

事例データに基づく人物領域の抽出

山本文香^{†1} 岩井儀雄^{†1} 石黒浩^{†1}

動画像から人などの移動物体を抽出するために背景差分法が広く用いられているが、背景色と同じ色を持つ物体領域を抽出できないという本質的問題が存在する。本稿では、背景差分により得られる人物領域に欠損が生じた画像に対して、事例データを用いることで欠損部を補完し、安定して人物領域を得る手法を提案する。事例データとして、人物領域に欠損を含んだパターンと欠損が補完された完全なパターンを組にしたものを使用する。そして、背景差分により人物領域に欠損が生じた場合に、事例データベースから似た欠損パターンを探し出し、その欠損パターンと組になっている完全な人物領域パターンを用いることで、人物領域の補完を行う。我々はこの補完処理を、入力画像上の欠損と人物領域パターンの関係をブロック単位で定式化し、最適化することで実現する。本手法を実画像に適用して得られる推定完全画像とその評価について報告する。

Memory-Based Object Detection

AYAKA YAMAMOTO,^{†1} YOSHIO IWA^{†1}
and HIROSHI ISHIGURO^{†1}

Background subtraction is widely used for detecting moving objects; however, color similarity between a background and a moving object is still an important problem. In this paper, we present a memory-based approach to efficiently detect moving objects that are similar in color to the background. Our database is constructed from couples of missing and complete human silhouette patterns; the missing parts obtained by background subtraction are detected using missing patterns, and then complemented by complete patterns. In our approach, the relationship among the missing parts and the human silhouette patterns is formulated and optimized in a block-based manner. The experimental results and evaluations of our system are demonstrated.

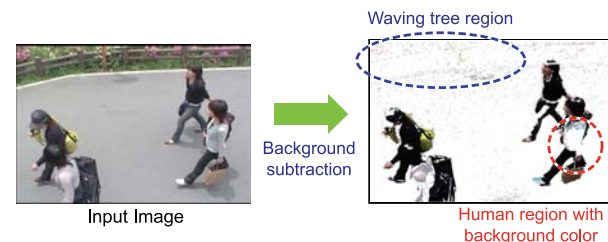


図1 従来法³⁾の問題点
Fig.1 Problems of the previous method³⁾

1. はじめに

動画像からの移動物体検出は、人物の追跡や行動認識などの様々なコンピュータビジョン技術において重要な役割を担っている。そのため、より正確に、またより高速に移動物体領域を取り出すための研究が盛んに行われている。画像中から人などの領域を抽出する手法の1つに背景差分法がある。背景差分法は、背景が変化しない静的な場合に、安定して移動物体領域を取り出すことができる。しかし、屋外環境下では照明変動や物体の影の影響を受けて、物体領域を過検出する恐れがある。また、背景の色と移動物体の色が似ている場合に、移動物体を検出できないという問題が存在する。

そこで、屋外環境で安定して移動物体領域を検出するために、映像の背景部分を推定する手法が数多く報告されている。多くの研究では、画素の時間的な変化に対して、混合正規分布などの分布モデルを仮定して、統計的に判断を下している¹⁾²⁾。我々も、太陽直射光と環境光を組み合わせた2色反射のモデルを明るさ可変背景モデルとして提案し、屋外において明るさの変動にロバストな人物領域の抽出を実現している³⁾。この手法では、太陽直射光と環境光の色成分をリアルタイムで推定しながら背景画像を生成できるため、高速に人物領域を抽出することができる。しかし、カメラ映像中の背景が動くことを仮定していないため、草木の揺れなどの微小な変動が生じると移動物体として検出してしまう欠点がある(図1)。また、背景画像と同じ色を持つ物体領域を抽出できないという背景差分法の本質的問題がなお存在する。この2つの問題はトレードオフの関係にあり、草木の揺れなどに対応しようと安易に背景の範囲を広げると、背景色が多くなり、移動物体の検出率が下がってしまう。よって、安定して物体領域を抽出するためには、物体領域に関する何らかの事前知識が必要であると考えられる。

^{†1} 大阪大学大学院基礎工学研究科システム創成専攻
Graduate School of Engineering Science, Osaka University

