

重複のないカメラ視野を移動する複数の歩行者の対応付け

赤塚久哉，高橋友和，井手一郎，村瀬洋
名古屋大学大学院情報科学研究科

本稿では，低コストな広域人流解析システムの実現を目的として，カメラ視野に重なりのない複数のカメラにより撮影された映像中の歩行者の追跡を行い，異なるカメラ視野間を移動する歩行者を対応付ける手法を提案する．はじめに各カメラ視野内の歩行者をパーティクルフィルタを用いて追跡する．次に追跡によって得られた位置や速度特徴を用いて，移動先のカメラ視野における出現位置・時間を推定し，対応付けに利用する．実験により，提案手法による各カメラ視野内の歩行者の追跡成功率は 71%，歩行者追跡成功時の異なるカメラ視野間における対応付け成功率は 86% となり，提案手法の有効性が示された．

1. はじめに

セキュリティシステムや，交通流解析，マーケティング，スポーツシーンの解析などを目的とした，カメラを用いた歩行者・車両などの移動物体の検出・追跡に関する研究が多く報告されている．藤本ら[1]は，店舗内に設置された全方位カメラを用いた顧客行動の解析システムを提案している．小林ら[2]は，2 地点で収集された車長・車高データから同一車両を対応付け，地点間の旅行時間を算出している．内海ら[3]は，サッカー映像において隠れが生じやすい選手領域の検出・追跡する手法を提案している．上條ら[4]は，時空間 MRF を交差点画像に適用し，隠れがある車両を追跡している．

我々は，図 1 に示すような，大学のキャンパス，大規模なイベント会場，商店街などでの利用を想定した，低コストな広域人流解析システムの実現を目指している．これが実現できれば，屋外における広域なマーケティングや交通流解析が可能になる．低コストなシステムの実現のためには，できるだけ少数のカメラを広範囲に離散的に配置し，各カメラは広角なカメラ視野を持つことが望ましい．このようにして配置された複数のカメラは，設置場所やコストの制約により，カメラ視野に重なりを持たない場合が多いことが想定されるため，異なる複数のカメラ視野間における歩行者の対応付けが必要となる．小林ら[5]は，カメラの位置や視野が既知である環境において，机などの障害物の知識を利用し，カメラで観測されない領域での歩行者の行動軌跡推定手法を提案している．池亀ら[6]は，歩行者の色ヒストグ

ラム特徴と，カメラ視野間の移動時間をベイズ則に基づいて確率的に照合し，歩行者を対応付けている．この手法では，画像中に歩行者の占める割合が比較的大きく，歩行者の特徴が安定して得られると考えられるが，我々の想定するシステムでは，カメラ視野が広範囲になるために画像中に歩行者領域の占める割合が小さくなり，歩行者の特徴を安定して得ることが困難となる．

そこで本稿では，システム実現のための要素技術として，各カメラ視野においてパーティクルフィルタを用いて歩行者領域を追跡し，異なるカメラ視野に出現する位置・時間を推定するための安定した特徴を得ることにより，重複のないカメラ視野を移動する歩行者を対応付ける手法を提案する．

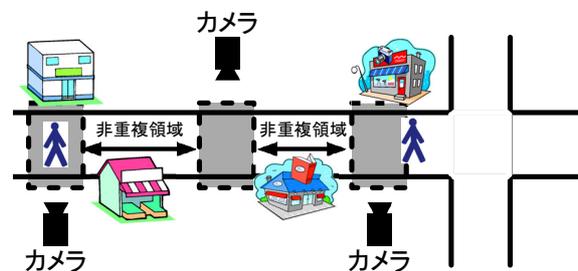


図 1. 広域人流解析システム

2. 提案手法

2.1 手法の概略

本手法は、1) 各カメラ視野で歩行者を追跡し、歩行者の特徴を抽出する段階、2) 得られた特徴を用いて異なるカメラ視野における同一歩行者の出現位置・時間を推定し、カメラ視野間で同一歩行者を対応付ける段階からなる。任意の2台のカメラ i, j を用いた処理手順を図2に示す。

