

Kinect による安定な動物体領域抽出のための カラー画像とデプスマップの時刻合わせ

中島 秀真^{†1} 満上 育久^{†1} 山添 大丈^{†1}
波部 齊^{†1} 榎原 靖^{†1} 八木 康史^{†1}

従来のカラー画像に対する背景差分ベースの前景領域抽出は、前景色と背景色が類似している場合や背景に変動が起こった場合に失敗するが、Kinect のように対象シーンに対するカラー画像とデプスマップを取得できるセンサを用いると、それぞれの背景差分を統合することにより、より安定に前景領域を取得できると考えられる。しかし、Kinect では、この 2 種類のセンサが非同期で動作しており、前景が動物体の場合に、両者を単純に統合することができない。そこで本論文では、Kinect の 2 種類のセンサの撮影挙動に関する調査について報告するとともに、その結果を踏まえ、背景差分結果のモーフィングによって擬似的に同期のとれた画像を生成し、それらを統合する手法を提案する。実環境で歩行者を対象とした領域抽出を行い、提案手法の有効性を確認した。

Pseudo-synchronization of Image and Depth Map of Kinect for Robust Extraction of Dynamic Object

HOZUMA NAKAJIMA,^{†1} IKUHISA MITSUGAMI,^{†1}
HIROTAKE YAMAZOE,^{†1} HITOSHI HABE,^{†1}
YASUSHI MAKIHARA^{†1} and YASUSHI YAGI^{†1}

Existing background subtraction methods often fail to extract a foreground region whose color is similar to that of the background. When we use Microsoft Xbox Kinect, since we can obtain both a color image and depth map simultaneously, it is expected to get a better foreground region by integrating extracted regions from the two kind of images. We have found, however, these two images are not captured synchronously so that the integration cannot be performed straightforward when a moving object is observed. In this paper, we first report investigation of the capturing behavior of Kinect, then propose a pseudo-synchronization method using a morphing technique. Experimental results about extraction of a walking human show its effectiveness.

1. はじめに

Microsoft 社の Xbox Kinect (以下 Kinect) は、同社製ゲーム機器における直感的インタフェースのためのセンサとして販売されたが、同社¹⁾ および OpenNI²⁾ などのグループによって PC から制御できる SDK が提供されたことで、最近ではロボットビジョン・ヒューマンインタフェースなど幅広い分野で様々な用途で利用されている。

Kinect によって得られるのは、これまでの一般的なカメラで得られていた二次元カラー画像と、そのシーンのメトリックなデプスマップである。多くの利用事例において、このデプスマップが簡単に得られるという点が重視されており、Shotton らはこのデプスマップからシーン中の人物の三次元姿勢推定を行なっている³⁾。この技術は、Microsoft の SDK の姿勢推定機能としても実装されており、容易に利用できる。また、Izadi らは Kinect を環境中で動かしながら撮影することで環境の三次元形状復元を行う手法を提案している⁴⁾。いずれの手法でも、推定や統合の対象はデプスマップであり、同時に撮影されているカラー画像はそれらの処理結果に対してデプスマップへのテクスチャマッピング用として付加的な形で利用されることが多い。

一方で、既存のカメラで撮影される二次元のカラー画像を対象として、人物の姿勢推定^{5),6)} や、Structure from Motion や Multi-view Stereo に代表されるような環境の三次元形状復元を行う手法^{7),8)} などが数多く行われていることを考えると、Kinect におけるカラー画像を単なるテクスチャマッピング用の情報としてだけではなく、より積極的に利用するのが有用であると考えられる。観測値の属性や観測原理が異なれば苦手とするシーンの特徴は異なるため、同一シーンのカラー画像とデプスマップを相補的に利用することができれば、処理全体のパフォーマンス向上が期待できる。例えば、背景差分による人物抽出処理を考えた場合、カラー画像に対する背景差分は、前景色と背景色が類似している場合や背景に変動が起こった場合に弱い。一方デプスマップに対する背景差分は、色の類似性や背景変動には強いが、人の接地部分は背景との奥行き差分値が小さいため抽出が不安定となり、また、その計測原理上投影光の反射を正しく取得できない頭髪部分やオクルーディングエッジ近くの領域は正しく計測を行うことはできないため、そもそもの観測値が正しく得られないケー

^{†1} 大阪大学
Osaka university

スが存在する。

以上の議論を踏まえ、本論文では、Kinect における背景差分ベースの前景領域抽出処理をカラー画像とデプスマップを相補的に用いることで高精度で行うことを目的とする。ただ、これを実現する上で、大きな足かせとなるのが、カラー画像とデプスマップの撮影が非同期であるという事実である。上述のとおり、Kinect を用いた既存研究の多くではカラー画像はデプスマップによって得られる三次元的情報に色を与えるために用いられていたため、その評価は主観的なものが多く、わずかな撮影タイミングのずれが問題視されてこなかった。Kinect を用いた研究が近年急速に増えているにもかかわらず、この問題を明示し取り扱った先行研究が存在しないのは、このような要因によるものだと考えられる。本論文では、まず予備調査によってこの非同期撮影の挙動を明らかにするとともに、その挙動を踏まえた疑似的な同期画像を生成し、それをグラフィックによって統合することで、より正確で安定な人物領域抽出を実現する手法を提案した。そして、実環境で歩行者を対象とした領域抽出を行い、提案手法の有効性を確認した。