物体領域に関する事前知識として、物体の特徴的な部位間に成り立つ幾何的制約を事前に学習し、検出された部位から全体像を推定する手法が提案されている^{4),5)}。しかし、これらの手法では幾何的制約により推定できる物体や姿勢が限定されるといった欠点がある。本研究では、事前知識として、人物領域に欠損が発生したパターンと欠損が補完された完全な人物領域パターンを組にしたものを使用する。そして、背景差分により人物領域に欠損が生じた場合に、事例データベースから似た欠損パターンを探し出し、その欠損パターンと組になっている完全な人物領域パターンを用いることで、人物領域の補完を行う。Collinsら⁶⁾も同様に、人物領域パターンを用いて画像中の歩行者を認識しているが、人物領域全体を1つのテンプレートとして扱っているため、人物の多様なシルエットに対応するためには大量の事例データが必要となってしまう。そこで、我々は欠損・完全パターン画像、及び入力画像を重複しないブロックに分割し、このブロック単位で上記の補完処理を行う。このように、人物領域パターンを分割して扱うことで、有限の事例データから多様なシルエットを表現することが可能となる。

画像データベースから所望の画像を構成するために、近年、画素ブロック間の隣接関係をマルコフ確率場などを用いて定式化し、グラフカット⁷⁾や信念伝播法(BP)⁸⁾により効率的に最適化する手法が数多く報告されている⁹⁾⁻¹¹⁾。Rotherらはフォトモンタージュの自動作成⁹⁾を、またKimらは人物を含んだ画像列から背景を推定する問題¹⁰⁾をそれぞれ多値のラベル付け問題として定式化し、グラフカットにより最適化している。またCaoらは、手書き文書のグレースケール画像から罫線とノイズを取り除いた二値画像を作成するために、文字を構成するパッチ間の隣接関係を学習し、BPにより二値画像を推定している。本研究でも、入力画像上の欠損と人物領域パターンの関係をブロック単位で定式化し、最適化することで、画像データベースから人物領域の欠損を補完した完全な人物領域を得る。以後、提案手法の詳細について述べる。

2. 最小化問題としての人物領域抽出

本章では、背景差分により生じた人物領域の欠損を事例データを用いて補完する問題を最小化問題として定式化する。まず、本研究で用いる事例データベースの構築方法について説明する。そして、補完により得られる推定完全画像を評価する目的関数を定義した後、推定完全画像を最適化する方法について述べる。

2.1 事例データベース

本手法で用いる事例データベースは、人物領域欠損パターン（以下、欠損パターンと記述

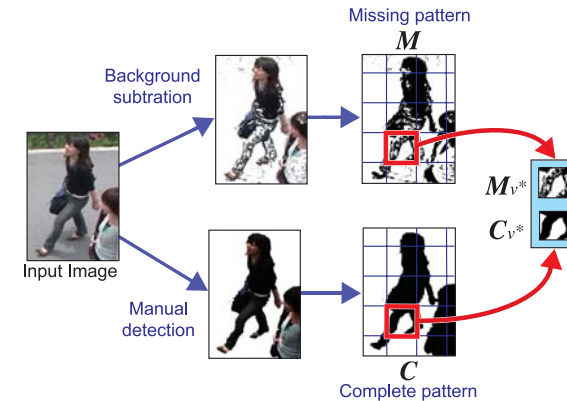


図2 事例データの作成方法
 Fig.2 Exemplar creation process

する)と人物領域完全パターン（以下、完全パターンと記述する)の組み合わせが登録される。データベースの検索キーは、背景差分により得られた欠損パターンである。この検索キーを用いて入力画像の人物領域と近い事例パターンを探し出し、その完全パターンと置き換えることで補完を行う。

入力画像と背景画像との差分により得られる人物領域の二値画像を M 、同じ入力画像から人物領域を手動抽出した二値画像を C とし、それぞれ $B \times B$ 画素の重複しないブロック $v \in \mathcal{V} = \{1, 2, \dots, N\}$ に分割する。ここで、サイト $v^* \in \mathcal{V}^* = \mathcal{I} \times \mathcal{V}$ を入力画像の通し番号 $i \in \mathcal{I}$ とブロックの位置 v で定義する。このとき、同じサイト v^* に対する欠損パターン画像上のパッチ M_{v^*} と完全パターン画像上のパッチ C_{v^*} を組にして事例データベースに登録する。事例データ作成の流れを図2に示す。この作業を大量の入力画像に対して行い、事例データベースを構築する。

2.2 定式化

事例データベースを用いて入力画像上の欠損を適切に補完し、完全な人物領域を得るには、入力画像の欠損情報から最も適した事例データを選び出す必要がある。そこで、補完により得られる推定完全画像を評価する目的関数を導入する。この目的関数を最小化することで、最適な推定完全画像を得る。

観測画像と背景画像との差分により得られる二値画像を I とし、 $B \times B$ 画素の重複しないブロック $v \in \mathcal{V}$ に分割する。ここでの問題は、各ブロック v に対して、入力画像パッチ I_v

