以上の処理により映像中の各人物について、情報の取得が完了する。

4 切り出し境界の決定

本章では、2.2節で述べた見やすい映像の定義に基づいた個人注目映像の切り出し境界決定手法について述べる。切り出し境界の決定には3章で求めた各人物の情報を用いる。

本論文では人物 i のフレーム t における切り出し境界 $Y_i(t)$ を以下のように記述する。

- $Y_i(t) = \{x_Y(t), y_Y(t), w_Y(t), h_Y(t)\}$
 - $x_Y(t)$: 切り出し境界の左上角の x 座標
 - $y_Y(t)$: 切り出し境界の左上角の y 座標
 - $w_Y(t)$: 切り出し境界の幅
 - $h_Y(t)$: 切り出し境界の高さ

また、切り出し境界の縦横比は予め個人注目映像生成プロトタイプシステムのパラメータとして与えられている。切り出し境界決定手法は各フレームにおける人物 i の情報から、個人注目映像が見やすい映像の定義に沿った映像となるように $x_Y(t)$, $y_Y(t)$, $w_Y(t)$, $h_Y(t)$ を各フレームで決定する。次節では切り出し境界を見やすい映像の定義に基づいて決定する提案手法についてその詳細を述べる。提案手法では、以下の処理を行い、切り出し境界を決定する。

- 人物の大きさに応じた処理
- 人物の動きに応じた処理

4.1 人物の大きさに応じた処理

まず各フレームで切り出し境界の高さ及び幅を決定し、境界を表現する矩形を作成する。見やすい映像の定義で述べたとおり、

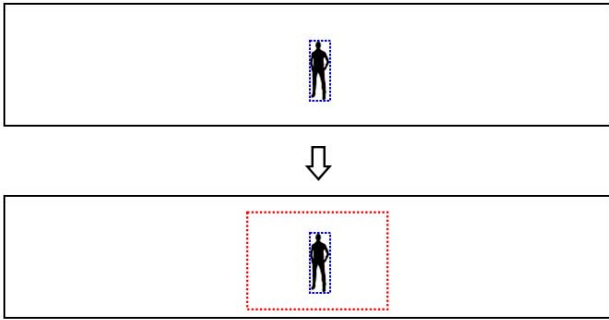


図10 提案手法における切り出し境界の高さの決定

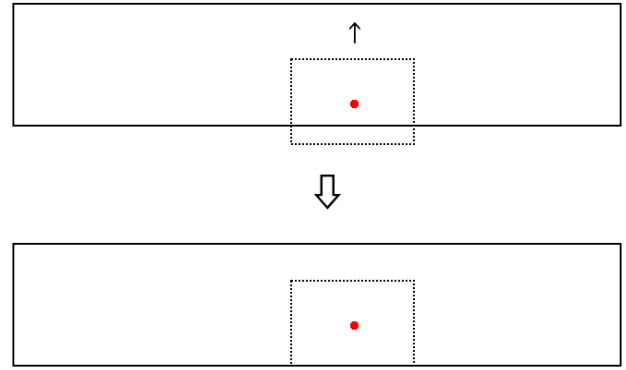


図11 切り出し境界が入力映像の上下にはみ出した場合の処理

映像中のズームイン・アウトを予め決められたズーム率を基に実行するために、ズーム率に相当する切り出し境界の高さは予め段階化されたものを用いる。そのために、各フレームにおける人物の外接矩形の高さに対応する境界の高さを段階化された境界の高さの中から選択する。高さの段階化は入力映像の高さを基に行い、段階化を何段階かで実行するかを決定する値 n はプロトタイプシステムのパラメータとして与えられる。入力映像の高さを h_{input} とし、 n 段階に段階化された切り取り境界の高さ H_k ($k = 1, \dots, n$) をそれぞれ式 (7) により求める。

$$H_k = \frac{h_{input}}{n} \cdot k \quad (k = 1, \dots, n) \quad (7)$$

フレーム t における人物の外接矩形の高さ $h_X(t)$ と n 段階に段階化された境界の高さ H_k ($k = 1, \dots, n$) について式 (8) を満たす H_k をフレーム t における境界の高さとする。ここでは $H_0 = 0$ とする。

$$H_{k-1} < h_X(t) \leq H_k \quad (k = 1, \dots, n) \quad (8)$$

図10に示すように、選ばれる境界の高さは必ず人物の外接矩形の高さより大きくなる。この処理で高さを決定した切り出し境界により切り出される映像は定義の条件2を満足する。