2. Kinect の撮影挙動の調査

Kinect (図 1) は、図 2 のように、観測シーンのカラー画像とデプスマップを同時に取得できるデバイスである。カラー画像は内蔵されたカメラで撮影され、デプスマップは赤外カメラと赤外光のドットパターン投影によって取得されており、いずれも仕様上は 30fps で動作する。カラー画像用のカメラとデプスマップ用の赤外カメラは物理的に異なる位置にあり、その内部パラメータも異なるため、取得されるカラー画像とデプスマップは単純には重ね合わせることはできないが、SDK により工場出荷時のキャリブレーション情報を利用して擬似的な位置合わせを行うことができる。この位置合わせは、同デバイスで取得可能な奥行きレンジでは十分な精度であることを実験的に確認している。

本論文では、Kinect で取得されるカラー画像とデプスマップそれぞれにおける背景差分を統合し、より正確な前景領域抽出を行うことを考える。対象とする前景物体が人などの同物体の場合は、カラー画像とデプスマップが同時刻に撮影されなければならない。しかし、公式な技術仕様には、この同期撮影が行われているかどうかに関する情報が存在しない。また、Kinect を用いた既存研究では、この同期撮影の有無の影響を受けない、あるいは受けにくいものが多かったため、厳密な調査を行ったものが存在しなかった。そこで本論文では、まずこの同期の有無を含む Kinect の撮影挙動に関する調査を行った。以下にその結果と得られた知見を述べる。



図 1 Microsoft Xbox Kinect
Fig.1 Microsoft Xbox Kinect.

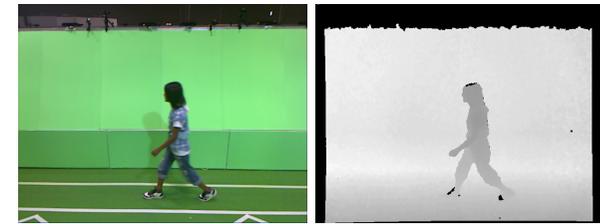


図 2 Kinect でキャプチャしたカラー画像 (左) とデプスマップ (右)
Fig.2 Color image (left) and depth map (right) captured by Kinect.

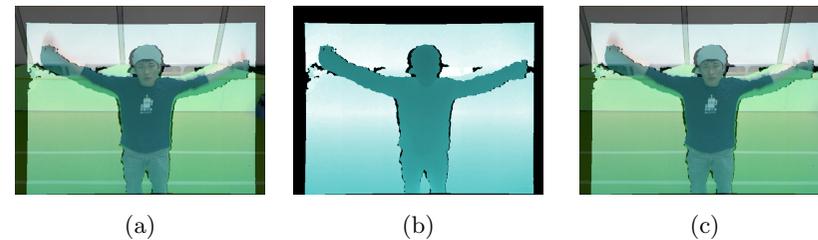


図 3 Kinect の撮影非同期問題
Fig.3 Asynchronization problem of Kinect.

まず、実際に Kinect で撮影を行ったとき、ソフトウェアレベルでの同期^{*1}を行ったカラー画像とデプスマップの間で動物体領域における両者のズレが観測された(図 3)。図 3(a)(b) は Kinect から同時に取得されたフレームのカラー画像とデプスマップであり、図 3(c) はその重ね合わせの結果である。図 3(c) で頭部・腰・足などの静止領域においては同様のズレが観測されていないことを踏まえると、ズレの原因はカラー画像とデプスマップの位置合わせのズレではなく、これらの撮影タイミングについてハードウェアレベルでの同期が行われていないためであると考えられる。さらに、撮影を継続すると、時間の経過に伴い図 4 のように動物体領域のズレの程度や前後関係が変化することが確認された。このことから、撮影時刻のズレが一定ではないことが分かる。以上を踏まえると、Kinect の撮影挙動は、以下のいずれかであると考えられる。

*1 ここでの「ソフトウェアレベルの同期」は OpenNI の関数によって行う。撮影時に WaitAndUpdateAll を用いてカラー画像とデプスマップの両方のノードで新しいデータが取得可能になるのを待ってから二つのノードを更新し、さらに FrameSyncWith を用いてフレームの同期機能を有効にする。

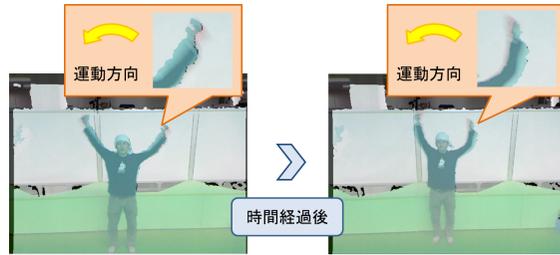


図 4 時刻ズレの前後関係の逆転
Fig. 4 Temporal change of capture time difference.