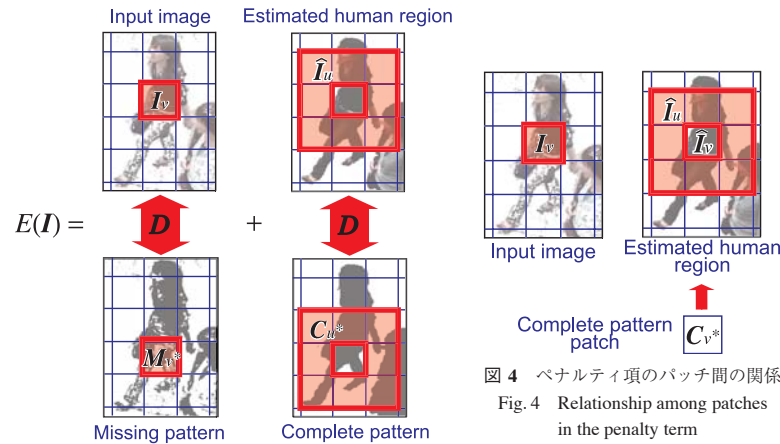


図3 目的関数
Fig.3 Objective function

の欠損を補完する完全パターンパッチ C_{v^*} を割り当てた推定完全画像 \hat{I} を生成することである。このとき、生成された推定完全画像 \hat{I} を評価する目的関数は次のように定義される

$$E(I) = \sum_{v \in V} [D(I_v, M_{v^*}) + \kappa \sum_{u \in N(v)} D(\hat{I}_u, C_{u^*})] \quad (1)$$

ここで、 κ は定数、 $u \in N(v)$ は v の近傍ブロック、 $u^* \in N(v^*) = \{(i, u); u \in N(v)\}$ はサイト v^* の近傍ブロックをそれぞれ表す。ここでの近傍 $N(v)$ は v 自身を含まないとする。また、 $D(\cdot, \cdot)$ ($0 \leq D \leq 1$) はパッチ間の距離（類似度）を表す。

式(1)において、第1項はデータ項であり、入力画像パッチ I_v と欠損パターンパッチ M_{v^*} の距離が加えられる。つまり、データ項を最小化することは、入力画像パッチ I_v の欠損と近い欠損パターン M_{v^*} を見つけることになる。

第2項は平滑化項であり、推定完全画像上の v の近傍パッチ \hat{I}_u と完全パターン画像上の v^* の近傍パッチ C_{u^*} との距離に応じて κ が加えられる。ここでは、推定完全画像上の隣接したブロックのパターンからブロック v に割り当たる可能性の高い完全パターン C_{v^*} を見つけることを意図する。式(1)のパッチ間関係を図3に示す。

平滑化項の効果を強めることでより滑らかな人物領域を得ることができるが、一方で、手足のような人物領域の境界部では、隣接する背景ブロックからの影響を受けて欠損を補完できない恐れがある。そこで、人物領域の境界部において、背景を表す事例データを割り当て

ようとした場合に次のようなペナルティ g を式(1)に加えることを考える

$$E(I) = \sum_{v \in V} \left[D(I_v, M_{v^*}) + \kappa \sum_{u \in N(v)} D(\hat{I}_u, C_{u^*}) + g(I_v, C_{v^*}) \right], \quad (2)$$

$$g(I_v, C_{v^*}) = \lambda \frac{|I_v|}{B^2} \quad \text{if } |\hat{I}_u| = 0, u \in N(v) \text{ and } |C_{v^*}| = 0. \quad (3)$$

ここで、 λ は定数、 $u \in N(v)$ は v の近傍ブロックをそれぞれ表す。また、 $|\cdot|$ はパッチ内の人物領域のピクセル数を表す。ここでは、推定完全画像上のブロック v の近傍パッチ \hat{I}_u の少なくとも1つが人物領域ピクセルを含まない場合、 v を人物領域の境界部を含むブロックと判断している。ここに人物領域を含まない完全パターンパッチ C_{v^*} を割り当てる場合に、入力画像パッチ中の人物領域のピクセル数に応じてペナルティ λ が加えられる。したがって、人物領域の境界部を含むブロックに人物領域を含まない完全パターンパッチが割り当たることを避けることが可能となる。ペナルティ項(3)のパッチ間関係を図4に示す。

2.3 最小化

2.2で定義した目的関数を最小化する方法について説明する。最小化の過程では効率的に最小値に達することが望まれており、近年では、グラフカット^{9),10)}やBP¹¹⁾を利用する手法が報告されている。本稿では、各ブロック v ごとにデータベース内を全探索し、目的関数の値をより小さくする事例データを選び出す方法をとる。この手続きを、端のブロックから順番に停止条件を満たすまで行い、目的関数の値を最小化するパッチの組み合わせを決定的に見つけ出す。