また、フレーム t において人物にオクルージョンが発生しており、かつ $k < n$ のときは H_k の代わりに H_{k+1} をフレーム t における境界の高さとして採用する。この処理は個人注目映像が定義の条件3を満たすための処理である。

ここまでの処理で求められた各フレームにおける境界の高さは離散的な値である。個人注目映像においてズームイン・アウトを滑らかに行うために値の変化を連続的にする必要がある。そこでフレーム t とフレーム $t+1$ の間で境界の高さの変化が発生するとき、フレーム t から α フレーム後のフレームまでその間のフレームにおける境界の高さについて式 (9) を用いて値の変化を連続的にする処理を行う。

$$h_Y(t+u) = h_Y(t) + u \cdot \frac{h_Y(t+1) - h_Y(t)}{\alpha} \quad (u = 1, \dots, \alpha - 1) \quad (9)$$

α の値を大きくするほど個人注目映像のズームイン・アウトは滑らかになり、小さくするほど急になる。 α の値は個人注目映像生成プロトタイプシステムのパラメータとして与えられる。

次に各フレームでの境界の幅を高さと縦横比から求める。以上の処理により境界の幅と高さが求められ、各フレームについて

切り出し境界を表現する矩形が作成される。

4.2 人物の動きに応じた処理

前節で作成した境界を入力映像の何処に設定するかを決定する。各フレームについてまず境界を境界の中心と人物の重心が一致するように配置する。このとき、境界の高さを人物の外接矩形の高さより大きく決定しているため、境界が入力映像に対して上下にはみ出す場合がある。その場合、境界の高さは変えず境界の y 座標を式 (10) により更新することで境界を上下に移動し、入力映像中に収める。更新後の境界の y 座標を $y'_Y(t)$ とする。

$$y'_Y(t) = \begin{cases} 0 & (y_Y(t) < 0) \\ h_{input} - h_Y(t) & (y_Y(t) + h_Y(t) > h_{input}) \end{cases} \quad (10)$$

この処理の内容を図11に示す。図中では実線で描かれた矩形を入力映像、破線で描かれた矩形を切り出し境界、赤点を境界の中心とする。

次に、個人注目映像を定義の条件1を満たす映像とするため、各フレームにおける切り出し境界を人物の重心の移動ベクトルの大きさに応じて左右に移動する。重心の移動ベクトルについて、人物の動きが緩やかなほどその値は小さく、激しいほどその値は大きくなり、さらにその変化は連続的である。よって各フレームで境界の x 座標を式 (11) により更新することで、境界を人物の速度及び移動方向に合わせて連続的に左右へ移動させることができる。更新後の境界の x 座標を $x'_Y(t)$ とすると、

$$x'_Y(t) = x_Y(t) + x_g(t) \quad (11)$$

となる。この処理により人物の移動方向の領域の情報も含まれた個人注目映像を生成する切り出し境界が決定される。処理の内容を図12に示す。図中では実線で描かれた矩形を入力映像、破線で描かれた矩形を切り出し境界、赤点を人物の重心とする

また、人物にオクルージョンが発生している場合、定義の条件3に基づき、人物を映像の中心に位置させるため式 (11) による境界の x 座標の更新は行わない。その結果、人物のオクルージョン発生時、終了時において境界の x 座標の値の変化が連続的ではなくなる。これは切り出し境界の位置の連続的な移動を阻害する問題となることから、次はオクルージョン発生時、終了

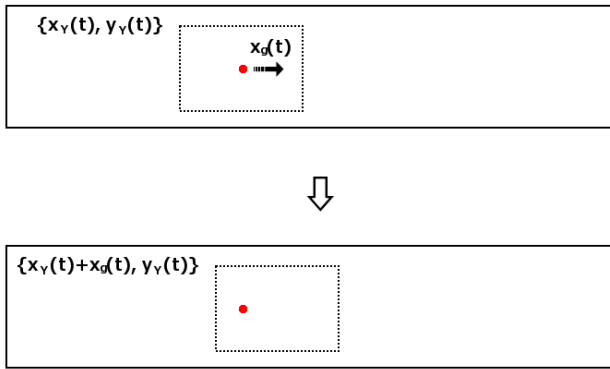


図12 人物の速さに応じた切り出し境界の左右への移動

後における境界の x 座標の値の変化を連続的にする処理を行う。オクルージョン発生フレームを t_s 、終了フレームを t_e 、処理を行うフレーム間隔を β フレームとして、式 (12) 及び式 (13) により値の変化の平滑化を行う。

$$x_Y(t_s - u) = x_Y(t_s) - u \cdot \frac{x_Y(t_s) - x_Y(t_s - \beta)}{\beta} \quad (u = 1, \dots, \beta - 1) \quad (12)$$

$$x_Y(t_e + u) = x_Y(t_e) + u \cdot \frac{x_Y(t_e + \beta) - x_Y(t_e)}{\beta} \quad (u = 1, \dots, \beta - 1) \quad (13)$$