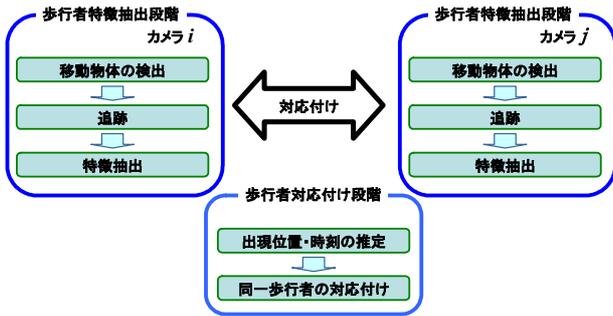


図2. 提案手法の流れ

広角なカメラで撮影された画像中を互いに隣接した複数の歩行者が移動している場合、画像中に占める歩行者領域の割合が小さいため、個々の歩行者領域の切り分けが困難となる。そこで本稿では、このような歩行者の集団（一人の場合を含む）を1つの歩行者として扱う（図3）。また、樹木や歩行者同士の交差などによる隠れの影響により、歩行者の追跡が失敗し、特徴が安定して得られない問題がある。そこで本稿では、このような状況下で優れた追跡特性を持つパーティクルフィルタを用いて歩行者領域の追跡を行い、フレームごとに抽出した特徴を用いることで安定した特徴抽出を行う。

異なるカメラ視野間における歩行者の対応付けには各カメラ視野に出現した歩行者の特徴を抽出し、それを利用する。

2.2 各カメラ視野内の歩行者特徴抽出

各カメラ視野において、パーティクルフィルタ[7]を用いた歩行者の追跡を行い、各歩行者の特徴を複数フレームから抽出することで、隠れなどによる局所的な変化に対してロバストな特徴抽出を行う。抽出する特徴は以下の3つである。

- ・ 歩行者の位置 (x, y)
- ・ 速度 $\mathbf{v} = (v_x, v_y)$
- ・ 色ヒストグラム



図3. 歩行者が集団として移動する例

2.2.1 抽出手順

- Step 1.** 事前にカメラ視野ごとに検出の対象領域を設定し、その領域を通過する歩行者の位置と大きさを背景差分法により検出する。
- Step 2.** 検出された位置から初期状態を設定して、パーティクルフィルタを用いた歩行者の追跡を行う。
- Step 3.** 追跡の結果から、各フレームにおける歩行者の速度・位置・色ヒストグラムを抽出し、それらを歩行者の特徴とする。

2.2.2 パーティクルフィルタ

パーティクルフィルタは、モンテカルロフィルタ、CONDENSATIONとも呼ばれ、ある状態から次の状態への遷移を推定する際、次の状態を表すパーティクルを状態遷移確率密度に従い遷移させ、観測された状態から次の状態を確率的に決定する手法である[7]。今、時刻 t における観測時系列を $Z_t = (\mathbf{z}_1, \dots, \mathbf{z}_t)$ とする。観測値から得られる状態ベクトルを \mathbf{x}_t とする。現在の状態の遷移は過去の状態遷移の履歴には影響されないというマルコフ連鎖の考え方に基づき、事前確率密度関数 $p(\mathbf{x}_t | Z_{t-1})$ は時刻 $t-1$ の事後確率密度関数 $p(\mathbf{x}_{t-1} | Z_{t-1})$ と時刻 $t-1$ から t への状態遷移確率密度関数 $p(\mathbf{x}_t | \mathbf{x}_{t-1})$ を用いて式(1)のように表すことができる。

$$p(\mathbf{x}_t | Z_{t-1}) = \int p(\mathbf{x}_t | \mathbf{x}_{t-1}) p(\mathbf{x}_{t-1} | Z_{t-1}) d\mathbf{x}_{t-1} \quad (1)$$

ここで、時刻 t における事後確率密度関数はベイズの法則に従い式(2)で表すことができる。

$$p(\mathbf{x}_t | Z_t) \propto p(\mathbf{z}_t | \mathbf{x}_t) p(\mathbf{x}_t | Z_{t-1}) \quad (2)$$

パーティクルフィルタでは、 t における確率密度関数 $p(\mathbf{x}_t | Z_t)$ を、状態ベクトルのパーティクル群

と各パーティクルに対応する重みによって離散的に近似する。

パーティクルフィルタを用いた追跡処理は以下で述べる「観測」、「選択」、「予測」という3つの操作を繰り返すことによって行われる。

- 選択

時刻 $t-1$ にて、 Z_{t-1} から得られた $p(\mathbf{x}_{t-1} | Z_{t-1})$ が、 N 個の重み付きパーティクル群で表されているとき、各パーティクルの重みに比例した確率に従いパーティクルを選択する。つまり、重みが低いパーティクルは消滅する可能性が高く、重みが高くと増殖する可能性が高い。また、選択後の各パーティクルの重みはすべて等しく設定する。