3. 事例データに基づく人物領域検出システム

提案システムの概要を図5に示す。本研究では、入力画像として、観測画像と明るさ可変背景モデル³⁾との差分により得られる二値画像を用いる。まず、入力画像 I から推定完全画像 \hat{I} の初期画像を生成する。推定完全画像 \hat{I} の初期画像が得られると、各ブロック v について、事例データベースから目的関数の値をより小さくする事例データを検索する。全てのブロック v で完全パターンパッチ C_{v^*} の割り当てを更新すると、その時点での推定完全画像 \hat{I} に対して目的関数の値 $E(I)$ を算出する。この目的関数の値 E が終了条件を満たすまで、推定完全画像 \hat{I} の更新を繰り返す。そして、推定完全画像 \hat{I} の更新処理が終了した時点で、これまで得られた解候補のうち目的関数の値が最も小さいものを最良解として出力する。

3.1 推定完全画像の初期化

入力画像 I の人物領域の情報を基に、推定完全画像 \hat{I} の初期画像を生成する。各ブロッ

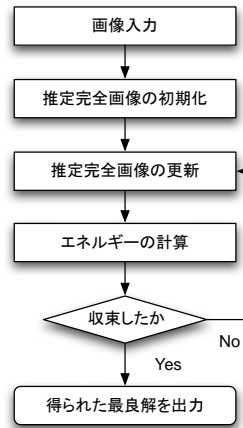


図5 システムの流れ
 Fig. 5 System overview

ク v で、式(4)を満たす欠損パターンパッチ M_{v^*} を見つけ出し、その組となっている完全パターンパッチ C_{v^*} を割り当てる

$$\operatorname{argmin}_{v^*} D(I_v, M_{v^*}). \quad (4)$$

同じ値をとる欠損パターンパッチ M_{v^*} が複数存在した場合は、人物領域の情報をできるだけ残すために、組となっている完全パターンパッチ C_{v^*} 中の人物領域のピクセル数が多い方を選択する。

3.2 推定完全画像の更新

推定完全画像 \hat{I} の初期画像が得られると、目的関数の値をより小さくするために各ブロック v への完全パターンパッチ C_{v^*} の割り当てを更新する。ブロック空間 \mathcal{V} の端のブロック v から順に式(5)を満たす事例データを見つけ出し、その完全パターンパッチ C_{v^*} を割り当てる

$$\operatorname{argmin}_{v^*} D(I_v, M_{v^*}) + \kappa \sum_{u \in \mathcal{N}(v)} D(\hat{I}_u, C_{u^*}) + \lambda \frac{|I_v|}{B^2}. \quad (5)$$

ここでは、同じ値をとる事例データが複数存在した場合に、第2項の値がより小さい方を選択する。

3.3 更新処理の終了条件

更新された推定完全画像 \hat{I} に対して、式(2)により目的関数の値 $E(I)$ を算出する。この値が更新前の値 E_{pre} と等しい場合、あるいはこれまでの最小値 E_{min} に等しい場合に、推定完全画像をこれ以上に改善できないと判断し、更新を終了する。また、反復回数の上限 k_{max} に達した場合にも更新を終了し、目的関数の値が最小値 E_{min} であったときの推定完全画像を最良解として出力する。

4. 実験

提案システムの性能を、ペナルティ項の効果とブロックサイズの観点から評価した。まず、ペナルティ項の有無による推定完全画像の差を評価し、次に、ブロックサイズを $B = 24, 32, 40$ と変化させた際に得られる推定完全画像を比較・評価した。本実験で用いた事例データベースは、図7に示すような人物一人の欠損パターン画像とその完全パターン画像それぞれ170枚の計340枚で構成される。画像は全て解像度 720×486 pixelの二値画像を用いた。実験に用いた動画の撮影時は晴れて風が吹いており、背景に草木の揺れが見られた(図6)。このため、明るさ可変背景モデル³⁾との差分により得られる人物領域には、背景色による欠損と草木の過検出が見られる(図7上段)。また、画面横方向に歩く人が観測され、画像上ではほぼ同じ大きさの人物領域が見られる。パッチ間の距離(類似度) $D(\cdot, \cdot)$ には、正規化ハミング距離を用いた。人物領域の検出結果においては、背景と識別された領域を白色で、人物と識別された領域を黒色でそれぞれ表示している。

4.1 推定完全画像の評価基準

本システムから得られる推定完全画像 \hat{I} を評価する際に、入力画像に対する真の人物領域は未知であるので、入力画像 I 中の人物領域と比較することで評価する基準を導入する

$$f(s, l, m, n) = \frac{1}{2} \left(\frac{l}{s-m} + \frac{n}{m} \right). \quad (6)$$