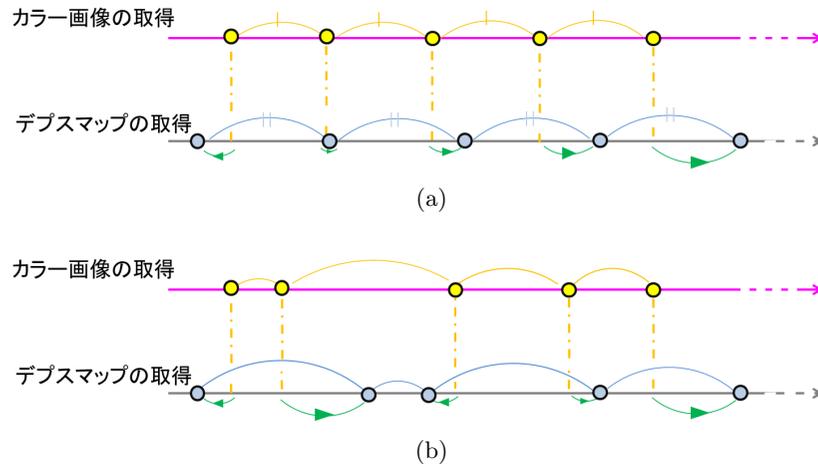


図 5 二つの仮説
Fig. 5 Two possible hypotheses.

仮説 A カラー画像・デプスマップの取得はそれぞれのセンサについては等間隔で行われているが、その周期がわずかに異なるため撮影時刻のズレの量・前後関係が連続的に変化する(図 5(a))

仮説 B カラー画像・デプスマップの取得はいずれも等間隔で行われておらず、撮影時刻のズレの量・前後関係がランダムに変化する(図 5(b))

2.1 予備実験

前節の 2 つの仮説のうちどちらが正しいかを判断するため、以下の予備実験を行った。図



図 6 予備実験
Fig. 6 Preliminary experiment.

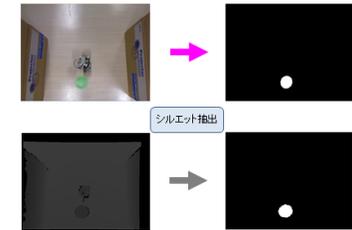


図 7 円板のシルエットの抽出
Fig. 7 Extraction of disk silhouette.

6のように、直径 8.5cm の薄い円板をモーターで等速円運動させ、その 0.7 m 上方に Kinect を設置して一定時間撮影を行った。これを異なる時間に合計 3 試行を行った。なお、撮影の際、SDK の設定でフレームレートは 30fps を指定した。そして、図 2.1 のようにカラー画像とデプスマップからそれぞれ円板のシルエットを抽出し、カラー画像とデプスマップの各系列内での一周分の軌跡を観察する。また、1 試行内におけるシルエットの重心の移動量の平均と偏差を求める。

2.2 実験結果

シルエットの一周分の軌跡(図 2.2)と重心移動量の平均と偏差(表 1)を示す。モーションブレイクの影響によりわずかな誤差はあるものの、シルエットの移動量はほぼ等しいことが分かる。また、重心移動量の偏差は移動量の平均と比べ十分小さいといえる。

また、可視化した両系列のシルエット間のズレ量・ズレの前後関係を図 8 に、その時間による変化を図 2.2 に示す。10 回転、25 回転後のほぼ同位置どうしのシルエットのズレの量・前後関係を比較すると、全ての位置で同じ方向に連続的な変化をすることが分かった。

2.3 考察

予備実験の結果より、両画像系列ともに等速運動を行う物体の隣接フレーム間での移動量に差が見られない。すなわち、2 デバイス内でのデータ取得は等間隔で行っているとみなせるので、仮説 1 が正しいと考えられる。すなわち、カラー画像・デプスマップの取得はそれぞれのセンサについては等間隔で行われているが、その周期がわずかに異なるため、Kinect から同時に取得した画像でも撮影時刻にズレがあり、そのズレ量・前後関係が連続的に変化することが分かった。図 2.2 の結果は時刻ズレが連続的に変化していくことを示しており、

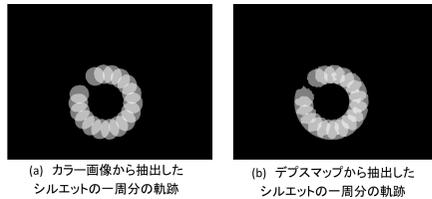


図 8 各系列のシルエットの一周分の軌跡
Fig. 8 Trajectories of silhouette.

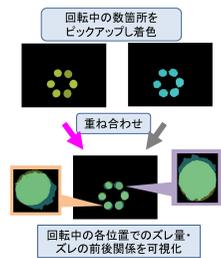


図 9 ズレ量・ズレの前後関係の可視化
Fig. 9 Visualization of displacement.