この処理によって個人注目映像中の人物の位置はオクルージョンが発生する β フレーム前から徐々に映像の中心に近づき、オクルージョンが終了した後はその後 β フレームかけて人物の速度に応じた位置に戻っていく。

β の値はプロトタイプシステムのパラメータとして与えられ、この値が大きいくほどオクルージョン発生時における映像中の人物の位置の変化が緩やかになり、小さいほど速くなる。

以上の処理で切り出し境界の x 座標、 y 座標、幅及び高さが求められ、入力映像中で切り出し境界が一意に決定される。最後に、決定された切り出し境界に従って各フレームで入力映像を切り出し、指定された出力サイズに映像を拡大し、プロトタイプシステムの出力とする。

5 実験

第4章で述べた提案手法とベースラインについて比較実験を行った。実験では、急な背景変化のない室内で撮影した1人または2人の人物が含まれる全方位映像について個人注目映像を提案手法、ベースラインそれぞれを用いて生成し、どちらが映像として見やすいかの評価を20人の被験者が行った。撮影の対象としたのは5人の人間で、各人物について異なる動作を行っている場面を撮影した。撮影した映像の内容及び評価を行ったデータ数を表1に示す。他人物の後ろに隠れる動きに関しては2つのパターンを用意し、それぞれをパターンA、パターンBとする。パターンAで人物は図13で示すように移動中に他人物の後ろに隠れ、再び映像に現れた後は隠れる前と同じ方向に移動する。それに対してパターンBでは人物は図14で示すように再び映像に現れた後隠れる前とは逆の方向に移動する。個人注目映像の評価は提案手法、ベースラインそれぞれの手法で生成した映

表1 撮影した映像内容

映像内容	秒数(秒)	データ数(個)
歩く	5	40
早歩きする	5	40
走る	5	40
カメラに接近した後離反する	5	38
他人物の後ろに隠れるA	5	20
他人物の後ろに隠れるB	5	20
自由に動く	20	20
		218

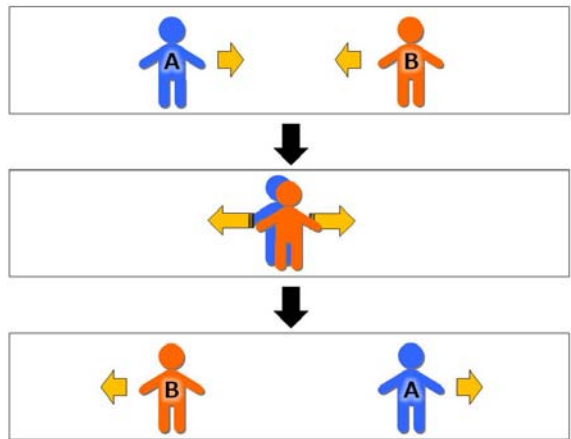


図13 他人物の後ろに隠れる動き(パターンA)

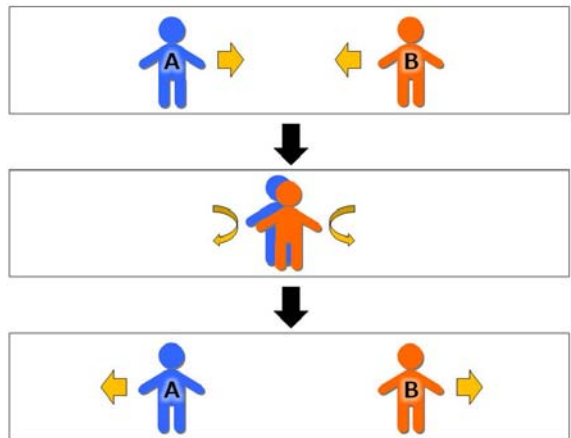


図14 他人物の後ろに隠れる動き(パターンB)

像に対してどちらが映像として自然で見やすいかを選ぶという2択方式で行った。

5.1 ベースライン

提案手法に対するベースラインとして、人物の映像中の位置が常に安定するように切り出し境界を決定する手法を用いる。各フレームにおいて境界の決定に用いる情報は人物の外接矩形及

び重心である。ベースラインでは、可能な限り人物が個人注目映像の大部分を占めるように、かつ人物が常に映像の中心となるように処理を行う。それぞれの処理の詳細を以下で述べる。