ここで、 s は画像の総ピクセル数(本稿では $s = 720 \times 486$)、 m は推定完全画像上の人物領域のピクセル数をそれぞれ表す。また、 l は入力画像上では人物領域であったが、補完処理後に背景領域となったピクセルの数、 n は背景領域であったが、人物領域と識別されたピクセルの数をそれぞれ表す。これら4つのパラメータの関係を表1に示す。

式(6)は、補完処理後で属している領域が前と異なるピクセルが多い場合に、より大きな



図 6 撮影場の背景
Fig. 6 Scene background



図 7 事例データの例；(上) 欠損パターン画像 (下) 完全パターン画像
Fig. 7 Examples of exemplars; (top) missing pattern images
(bottom) complete pattern images

表 1 評価式のパラメータの関係

Table 1 Parameter relationship of the evaluation formula

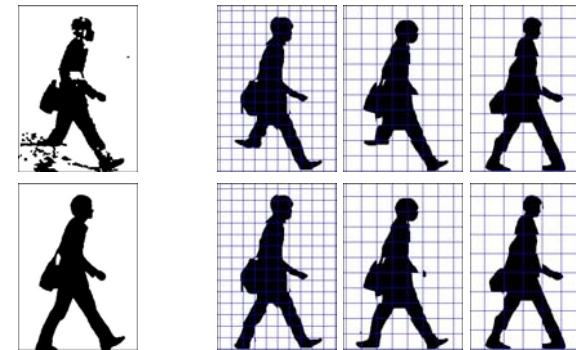
		入力画像 I		
		人物領域 (pixel)	背景領域 (pixel)	全体 (pixel)
推定完全画像 \hat{I}	人物領域 (pixel)	$m - n$	n	m
	背景領域 (pixel)	l	$s - m - l$	$s - m$
	全体 (pixel)	$l + m - n$	$s - m + n - l$	s

値をとる。ここでは、本システムを通じて補完がなされれば常に真の人物領域に近づくことと仮定し、より多くのピクセルの識別結果が変化した場合に良い評価をすることとしている。

4.2 ペナルティ項の効果

ペナルティ項が有る場合と無い場合の推定完全画像を比較し、ペナルティ項の有効性を検証する実験を行った。実験に用いた入力画像と、手動抽出により得た目標の人物領域画像、ブロックサイズ B を 24, 32, 40 の 3 通りに変化させてペナルティ項が有る場合と無い場合に得られる推定完全画像を、人物領域部分を拡大して図 8 に示す。図 8 の目標完全画像において、人物領域の総画素数は 21345 画素であり、入力画像では人物領域 3285 画素が欠損し、背景領域 3360 画素が過検出されていた。表 3 は、図 8 に示した場合の人物領域ピクセル数と過検出ピクセル数、欠損ピクセル数、評価値を示す。また図 13 に、ブロックサイズ B を 3 通りに変化させた場合の評価値をグラフに示す。

表 3 より、ペナルティ項を加えることで人物領域の欠損の減少しているのが見て取れる。このことは、図 8 の画像左下の脚部からも見て取れる。また表 3 より、人物領域を過検出したピクセル数の増加が見られるが、これは草木の揺れのようなノイズに対しても背景パ



(上) 入力画像
(下) 目標完全画像
(1) $B = 24$ (2) $B = 32$ (3) $B = 40$

図 8 ペナルティ項の有無による推定完全画像の比較；(上) ペナルティ無し (下) ペナルティ有り
Fig. 8 Comparison of estimation results; (top) without penalties (bottom) with penalties

ターンを割り当てるとペナルティがかかるため、大きなノイズ領域が人物領域として残ったことが原因と考えられる。しかし、欠損部の減少と比較するとその数は微小であり、また縮小・拡大などの事後処理で取り除くことができると考えられる。図 13 から、ペナルティ項を用いた場合に評価値が高いことが見て取れる。これら結果から、ペナルティ項を用いることで、手脚などの人物領域の境界部が隣接する背景領域の影響を受けて欠損するのを防げることがわかる。また図 8 を見ると、 $B = 32, 40$ の場合に、本来とは異なる方向を向いた脚パターンが用いられているが、これは本手法が近傍のブロックとの関係のみを用いており、人物領域全体の情報を踏まえていないためである。

4.3 ブロックサイズの評価

最適なブロックサイズ B の決定方法を検討するために、更に 2 つの場面に対して実験を行った。まず、複数の人物が重なったパターンが入力された場合に、その人物領域を適切に補完できるかを調べた。次に、大きさが異なる 3 つの人物領域パターンを入力し、人物領域の大きさとブロックサイズ B の関係を調べた。2 つ目の実験では、新たに撮影した動画画像を用い、事例データベースにもこの動画画像から欠損・完全パターン画像を各 10 枚、計 20 枚を加えた (図 10)。この動画画像の撮影時は晴れで (図 9)、画面の前方あるいは後方へ歩く人が観測された。このため、画像上では様々な大きさの人物領域が見られる。