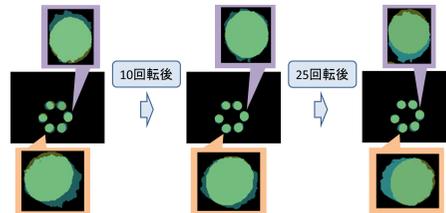


図 10 ズレ量・ズレの前後関係の時間変化
Fig. 10 Time variation of displacement.

これも同結論を裏付けている。

3. 提案手法

本論文では、前節での調査結果を踏まえ、カラー画像とデプスマップが非同期で撮影されるという問題を解消するために、一方の前景抽出結果の時系列に対してモーフィング処理を施して擬似的にフレームレートを向上させ、他方の時系列と時刻合わせを行う手法を提案す

表 1 シルエットの重心移動量の平均と偏差
Table 1 Average and deviation of silhouette movement.

	フレーム数	移動量平均 (Pixel)		移動量偏差 (Pixel)	
		カラー画像	デプスマップ	カラー画像	デプスマップ
試行 1	938	29.0505	29.2663	1.30958	1.71071
試行 2	872	28.1577	28.3738	1.22221	1.58679
試行 3	924	28.9114	29.1237	1.33953	1.75070

る。さらに、その結果得られる擬似的に同期のとれたカラー画像とデプスマップから得られる前景抽出結果を用いて、グラフカットによって正確な前景領域抽出を行う。以下にその詳細について順に述べる。

3.1 前処理

まず、カラー画像・デプスマップそれぞれからの前景抽出については、既存の単一ガウス分布による背景モデルを用いた背景差分手法とノイズの除去によって行う。なお、デプスマップに対しては事前に撮影した背景シーンから生成される固定の背景モデルを用い、カラー画像に対しては照明変化等による背景画素値変化に対応するために背景モデルを動的に更新する。ノイズの除去は、背景差分結果の 2 値画像からある一定の大きさ以下の領域を削除することで行う。カラー画像についてはさらに、観測画素値と背景モデルの色ベクトルの角度差を評価することで影領域を検出し、その領域を除去する処理を加えている。

3.2 Earth Mover's Morphing による擬似的な時間分解能向上

前節で述べた前処理によって、カラー画像・デプスマップそれぞれの時系列に対してシルエット画像が得られるが、前述のとおりこれらは同期がとれておらず、あるカラー画像に対して厳密に同時刻に撮影されたデプス画像は一般には存在しない (図 11)。そこで、一方の系列に対して、横原らの Earth Mover's Morphing⁹⁾ を用いてモーフィングを行い、系列中の隣接した 2 枚に対し遷移率 $1/M, 2/M, \dots, (M-1)/M$ の画像、つまり擬似的な M 分割した時刻の画像を作成し、一方の画像系列を密にする (図 12)。これにより最大の時刻ズレ幅が $1/M$ になるので、M を大きくすれば、両系列の位置合わせの時間分解能を細かくできることを意味する。

3.3 系列間の対応付け

前章で明らかにした「カラー画像・デプスマップの両系列の周期は系列内では不変だが、系列間では異なる」という性質により、系列全体に対して並行移動量 (以下「トランスレーション」) T^{*1} と時間伸縮量 (以下「スケーリング」) S を変化させて他方の系列の全体と時間的に一致するところを探索する (図 13)。両系列が時間的に一致しているかを評価するには、理論上異なる 2 時刻での対応関係を取得すればよいが、背景差分によって抽出されたシルエットにはノイズが乗ることや一方の画像系列はモーフィングによって生成された画像であることを考えると、時刻によっては対応関係の推定を大きく誤る可能性がある。そこ

*1 T の単位は「フレーム」である。ただし、実際には時間分解能を向上させた系列での時間合わせを行うため、値は 1 未満の非常に小さな実数値をとる。

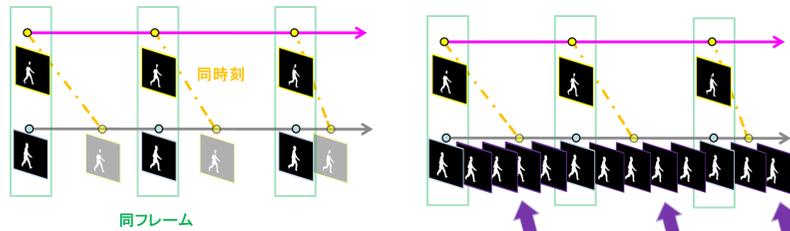


図 11 カラー画像とデプスマップのシルエット画 図 12 モーフィングによる高フレームレート画像
像系列 像系列

Fig. 11 Silhouette image series of color images and depth maps. Fig. 12 Generating high frame rate images.