- 予測

選択された各パーティクルを $p(\mathbf{x}_t | \mathbf{x}_{t-1})$ に従い伝搬させ、 $p(\mathbf{x}_t | Z_{t-1})$ に相当する t における N 個のパーティクル群を生成する。

- 観測

パーティクルの重みを尤度評価により推定し、各パーティクルの重みの合計が 1 になるように正規化をする。

本手法において、 \mathbf{x}_t は時刻 t におけるパーティクル群の位置を、観測量 \mathbf{z}_t は t におけるカメラ視野中の歩行者の位置を意味し、式(3)で表される。

$$\begin{cases} \mathbf{x}_t = \{ x_t^{(n)}, y_t^{(n)}, n = 1, \dots, N \} \\ \mathbf{z}_t = (\hat{x}_t, \hat{y}_t) \end{cases} \quad (3)$$

また、各パーティクルを重みづけるための尤度評価には、他の歩行者と区別するために、時刻 $t-1$ における歩行者領域中の色ヒストグラムと、 t において予想される歩行者領域中の色ヒストグラムの類似度を歩行者ごとに尤度として与える。HSV 色空間を用い、照明変化の影響を減らすために明度を無視し、色相と彩度成分のみを用いる。

$p(\mathbf{x}_t | \mathbf{x}_{t-1})$ は式(4)で与える。

$$\begin{cases} x_t = x_{t-1} + dx_{t-1} + N(\sigma_x) \\ y_t = y_{t-1} + dy_{t-1} + N(\sigma_y) \end{cases} \quad (4)$$

ここで、 dx_{t-1} 、 dy_{t-1} は $t-1$ におけるパーティクル

の速度、 $N(\sigma_x)$ 、 $N(\sigma_y)$ はそれぞれ平均 0、標準偏差 σ_x 、 σ_y のガウシアンノイズである。

2.2.3 初期状態の決定

パーティクルフィルタを用いて物体を追跡する場合、初期状態を与える必要がある。1つの物体を追跡する場合、初期状態は時刻 $t = 0$ において一様乱数を用いてパーティクルを散布することによって与える場合が多い。しかし、本研究で想定するように、歩行者が出現・退出を繰り返し、歩行者数も不定である場合、 $t = 0$ においてパーティクルをランダムに散布すると、異なる歩行者の色ヒストグラム類似度が偶然高くなる場合があり、追跡失敗の原因となる。そこで、各歩行者を検出し、追跡開始時間・位置を決定し、それをもとに歩行者ごとに初期状態を与える。

本稿では、図4のような屋外での2台のカメラを用いた歩行者追跡を想定し、道路に沿った成分を x 、垂直成分を y とする。そのため、各カメラ視野に歩行者が新たに出現する位置は画像の周辺領域に限られる。カメラ視野ごとに検出の対象領域を事前に設置し、背景差分法により歩行者領域を検出することで、パーティクルフィルタの初期状態を自動的に決定する。設定方法を以下に述べる。移動物体の存在しない2枚の画像を取得し、一方を背景画像としてRGBそれぞれ独立に差分を求め、次にそれぞれ標準偏差 σ_R 、 σ_G 、 σ_B を求め、それらを定数倍した値を背景差分値の閾値とする。

$$\theta_{color} = \alpha \sigma_{color} \quad (5)$$

ここで α は定数を示し、 $\sigma_{color} \in \{\sigma_R, \sigma_G, \sigma_B\}$ とする。日照の時間変化を考慮して、一定時間ごとに背景画像と θ_{color} を更新する。