人物の大きさに応じた処理

各フレームで切り出し境界の幅及び高さを決定し、境界を表現する矩形を作成する。式(14)に表す通り、境界の高さの値は人物の外接矩形の高さの値を用いる。

$$h_Y(t) = h_X(t) \quad (14)$$

人物の外接矩形の高さは連続的に変化することから、出力映像が緩やかなズームイン・アウトが常に行われるような映像となる。各フレームでの境界の幅は高さと同様比から求める。以上の処理により境界の幅と高さが求められ、各フレームについて切り出し境界を表現する矩形が作成される。

人物の動きに応じた処理

次に、作成した境界を入力映像の何処に設定するかを決定する。人物を常に映像の中心に位置させるため、外接矩形のy座標を境界のy座標とし、境界の中心のx座標を人物の重心のx座標と一致させる。それぞれの処理を式(15)及び式(16)に示す。

$$x_Y(t) = x_G(t) - \frac{w_Y(t)}{2} \quad (15)$$

$$y_Y(t) = y_X(t) \quad (16)$$

以上の処理で切り出し境界のx座標、y座標、幅及び高さが求められ、入力映像中で切り出し境界が一意に決定される。

5.2 評価結果

映像データをその内容の特徴に応じて以下で示す3つの種類に分ける。

- *Line*
 - 人物が直線的で単純な動きをする映像。人物が歩く、早歩きする、走る映像がこれに含まれる。
- *Turn*
 - 人物が映像中でターンを行う、人物に向けて頻りにズームイン・アウトが行われる、*Line*よりも人物の動きが複雑な映像。人物がカメラに接近した後離反する、他人物の後ろに隠れる動きがこれに含まれる。
- *Free*
 - 人物の動きに規則性がなく、さらに*Turn*のような人物の複雑な動きを含む映像。人物が自由に動く映像がこれに含まれる。

Line、*Turn*及び*Free*の3種類の映像について、人物の背景に注目した手法と人物の位置の安定に注目した手法の比較評価結果を表2に示す。表中では、3種類の映像について各手法で見やすいと評価された数がカウントされている。

5.3 考察

表2で示される評価結果から、まず提案手法が*Turn*の映像に対して見やすい映像を生成できていない事が分かる。これは、提案手法で行う処理の目的である2.2節で述べた見やすい映像の定義の条件付けが間違っていたわけではなく、提案手法について、

表2 2つの手法の比較評価結果

種類	提案手法	ベースライン	総データ数(個)
<i>Line</i>	65	55	120
<i>Turn</i>	27	51	78
<i>Free</i>	12	8	20
<i>Total</i>	104	114	218

その処理を決定する際のプロトタイプシステムのパラメータの設定が不適切だったことが原因であると考えられる。具体的には、まず4章で述べた切り出し境界の高さの段階化を何段階で行うかを決定するパラメータ*n*の値を大きくし過ぎた事によりズームイン・アウトの回数が増加した。また、同章で述べたズームイン・アウトの緩急を決定するパラメータ*α*の値を小さくし過ぎたことによって急なズームイン・アウトが発生した。さらに同節で述べたオクルージョン発生及び終了時の映像中の人物の位置の変化の緩急をするパラメータ*β*の値を小さくし過ぎたことによりオクルージョン発生時の急激な映像の動きが発生した。以上の要因により人物の背景に注目した手法による個人注目映像が、急なズームイン・アウトが多く、また動きが激しい部分も見られる映像になったことが、人物の背景に注目した手法が*Turn*の映像に対して見やすい映像を生成できなかった原因であると考えられる。また、評価用データを確認したところ、前述したような映像になっている個人注目映像が実際に多数確認できた。

今回の実験では前述したパラメータの設定をプロトタイプシステムの出力映像を用いて目測で行った。しかし、手法の正確な比較のためには、パラメータを定量的な評価に基づき決定した上で個人注目映像を生成する必要がある。定量的な評価の例としては、個人注目映像のズーム率の変化の回数及び速度、映像中で人物が占める領域の大きさ及び場所等が挙げられる。

また、評価用データ中の人物がカメラに接近した後離反する映像について、どの映像も人物がカメラに寄り、離れる間隔が短かった。このことも映像のズームイン・アウトの回数を増やす一因となり、人物の背景に注目した手法で生成された映像の見やすさに今回の実験では悪影響を与えている。

ズームイン・アウトがあまり発生しない*Line*、*Free*の映像に関しては、僅かながらも人物の背景に注目した手法が評価が高い。しかしこの結果もパラメータ及び評価データの見直し、改善を行い追加実験を行えばさらに評価に差が出てくることが予想できる。