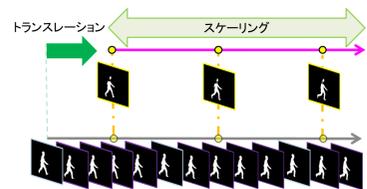


図 13 トランスレーションとスケールリングによる系列全体の時刻合わせ
Fig. 13 Time adjustment of two series.

で、両シルエット画像系列全体の一致度を評価して、その評価値が最大となる S, T を求め、それを最適なトランスレーションとスケールリングとする方法をとる。

3.3.1 シルエット画像系列間の時間一致度の評価

両シルエット画像系列の時間一致度は、各時刻の画像間の時間一致度を累積して求めることとする。各時刻の時間一致度は、図 14(a)(b) のようなシルエット対に対して定義するが、尺度として考えられる最も単純なものは、シルエットの共通部分の面積の大きさである。時刻の差が大きいほどシルエット全体のズレが大きくなり、共通部分の面積は減少するため、この尺度でも一定の評価は可能である。しかし、図 14(b) のようにシルエットの一部が大きく欠落してしまうなどの外乱が発生したとき、その外乱が評価値に強い影響を与えてしまうと考えられる。図 14 の (c) は (a) と (b) を重ね合わせたもので、緑色の領域が両者の共通部分、青と黄色の領域が非共通部分になっている。シルエット全体のズレによる影響と比べ外乱による影響が大きいことが分かる。本手法では、極力このような外乱の影響が小さくなるような評価尺度として、シルエットの輪郭線（以下「エッジ」）と中心線（以下「ス

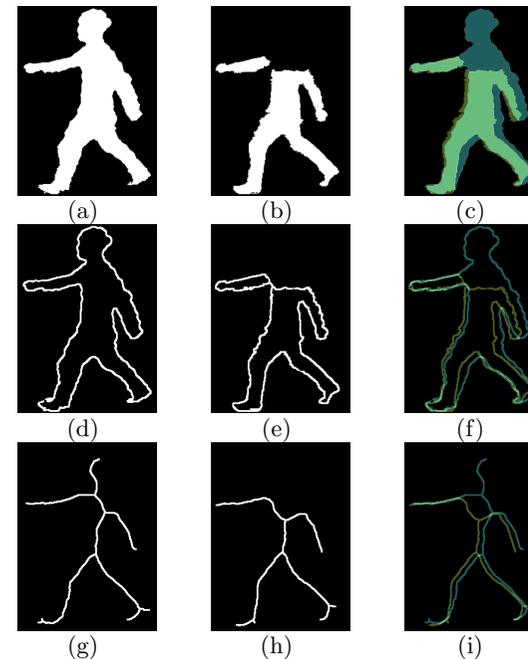


図 14 シルエットの欠損時における時間一致度の評価尺度の違い
Fig. 14 Difference of matching measurements for partially cracked silhouette.

ケルトン」) を用いる。エッジまたはスケルトンの共通部分を尺度とすると、図 14(d)(e)(f) と (g)(h)(i) で示すように、シルエット全体のズレによる影響が非常に大きくなり、欠落による影響は相対的に小さくなる。なお、実際には、エッジやスケルトンそのままでは共通部分の画素数が少なすぎて不安定なため、それぞれ一回膨張処理を行ったものの共通部分の大きさにより評価している。

3.3.2 分布のピーク位置の算出

系列間の時間一致度が最大となる S, T は、探索範囲内から全探索的に求める。図 15 に S, T の変化に対する系列間時間一致度の分布を示す。分布は巨視的に見て単峰性を示すことが分かるが、微視的に見ると図中の拡大図のように細かい増減が確認される。これはエッジやスケルトン画像の離散化誤差によるものと考えられ、この誤差に影響されず巨視的な分布のピークを選択するのが望ましい。そこで本論文では、この分布に以下の二次元ガウス分

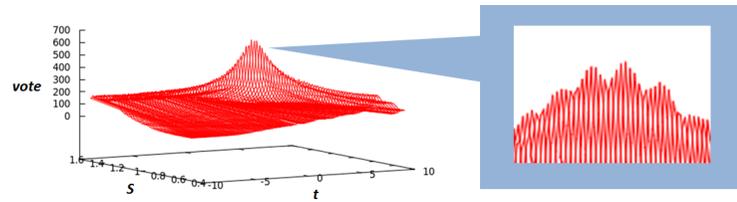


図 15 各 S, T における系列間時間一致度
Fig. 15 Matching score of two series for each (S, T) .

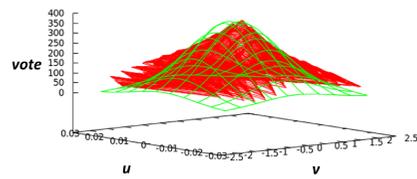


図 16 初期フィッティング
Fig. 16 Initial fitting result.

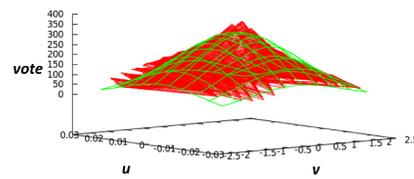


図 17 最適化後のフィッティング
Fig. 17 Fitting result after optimization.