各場面について、実験に用いた入力画像と、手動抽出により得た目標の人物領域画像、ブ



図 9 撮影場の背景
Fig.9 Scene background

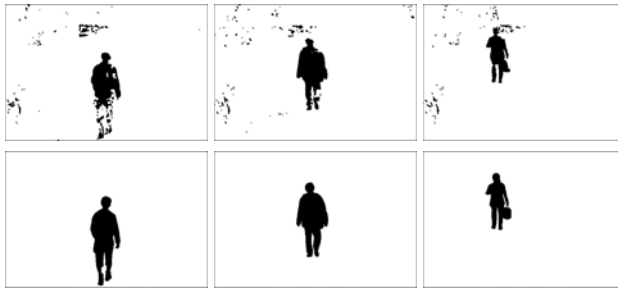


図 10 事例データの例; (上) 欠損パターン画像 (下) 完全パターン画像
Fig.10 Examples of exemplars; (top) missing pattern image (bottom) complete pattern image

表 2 目標完全画像の人物領域と入力画像の欠損/過検出領域
Table.2 human regions from ideal images and under/over-detected regions from test images

	(a)	(b)	(c)	(d)	(e)
目標完全画像の人物領域 (pixel)	25284	36093	19424	15675	7547
入力画像中の人物領域の欠損 (pixel)	4689	11364	3155	4485	2222
入力画像中の背景の過検出 (pixel)	3624	2936	5868	3276	2814

ブロックサイズ B を 24, 32, 40 の 3 通りに変化させて得られる推定完全画像を、人物領域部分を拡大して図 11, 図 12 にそれぞれ示す。図 11, 図 12 のそれぞれについて、目標完全画像中の人物領域の総画素数、入力画像中の人物領域の欠損画素数、過検出画素数を表 2 に示す。表 4, 表 5 は、それぞれ図 11, 図 12 に示した場合の前景ピクセル数と過検出ピクセル数、欠損ピクセル数、評価値を示す。また図 14, 図 15 で、評価値をそれぞれグラフに示す。

表 2(a), (b) と表 4 より、複数人物が重なったパターンに対しても人物領域の欠損を減少させる効果があることが見て取れる。図 11 から、頭部や脚部などの人物の特徴的な部位を捉えて補完をしていることがわかる。一方で、表 2(a), (b) と表 4 から、人物領域を過検出したピクセル数が増加しているのが見て取れる。これは、本実験で用いた事例が全て人物一人のパターンであるため、図 11(a) の画像右中央部や (b) の画像中央部に見られる人物間の隙間のように、人物二人から成る複雑なシルエットを適切に捉えられず、欠損として補完してしまうことが原因として考えられる。また同様の理由で、図 11(b) の画像中央部の女性の脚部がノイズと捉えられ、欠損していることがわかる。これらの問題を解決し、より詳細な人物領域を得るためには、事例に人物二人のパターンを加えて評価実験をする必要があ

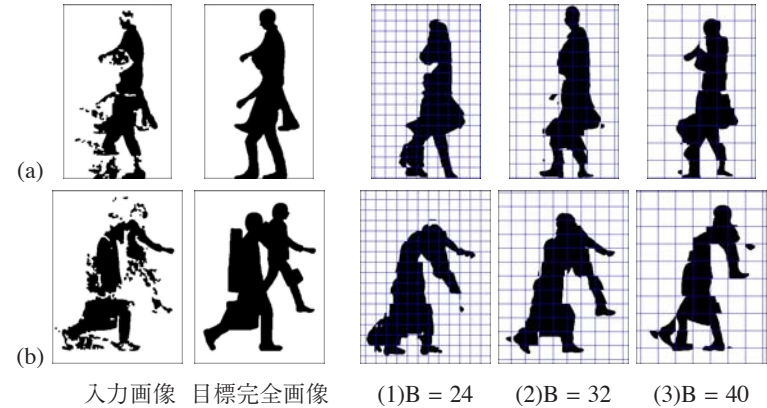


図 11 複数の人物領域に対する推定完全画像の比較
Fig.11 Comparison of estimation results for human regions with occlusion

る。また、ブロックサイズの観点からみると、表 4 より、ブロックサイズが小さいほどより人物領域の欠損ピクセル数が減少するが、過検出ピクセル数が増加する傾向にあることが見て取れる。図 11 から、ブロックサイズが小さいほどより詳細に人物領域を補完できているが、図 11(b) の画像左下に見られるノイズ領域まで補完していることがわかる。本実験で用いた入力画像に対する最適なブロックサイズは、提案する評価基準を用いると図 14 より、(a), (b) の場合共に $B = 32$ となっている。このことから、最適なブロックサイズは入力画像中の人物領域パターンの複雑さに依存しないで決まることが考えられる。