歩行者検出の際には、RGB すべてにおいて背景差分値が閾値以上の場合移動物体と判定する。ただし、検出された領域がある大きさ以下ならば、雑音領域として除去する。

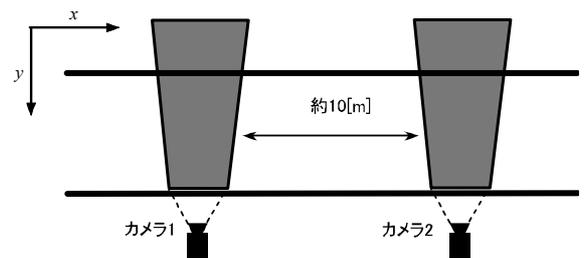


図4. 撮影環境

2.3 カメラ視野間における歩行者対応付け

異なるカメラ視野間で歩行者の対応付けを行うため、一方のカメラ視野内の歩行者の移動経路と速度から、同一の歩行者がもう一方のカメラ視野に出現する時刻と位置を推定する。

本稿では、カメラの位置関係、カメラ視野間の距離が既知であるとする。そのため、対象とする歩行者がカメラ視野から退出する時刻とカメラ視野を歩行する速度のばらつきから、異なるカメラ視野内に出現する時刻が推定できる。また、位置の y 方向成分のばらつきから歩行者の出現する位置が推定できる (式 (6))。

推定後は、推定結果から初期状態を決定し、同様にパーティクルフィルタを用いた追跡を行う。このとき、最初の尤度評価に以前のカメラ視野内で得られた歩行者領域の色ヒストグラムの平均との類似度を用いることで、異なるカメラ視野間における歩行者の対応付けを行う。

$$\begin{cases} t_{next} = t_{now} + D / |\mathbf{v}| + \mathbf{N}(\sigma_{|\mathbf{v}|}) \\ y_{next} = \mathbf{N}(\sigma_y) + C \end{cases} \quad (6)$$

ここで、 t_{now} は歩行者が一方のカメラ視野内から退出した時刻、 t_{next} はもう一方のカメラ視野内に出現する時刻、 D はカメラ視野間の距離、 y_{next} は一方のカメラ視野内に出現する位置、 $\sigma_{|\mathbf{v}|}$ は速度ベクトルの大きさの標準偏差、 C は定数、 σ_y は位置の y 方向成分の標準偏差を表す。

3. 実験

3.1 実験条件

実験環境下において 2 台のカメラを用いて撮影された映像に対して本手法を適用し、実験を行った。

図 4 に実験で用いた撮影環境を上から見た概略図を示す。図 5 はカメラの配置関係を示すために、撮影環境周辺を撮影したものである。カメラは上から見下ろす形で設置し、実線で囲まれた領域は、それぞれカメラ 1、カメラ 2 の視野を示す。カメラ視野中の斜線領域は、歩行者検出領域を示す。カメラ視野間には平均約 10[m] の重なりのない領域が存在する。実験は以下の条件で行い、歩行者の追跡及び対応付け精度を調べた。

- ・ 撮影映像 : 2 台のカメラを同期して撮影した 10 分の映像セット 8 つ
- ・ 撮影時間帯 : A.M. 10:00 ~ P.M. 5:00
- ・ カメラ : CCD-V17 (SANWA 製)
- ・ 映像のフレームレート : 10 frames/sec

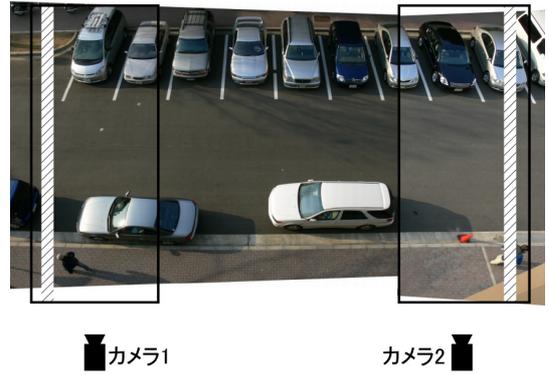


図 5. 実験環境概観

- ・ カメラ視野の画像サイズ : 120×320 pixels
- ・ 背景画像の更新間隔 : 1 分 (600 frame)
- ・ パーティクル数 : 100
- ・ 予測モデル(式 4)で用いる標準偏差 σ_x , σ_y : 10 pixels (約 0.4 m)
- ・ 色ヒストグラムのビン数 : 396 (色相 36×彩度 11 段階)
- ・ 閾値設定パラメータ α (式 5) : 10
- ・ 色類似度算出手法 : ヒストグラムインターセクション(式 7)

$$S_{HM} = \sum_{i=1}^I \min(H_i, M_i) \quad (7)$$

H , M を正規化ヒストグラムとしたとき、その類似度 S_{HM} は 0 から 1 までの値をとる。

3.2 各カメラ視野内における歩行者追跡結果

2 台のカメラの両方で観測された歩行者は 177 組であった。どちらか片方のカメラでのみ観測された歩行者は目的とする歩行者対応付けが行えないため除外した。評価は目視で行い、各歩行者の位置を正しく追跡できていたら成功とした。