布をフィッティングすることでこのピークを算出する．

$$F(S, T) = \alpha \frac{1}{2\pi\sigma_S\sigma_T\sqrt{1-\rho_{ST}}} e^{-\frac{1}{2(1-\rho_{ST}^2)} \left\{ \frac{(S-\mu_S)^2}{\sigma_S^2} + \frac{(T-\mu_T)^2}{\sigma_T^2} - \frac{2\rho_{ST}(S-\mu_S)(T-\mu_T)}{\sigma_S\sigma_T} \right\}} \quad (1)$$

このフィッティングは、非線形最適化によって行う．最適化のための初期値は以下の方法によって定める．まず、分布のピークから離れた裾野部分はピークの推定には無関係なため、あるオフセット値を定め、それ以下の評価値の領域は削除する．次に、各 (S, T) における評価値を、その (S, T) の点の個数とすることで、この 3 次元曲面を (S, T) の二次元平面上の点群とみなし、その点群に対する主成分分析を行う．これによって、この分布の中心と軸が算出され、式 (1) の各パラメータの初期値が決定される．この初期値を用いてレーベンバーグ・マーカート法によって最適化を行い、分布の中心を決定する．

最適化前後のフィッティングをそれぞれ図 16, 図 17 に示す．この処理によって、分布の大規模なピークが正しく得られていることが分かる．

3.4 グラフカットによるシルエット統合

前節までで対応付けられた両シルエット系列を利用して、最終的な前景領域を取得する．画

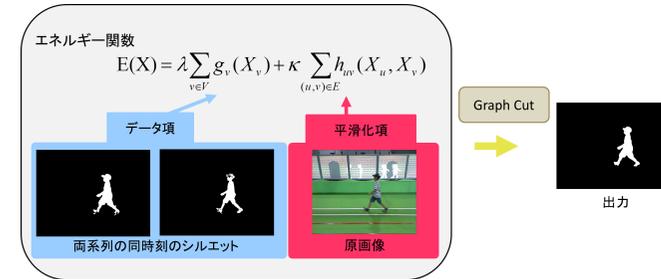


図 18 グラフカットによる 2 つのシルエットの統合
Fig. 18 Integration of two silhouette based on graph cut.

像 V の各画素を v としたとき、ラベルを $\mathbf{X} = \{X_1, X_2, \dots, X_v, \dots, X_{|V|}\}$ として、各 X_v には 0 (背景) か 1 (前景) の値を与える．以下の式で定義されるエネルギー関数 $E(\mathbf{X})$ を最小化するような \mathbf{X} をグラフカットアルゴリズム¹⁰⁾ に従って計算することで、前景領域を決定する (図 18) ．

$$E(\mathbf{X}) = \lambda \sum_{v \in V} g_v(X_v) + \kappa \sum_{(u,v) \in E} h_{uv}(X_u, X_v) \quad (2)$$

$$g_v(1) = \exp(-(\alpha_c c_v + \alpha_d d_v)) \quad (3)$$

$$g_v(0) = 1 - g_v(1) \quad (4)$$

$$h_{u,v} = ||I_u - I_v|| \quad (5)$$

式 (2) の第一項はデータ項、第二項は平滑化項である．データ項の c_v, d_v はそれぞれカラー画像、デプスマップから抽出したシルエットの位置 v の画素の値、平滑化項の I_u, I_v は原画像における位置 u, v の画素の輝度値である．また、 α_c, α_d はバランス調整のためのハイパーパラメータである．

4. 実験結果

実環境下で歩行者を含むシーンを、SDK でフレームレート 30fps を指定し、撮影を行った．そして、得られたカラー画像・デプスマップの時系列に対して本手法による時刻合わせと前景領域抽出を実行した．なお、提案手法におけるモーフィングの分割数 M は 50, 探索する S, T の範囲は S が $[0.5, 1.5]$, T が $[-10, 10]$, 刻み幅は S が 0.01, T が 0.02 とした．

まずは、時刻合わせによって対応付けられた画像を重ね合わせた結果を図 4 に示す．図中の白い領域が共通部分の領域、黄色の領域がカラー画像のシルエットのみの領域、青い領

域がデプスマップのシルエットのみの領域である。(a)は Kinect から同時に取得されたフレームの画像の重ね合わせ、(b)(c)(d)(e)は提案手法で時刻合わせを行い対応付けた擬似的な同時刻の画像との重ね合わせである。(b)(c)はデプスマップの系列でモーフィングを行い、各カラー画像と同時刻のものを対応付けており、(d)(e)は逆にモーフィングしたカラー画像の系列を各デプスマップに対応付けている。また、(b)(d)ではエッジを、(c)(e)ではスケルトンをそれぞれ時間一致度の評価尺度とした。