更に、大きさが異なる 3 つの入力パターンに対して得られる結果から、最適なブロックサイズについて考察する。図 12(c) より、人物領域に対してブロックサイズが小さすぎると、首部分のような人物シルエットの特徴を適切に捉えられないことがわかる。また図 12(e) より、人物領域に対してブロックサイズが大きすぎても、人物シルエットを適切に捉えていないのが見て取れる。これらの結果から、最適なブロックサイズと人物領域の大きさには相関があることが考えられる。図 15 より、我々の評価基準を用いても (c) では $B = 40$ の場合に、(d) では $B = 32$ の場合に良い評価をしていることがわかる。しかしながら、(e) に対しては $B = 40$ の場合に最も評価値が高く、図 12 で見られる結果とそぐわない評価をしている。これは、 $B = 40$ の場合に、我々の仮定に反して真の人物領域から遠ざかるように補完がなされたためである。よって、このような場合にも適切なブロックサイズを選択するため

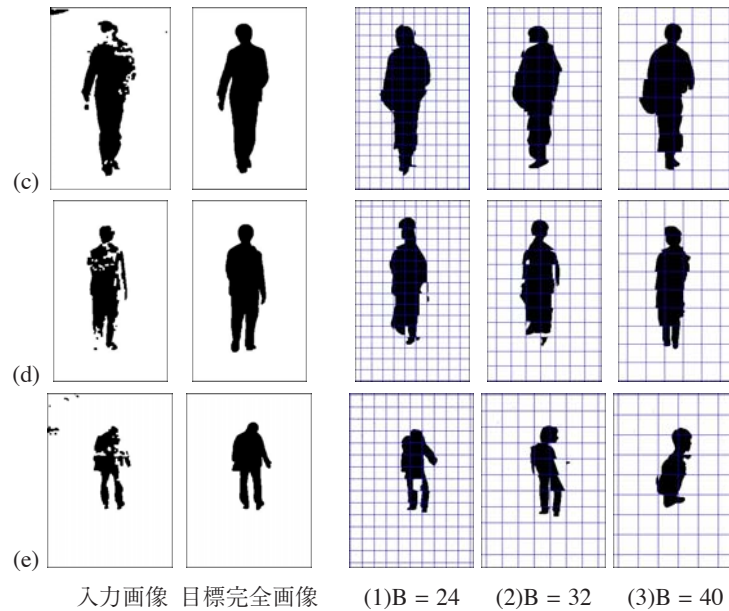


図 12 異なる大きさの人物領域に対する推定完全画像の比較
 Fig. 12 Comparison of estimation results for human regions of different sizes

に、4.1 で述べた仮定を反映しない別の評価基準を導入する必要がある。また、真の人物領域から遠ざかるように補完がなされた理由として、 $B = 40$ とした場合に、事例データの量が不十分であったことが考えられる。これは、ブロックサイズが大きくなる程、1つのパッチが保有する情報量が増える一方で、1枚の画像データから得られるパッチの数が減ることに起因する。

5. おわりに

我々は事例データを用いて、背景差分により得られる欠損を含んだ人物領域画像から欠損を補完した完全な人物領域画像を得る手法を提案した。事例データとして、人物領域の欠損パターンと完全パターンを組にしたものを作成し、これらをブロック単位で用いることで有限の事例データから多様な姿勢の補完を可能にした。そして、入力画像中の欠損に対する最適な事例データを選び出すために、入力画像と事例データ中のブロック間の関係を定義した

表 3 ペナルティ項の有無による欠損/過検出ピクセル数と評価値の比較
 Table. 3 Comparison of the number of under/over-detected pixels and evaluation values with/without penalties

	ペナルティ	人物領域 (pixel)	欠損 (pixel)	過検出 (pixel)	評価値
(1) B = 24	無	19347	3146	1346	0.058
	有	20536	1785	1174	0.065
(2) B = 32	無	19063	3434	1350	0.064
	有	20626	2052	1531	0.074
(3) B = 40	無	20465	2833	2151	0.087
	有	20516	2787	2156	0.087

表 4 複数の人物領域に対する欠損/過検出ピクセル数と評価値の比較
 Table. 4 Comparison of the number of under/over-detected pixels and evaluation values for human regions with occlusion

	B	人物領域 (pixel)	欠損 (pixel)	過検出 (pixel)	評価値
(a)	24	24957	4039	4025	0.102
	32	25274	2891	3194	0.111
	40	24754	3589	3372	0.102
(b)	24	30806	9547	4260	0.107
	32	31190	9598	4695	0.129
	40	25919	13640	3466	0.111