追跡結果を表 1 に示す。追跡成功率は 71% と良好な結果を示した。失敗の原因の大半を占めた、歩行者領域検出漏れの原因は以下の 3 つが考えられる。

1 つ目は、背景と類似した色特徴を持つ歩行者が存在し、背景差分による検出が困難となったためである (図 6 (a))。2 つ目は、背景画像の推定方法と関連するものである。背景画像の更新は 1 分に一度行っており、この間に風による樹木の揺らぎ、停車車両の移動などによる大きな背景の変化が発生したためであると考えられる。3 つ目は、検出の際に歩行者の一部が停車中の車などにより遮蔽されたことによる。この場合、検出される領域が小さいため、雑音領域として除去されたと考えられる (図 6 (b))。

歩行者領域追跡失敗は、歩行者検出の際に歩行者同士の交差が発生し、正しい歩行者の色ヒストグラムが取得できなかったために生じたと考える。

表 1. 歩行者追跡結果

歩行者領域追跡成功	125
歩行者領域追跡失敗	1
歩行者領域検出漏れ	51
合計	177



(a) 歩行者が背景と類似した色特徴を持つ場合



(b) 歩行者に隠れが生じた場合

図 6. 歩行者領域検出漏れの原因

3.3 カメラ視野間における歩行者対応付け結果

表 2 に一方のカメラ視野で追跡が成功した歩行者に関する対応付け結果と対応付け失敗の要因をまとめる。歩行者対応付け成功例を図 7 に示す。ラベル 35 の歩行者が正確に対応付けていることが確認できる。図中の矢印は、歩行者の進行方向を表すために対応付け結果に人手で付け加えたものである。

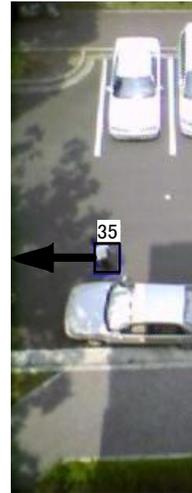
表 2. 歩行者対応付け結果

歩行者対応付け成功	108
対応付け誤り	11
出現位置推定誤り	5
出現時刻推定誤り	1
合計	125

位置推定誤りは、歩行者がカメラ視野間の非撮影領域において道路と垂直方向に大きく移動することで発生した(図 8)。カメラ 1 で観測された歩行者 45 は、カメラ 2 では推定位置より上部に出現している。

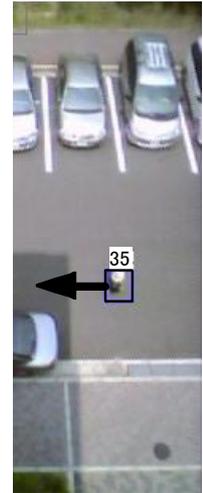
時刻推定誤りは、カメラ視野間の非撮影領域において歩行者が停止するなどして、移動速度が大きく変化した場合に発生した。これは、実験で歩行者がカメラ視野外でもカメラ視野内と同様の速度で歩行

2677



カメラ 1

2550



カメラ 2

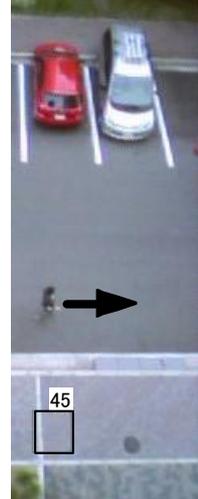
図 7. 歩行者対応付け成功例(画像上部の数字はフレーム番号を示し、画像中の数字は歩行者のラベルを示す)

3053



カメラ 1

3176



カメラ 2

図 8. 出現位置推定誤り例(カメラ 2 中のラベル 45 の矩形は、推定位置を表している。図中の矢印は図 7 と同様に歩行者の進行方向を示す。)