それぞれ時刻合わせによって、Kinect から同時に取得されたフレームのものよりも同時刻に近いものと対応付けられているといえる。(d)(e)の最下段では人物シルエットの首部分にノイズのようなものが観測されているが、これはカラー画像の連続フレーム間で図20のようにシルエットの欠落の仕方が異なるため、このフレーム間でモーフィングした場合、大きくノイズが乗ったようになってしまうためである。本手法では系列全体で対応付けを行うため、ある程度このような欠落の差があってもさほど時間合わせの結果には影響しない。

さらに、カラー画像から前景抽出を行った結果を、図21に示す。(b)の画像はカラー画像のみを用いて前景抽出を行った場合の結果である。前景色と背景色との類似により上半身が大きく欠けてしまっている。また、(c)は同時刻のデプスマップのシルエットを統合して抽出を行った結果である。より上半身の大きな欠落は解消されているが、本来同時刻には撮影されておらずズレのあるシルエットを統合したために、ノイズの領域が大きい。それと比較して提案手法(d)(e)では、モーションブラーなどの影響で多少ノイズは残るものの、(c)で見られたような大きなノイズの削減には成功している。これより前景抽出の精度を向上させることができたといえる。

5. おわりに

本論文では、非同期撮影されたカラー画像とデプスマップの背景差分ベースの前景領域抽出結果の系列から、一方の系列の擬似的な中間時刻の画像列を生成し、両系列の画像を適切に対応付けて統合することで、Kinect による前景領域抽出処理を高精度に行う方法を提案した。また、本手法を提案するにあたり、Kinect のカメラとデプスセンサの非同期撮影の挙動についての予備調査を行い、「両デバイス内でのデータの取得は等間隔で行われているが、その周期がデバイス間で異なる」という性質を明らかにした。実環境で歩行者を対象に抽出を行い、カラー画像のみを用いたときの抽出結果や、Kinect から同時に取得されたフレームのデプスマップと対応付け統合した場合の抽出結果と比較することで、その有効性を確認した。

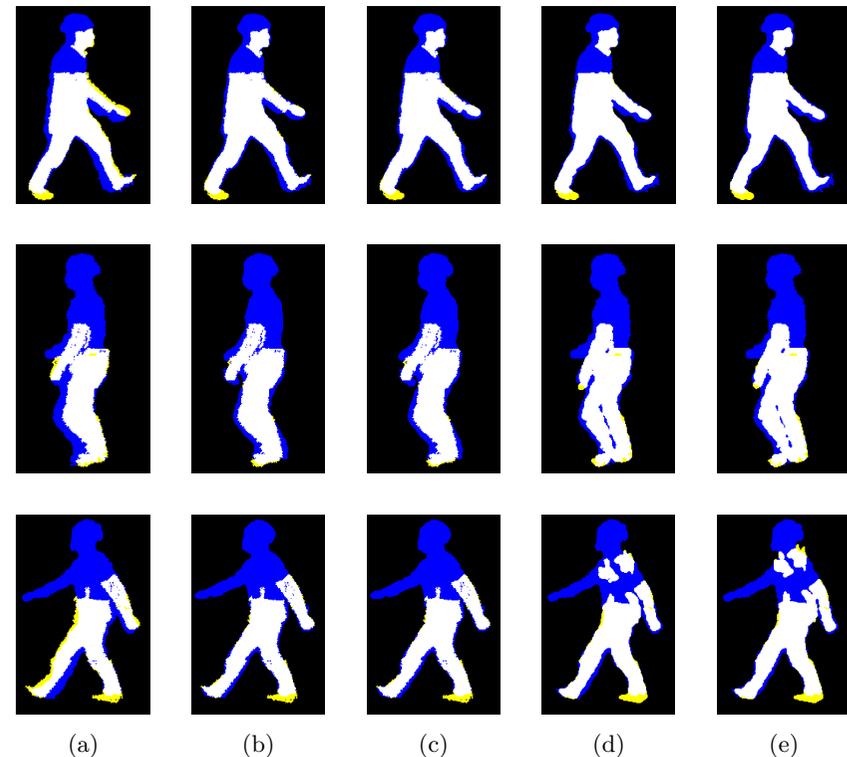


図 19 両系列のシルエットの重ね合わせ
Fig.19 Results of overlay.

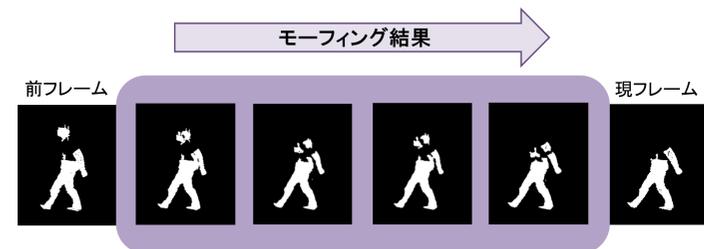


図 20 欠落の仕方が違う隣接フレームのモーフィング結果
Fig.20 Two silhouettes aligned by the proposed method.

