

固有空間法を用いた照明成分と構造成分の分離に基づく背景画像生成

川西 康友[†] 満上 育久^{††} 棕木 雅之^{†††} 美濃 導彦^{†††}

[†] 京都大学 大学院 情報学研究科 知能情報学専攻 〒 606-8501 京都市左京区吉田本町

^{††} 大阪大学 産業科学研究所 〒 567-0047 大阪府茨木市美穂ヶ丘 8-1

^{†††} 京都大学 学術情報メディアセンター 〒 606-8501 京都市左京区吉田本町

E-mail: †kawanishi@mm.media.kyoto-u.ac.jp, ††mitsugami@am.sanken.osaka-u.ac.jp,

†††{mukunoki,minoh}@media.kyoto-u.ac.jp

あらまし 本論文では、観測画像中の看板など背景中の動かされうる物体や日照の様子といったリアルタイムな情報をできるだけ保ちつつ、人物を除去した背景画像を生成する手法を提案する。従来手法において、画像中から人物を確実に除去するためには長期間の観測で得られる大量の画像に対する平均化処理が必要があるが、大量の画像を用いているために、太陽の移動に伴う日照の変化や、看板の移動のようなシーンの構造の変化を同時に再現できなかった。そこで我々は、画像中の日照成分と構造成分を固有空間を用いて分離することにより、それぞれを個別に再現した後、固有空間を用いてそれらの結果を統合することによりシーンの日照と構造を保存した背景画像生成を実現する。複数のシーンにおいて従来手法と提案手法により背景画像を生成し、観測画像との誤差の大きさによって本手法の有効性を確認した。

キーワード 背景画像生成, 固有空間法, クラスタリング, 長時間画像解析

Background Image Generation Based on Separation of Lighting and Structural Components by Eigenspace Method

Yasutomo KAWANISHI[†], Ikuhisa MITSUGAMI^{††}, Masayuki MUKUNOKI^{†††}, and Michihiko MINOH^{†††}

[†] Graduate School of Informatics, Kyoto University Yoshidahonmachi, Sakyo-ku, Kyoto, 606-8501 Japan

^{††} The Institute of Scientific and Industrial Research, Osaka University 8-1 Mihogaoka, Ibaraki, Osaka, 567-0047, Japan

^{†††} Academic Center for Computing and Media Studies, Kyoto University Yoshidahonmachi, Sakyo-ku, Kyoto, 606-8501 Japan

E-mail: †kawanishi@mm.media.kyoto-u.ac.jp, ††mitsugami@am.sanken.osaka-u.ac.jp,

†††{mukunoki,minoh}@media.kyoto-u.ac.jp

Abstract In this paper, we propose to generate a background image which contains no people and preserves real-time information including lighting condition and structure of outdoor scene observed by a stationary surveillance camera. Although there have been a lot of traditional methods that can generate a background image of the observed image, they can not preserve both of its lighting and structure. The proposed method separates an image into the lighting and structural components by the Eigenspace method, generates a background image preserving each of the components independently, and integrates them by the Eigenspace method. We experimentally evaluate the results of this approach using images captured by several surveillance cameras.

Key words Background Estimation, Eigenspace Method, Clustering, Long-term Image Sequence Processing

1. はじめに

近年、交通監視、セキュリティ、天候観測など様々な目的で身の回りに固定カメラが多数設置されるようになってきた。通常、こうした固定カメラの映像は設置者のみ

がある目的にのみ用いているが、特定の目的以外にも役立つ情報を固定カメラの映像は多く含んでいると考えられる。もし、こうした固定カメラのリアルタイムな映像をインターネットを介して共有することが出来れば、一層固定カメラを活用することが可能になると考えられる。



図1 変身カメラ

例えば、交通監視カメラの映像から天候情報を得られたり、商業施設の防犯カメラから店の様子を得られたりすることなどが考えられる。こうした使い方ができるため、インターネットを介したリアルタイムな固定カメラ映像の共有は非常に有用である。こうした考えに基づくセンシング Web プロジェクト [1] [2] では、固定カメラを含むあらゆるセンサをインターネットに接続しそのデータを公開することにより、誰もがこれらのセンサデータを利用できるような枠組みを提案している。