すると仮定したためである。

異なる歩行者への対応付け誤りは、対応付け候補が複数存在した場合かつ、異なる歩行者の色ヒストグラム間の類似度が高い場合に発生することが多かった。また、類似した色特徴を持つ車両など、ほかの移動物体に対応付く場合もあった(図 9 (a) 歩行者ラベル 23)。また、樹木の影の揺らぎなどの背景の一時的な変化により、その影領域に対応付く場合もあった(図 9 (b) 歩行者ラベル 256)。

4. おわりに

本稿では、重複のない複数のカメラ視野間を移動する歩行者の対応付け手法を提案した。提案手法は、パーティクルフィルタを用いた隠れや交差にロバストな歩行者追跡により、安定した歩行者特徴の抽出を行い、それらの特徴を用いてカメラ視野移動後の歩行者の出現位置・時刻の推定を行うことで異なるカメラ視野間における歩行者の対応付けを行うものである。2台のカメラを用いた実環境下での実験により、提案手法の有効性を示した。

各カメラ視野における歩行者追跡に関して、歩行者検出の対象領域の大きさや背景画像更新頻度を調整することにより精度低下の主要原因の検出漏れを減少させることができると考える。カメラ視野間における歩行者対応付けに関して、本稿ではカメラ視野中にて取得した特徴のみを用いて歩行者の出現位置・時刻の推定を行ったが、今後はカメラ視野間の非撮影領域を移動する歩行者の行動パターンを学習することによって、より高精度な推定を目指す。

謝辞

日頃より熱心に御討論頂く名古屋大学村瀬研究室諸氏に感謝する。本研究の一部は、日本学術振興会科学研究費補助金、21世紀COEプログラム「社会情報基盤のための音声・映像の知的統合」による。本研究では、画像処理にMISTライブラリを使用した。(http://mist.s.m.is.nagoya-u.ac.jp/)

文献

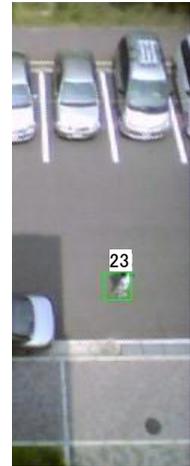
- [1] 藤本喜人, 小原ゆう, 柴田史久, 馬場口登, 八木康史, “店舗内に設置した全方位視覚センサによる顧客の行動解析”, 情処学研報 CVIM2004-142, Jan. 2004.
- [2] 小林弘忠, 石関隆幸, 遠藤雅也, 今井浩, 西村茂樹, 下浦弘, 天目健二, “2 地点車両観測情報からの全域的交通流解析システム”, 情処学研報 SIG-ITS2000, Mar. 2000.
- [3] 内海慎久, 浜田玲子, 井手一郎, 坂井修一, 田中英彦, “色頻度・局所エッジ成分を総合的に用いた物体検出手法とサッカー映像における選手検出への応用”, 信学技報 PRMU2001-241, Feb. 2002.
- [4] 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車両トラッキング”, 信学論(D-II), Vol.J83-D-II, No.12, pp.2597-2609, Dec. 2000.
- [5] 小林貴訓, 佐藤洋一, 杉本晃宏, “疎分散カメラ群を用いた人物行動軌跡の推定”, 画像の認識・理解シンポジウム(MIRU)2006 講演論文集, pp.1213-1218, July 2006
- [6] 池亀幸久, 平野真, 玉木徹, 山本正信, “視野重複のないマルチカメラを用いた人物の確率的歩行経路推定”, 信学技報 PRMU2004-186, Feb. 2005.
- [7] Michael Isard and Andrew Blake, “CONDENSATION - Conditional Density Propagation for Visual Tracking”,

2677



カメラ 1

2550



カメラ 2

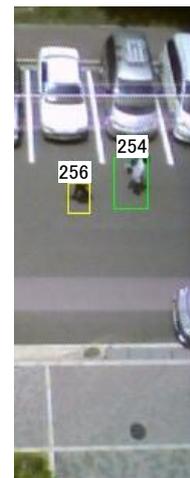
(a) 類似した色特徴を持つ移動物体に誤対応

5095



カメラ 1

4963



カメラ 2

(b) 影領域に誤対応

図 9. 対応付け誤り例