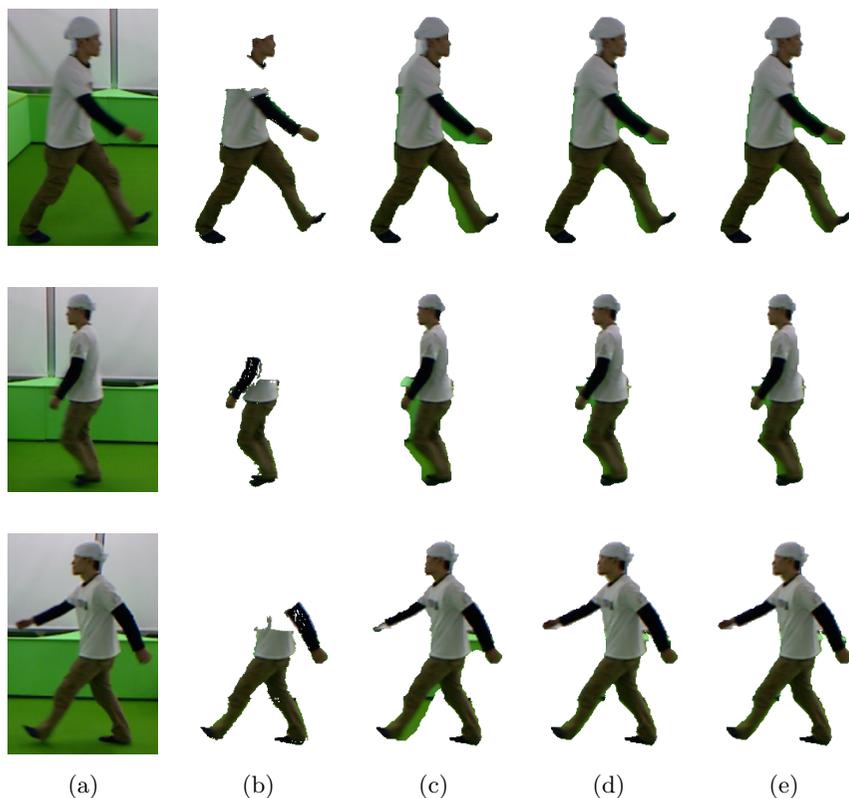


図 21 カラー画像からの抽出結果
Fig. 21 Results of foreground extraction from color images.

今後の課題としては、最終的な前景領域抽出処理における時系列情報の利用が挙げられる。現在は、系列間の時間合わせが済んだ後は、同時刻のカラー画像・デプスマップの統合は各時刻ごとに独立に行なっているが、時間方向の連続性を利用することで、さらに抽出精度を向上させられる可能性がある。また、提案手法の中には、分布の閾値処理やグラフカットのためのパラメータがいくつか存在しており、その値によって結果が変化する。これらのパラメータの自動決定法、あるいは、これらの値の変化に頑健性の向上なども今後の課題として考えられる。

謝辞

本研究は、科学技術振興機構 (JST) 戦略的創造研究推進事業 (CREST) の支援のもとに推進された。

参 考 文 献

- 1) <http://kinectforwindows.org/>
- 2) <http://www.openni.org/>
- 3) J. Shotton, A. Fitzgibbon, M. Cook, A. Blake, "Real-time Human Pose Recognition in Parts from Single Depth Images," Proc. IEEE Conference on Computer Vision and Pattern Recognition, 2011.
- 4) S. Izadi, R. Newcombe, D. Kim, O. Hilliges, D. Molyneaux, P. Kohli, J. Shotton, S. Hodges, D. Freeman, A. Davison, A. Fitzgibbon, "KinectFusion: Real-Time Dynamic 3D Surface Reconstruction and Interaction, Proc. SIGGRAPH, 2011.
- 5) Y. Yang, D. Ramanan, "Articulated Pose estimation with Flexible Mixtures-of-parts," Proc. IEEE Conference on Computer Vision and Pattern Recognition, 2011.
- 6) Y. Wang, D. Tran, Z. Liao, "Learning Hierarchical Poselets for Human Parsing," Proc. IEEE Conference on Computer Vision and Pattern Recognition, 2011.
- 7) N. Snavely, S. M. Seitz, R. Szeliski, "Modeling the World from Internet Photo Collections," International Journal of Computer Vision, Vol.80, No.2, pp.189-210, 2007.
- 8) Y. Furukawa, J. Ponce, "Accurate, Dense, and Robust Multi-View Stereopsis," Proc. IEEE Conference on Computer Vision and Pattern Recognition, 2007.
- 9) Y. Makihara and Y. Yagi, "Earth Mover's Morphing: Topology-Free Shape Morphing Using Cluster-Based EMD Flows," Proc. the 10th Asian. Conf. on Computer Vision, pp. 2302-2315, Queenstown, New Zealand, Nov. 2010.
- 10) 石川博: "グラフカット(チュートリアル)", 情報処理学会研究報告. CVIM, pp. 193-204, 2007.