結論として、前述したプロトタイプシステムへの不適切なパラメータや評価データの撮影方法の問題もあり、提案手法の長所と短所を考察できるまでの実験データは得られなかった。

6 おわりに

本論文では全方位映像から個人注目映像を生成するプロトタイプシステムの中で、映像の切り出し境界を決定する手法を見やすい映像の定義に基づいて提案し、ベースラインとの比較実験を通して長所及び短所を考察し、改善策を論じることを試みた。しかし今回行った実験のデータはそのような考察を行うには不

十分なものとなった。今後の課題として、まずプロトタイプシステムのパラメータ及び評価用データの見直し、改善を図ったうえで提案手法とベースラインについて改めて比較実験を実施することが挙げられる。また、比較実験において信頼できる実験結果を得るためには5.3節で述べたパラメータの問題以外にも、3.1節、3.3節で述べた処理を改善する必要がある。具体的には、正確でノイズがなく、オクルージョン発生時にも高精度に補間された人物の外接矩形及び重心を得るため、現在の人物抽出及び追跡手法ではなく、背景の変化及びオクルージョンの発生にロバストな手法を採用することが必要である。

個人注目映像生成プロトタイプシステムの今後の展望としては、1章において述べたように監視カメラ、会議の記録、そして幼稚園児の活動記録等がその応用例として挙げられる。その中で、幼稚園児の活動記録等、映像が観賞して楽しいものになっていることが期待されるケースでは個人注目映像が人物と周りの環境との関わりが良く分かる映像となる提案手法を採用した方が良い事が予想できる。一方、監視カメラ等、人物について安定して映像を記録でき、かつその情報量が多いことを期待されるケースでは個人注目映像が人物の位置、大きさについて安定したものとなるベースラインを採用した方が良い事が予想できる。以上のように、プロトタイプシステムが使われるケースで重要視される事柄に着目して手法を選択する場面が現れる事が考えられる。また、プロトタイプシステムが運用されるケースによっては現在の手法では利用者が最も欲しがっている情報が含まれる映像を生成できない事も予想される。そのため、特定の用途に沿った個人注目映像を切り出す手法を新たに考えることも今後の課題の1つとして考えられる。

参考文献

- [1] 川西亮輔, 山下淳, 金子透: 全方位カメラを用いた3次元環境モデリング, 画像の認識・理解シンポジウム(MIRU2008), (2008-7)
- [2] Lu Boun Vinh, 角田哲也, 川上玲, 大石岳史, 池内克史: 全方位カメラを用いた屋外複合現実感におけるオクルージョンを考慮した移動物体の検出と影の除去
- [3] 杉浦一成, 榎原靖, 八木康史: 全方位カメラを用いた複数方向の観測による歩容認証, 情報処理学会, (2007-5)
- [4] 川崎洋, 池内克史, 坂内正夫: 時空間画像解析を用いた全方位カメラ映像の超解像度化, 電子情報通信学会論文誌, D-2, Vol.J84-D-2, No.8, pp.1891-1902, (2001-8)
- [5] H.Kawasaki, K.Ikeuchi, and M.Sakauchi: Spatiotemporal analysis of omni image, CVPR, vol.2, pp.577-584, (2000-6)
- [6] 佐藤辰雄, 後藤和弘: 全方位カメラのためのパノラマ画像展開
- [7] 上條俊介, 松下康之, 池内克史, 坂内正夫: 時空間 Markov Random Field モデルによる隠れにロバストなトラッキングアルゴリズム, 電子情報通信学会論文誌, D-2, Vol.J83-D-2, No.12, pp.2597-2609, (2000-12)
- [8] 上條俊介, 坂内正夫: 時空間 MRF モデルに基づく車両と歩行者の統一的トラッキング, 情報処理学会論文誌, vol.45,

(2004-12)

- [9] 石川友哉, Yu Wang, 加藤ジェーン: 監視カメラ映像からの幼稚園児の1日ダイジェスト自動生成, 電子情報通信学会, (2010-3)
- [10] W.Wharton, D.Howorth: Principles of television reception, London, Pitman, 1971
- [11] N.Ohtsu: A threshold selection method from gray-level histograms, IEEE Trans.Sys, Man., Cyber, vol.9, no.1, pp62-66, 1979
- [12] J.Serra: Image analysis and mathematical morphology, New York, Academic Press, 1983
- [13] A.Bhattacharyya: On a measure of divergence between two statistical populations defined by probability distributions, Bulletin of the Calcutta Mathematical Society, No.35(1943), 99-109