表 5 異なる大きさの人物領域に対する欠損/過検出ピクセル数と評価値の比較
 Table. 5 Comparison of the number of under/over-detected pixels and evaluation values for human regions of different sizes

	B	人物領域 (pixel)	欠損 (pixel)	過検出 (pixel)	評価値
(c)	24	19837	1287	1700	0.089
	32	18856	1869	1301	0.082
	40	19938	1821	2335	0.110
(d)	24	14335	2507	1167	0.104
	32	14245	2790	1365	0.111
	40	13191	3249	765	0.091
(e)	24	6610	1567	630	0.116
	32	6823	1916	1192	0.146
	40	6196	3460	2109	0.221

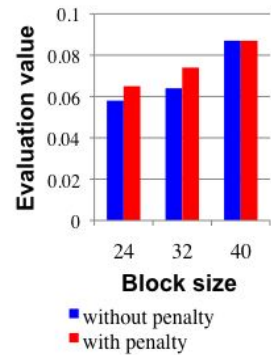


図 13 ペナルティ項の有無による評価値の比較
 Fig. 13 Comparison of evaluation values with/without penalties

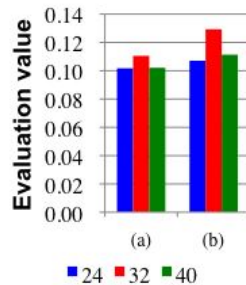


図 14 複数の人物領域に対する評価値の比較
 Fig. 14 Comparison of evaluation values for human regions with occlusion

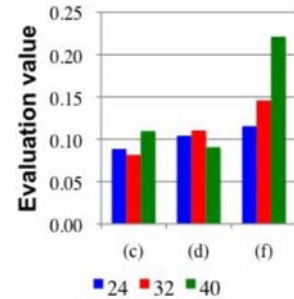


図 15 異なる大きさの人物領域に対する評価値の比較
 Fig. 15 Comparison of evaluation values for human regions of different sizes

目的関数を導入した。実験により、この目的関数を最小化することで、欠損を補完した人物領域が得られることを確認した。

今後の課題として、場面に応じて最適なブロックサイズを選び出すために、新たな評価基準を導入する必要がある。また、現在のシステムはデータベースの探索に多くの時間を要するため、ハッシュ関数などを用いて解空間を限定することで、計算時間を改善することが望まれる。

参 考 文 献

- 1) C. Stauffer and W. E. L. Grimson, "Adaptive background mixture models for real-time tracking", in *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 2, Fort Collins, CO, USA, Jun. 1999, pp. 246–252.
- 2) A. Elgammal, R. Duraiswami, D. Harwood and L. S. Davis, "Background and foreground modeling using nonparametric kernel density estimation for visual surveillance", *Proceedings of the IEEE*, Vol. 90, No. 7, pp. 1151–1163, 2002.
- 3) A. Yamamoto and Y. Iwai, "Real-time object detection with adaptive background model and margined sign correlation", in *Proc. the 9th Asian Conference on Computer Vision*, Xi'an, Sep. 2009.
- 4) B. Leibe, A. Leonardis and B. Schiele, "Combined object categorization and segmentation with an implicit shape model", in *Proc. ECCV workshop on statistical learning in computer*

- vision*, 2004, pp. 17–32.
- 5) A. Mohan, C. Papageorgiou and T. Poggio, "Example-Based Object Detection in Images by Components", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 23, pp. 349–361, 2001.
- 6) R. Collins, R. Gross and J. Shi, "Silhouette-based human identification from body shape and gait", in *Proc. 5th Intl. Conf. on Automatic Face and Gesture Recognition*, 2002.
- 7) 石川博, "グラフカット," コンピュータビジョン最先端ガイド 1, 八木康史・斎藤英雄(編), 第2章, アドコム・メディア, 2008.
- 8) Felzenszwalb, Pedro, Huttenlocher and Daniel, "Efficient belief propagation for early vision," *International Journal of Computer Vision*, Vol. 70, No. 1, pp. 41–54, 2006.
- 9) C. Rother, S. Kumar, V. Kolmogorov and A. Blake, "Digital tapestry," in *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 1, Jun. 2005, pp. 589–596.
- 10) D. W. Kim and K. S. Hong, "Practical background estimation for mosaic blending with patch-based Markov random fields," *Pattern Recognition*, Vol. 41, No. 7, pp. 2145–2155, 2008.
- 11) H. Cao and V. Govindaraju, "Preprocessing of low-quality handwritten documents using markov random fields," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 31, No. 7, pp. 1184–1194, 2009.