リアルタイムなセンサ情報を公開する際、最も重要な問題となるのがプライバシーの問題である。特にカメラ映像は、そこに映った人物の多くの情報を含むため、それをそのまま公開するとプライバシー侵害を引き起こす。実際、映像を見れば、顔や服装、動作などのアピランス情報からその人物を特定することが可能である。そのため、プライバシー侵害の問題がセンサデータ公開の際の大きな障害となっている。

こうしたプライバシー侵害の問題があるため、カメラの映像を公開する場合、人物のアピランス情報をその映像中から除去し、プライバシー侵害の恐れがない映像にして公開する必要がある。このとき、元のカメラの映像からできるだけ情報を失わないで映像を公開するためには、背景の様子や人物が映像中のどこに映っているかなど人物のアピランス情報以外はできるだけ保つような処理をすべきである。こうした目的で、従来、映像中から検出した人物に対しマスクをかけるといった方法が採用されている。マスクを掛けることにより、人物の存在はわかるがプライバシーは保護された映像が得られる。このアプローチは人物検出の精度に大きく依存するが、人物検出の精度は不完全であるため、人物を正しく検出できない場合があり、そのときには人物のアピランス情報を完全に除去することができない。

そこで我々は、プライバシーを確実に保護した映像を出力する「変身カメラ」を提案している(図1)。「変身カメラ」は、人物をキャラクターに置き換えることによって、人物の位置情報は保ったままアピランス情報を除去した映像を出力する。このとき人物を全く含まない背景画像を生成し、人物を検出できた時のみ背景画像上にシンボルを描画する。この方法ならば、もし人物検出に失敗したとしても、背景画像に描画されるシンボルの位

置がずれたり、そもそも描画されなかったりするだけで、プライバシーが漏洩することはない。これを実現するために、本研究では、人物が確実に背景画像中に現れず、また背景の再現性が高い背景画像を生成することを目標とする。

2. 従来の人物を含まない背景画像生成手法

2.1 人物検出に基づく手法

もし映像中の人物が確実に検出できるのであれば、人物以外の領域はそのまま観測画像を背景画像として用いれば良いのであるが、もし人物が存在した場合、人物がいた場所は空白になってしまうため、完全な背景画像を得ることは出来ない。このとき、画像の欠損を修復する Image Inpainting [3] などの手法を用いることにより、人物がいた場所の空白部分を復元することは可能であると考えられる。こうした手法では、人物以外の領域はそのままの観測画像であるため、背景の再現性は非常に高い。しかしこれまでに提案された人物検出手法では、任意の環境下で確実に人物を検出することができる手法は存在しないため、検出に失敗した場合、背景画像に人物が現れてしまう可能性がある。

2.2 時間方向の平均化処理による背景画像生成

従来、固定カメラにおける背景画像生成手法としては、観測画像の直前数枚の観測画像の画素ごとの時間平均やメディアン、モードなどを求めることにより、観測画像に類似した背景画像を生成する手法 [4] [5] [6] が特に物体検出等に用いられている。こうした手法は、抽出対象となる物体が常に移動していると仮定し、一定期間変化しないものは背景であるという考えに基づいている。しかし人物が映り得る屋外の固定カメラの場合、待ち合わせや立ち話等でしばらく立ち止まっている人物が背景と見なされてしまうといったことが起こる。この場合、背景画像中に人物が現れてしまう。

一定期間停止している人物が背景と見なされてしまう問題に対処するためには、人物が立ち止まり得る期間よりも十分長い期間観測された大量の画像に対して、画素毎の平均やメディアン等の平均化処理をすることが考えられる。こうすることにより人物が背景画像中に現れないという条件を満たすことができる。しかし、こうした長期間の観測画像を単純に平均化処理するアプローチでは、その長期間の間に起こるシーン中の変動が平均化されてしまい、背景画像の再現性が著しく低下してしまう。本論文で想定するような屋外シーンにおいては、まず天候や昼夜の違いによりシーン全体の明るさが変化する。さらに天気の良い日中は観測画像中で鋭いエッジを持つ影が存在し、またそのような影は時間とともに徐々に移動する。また、店頭の看板など、毎日置き換えられたり移動されたりする物体が存在し、その移動によってシーンの構造が変動する。こうした変動を含む背景画像に対

して平均化を行うと、天候ごとの日照の様子が失われたり、存在しない物体がゴーストとして生成した背景画像中に現れたり、存在するはずの物体が生成した背景画像中に現れなかったりする。したがって、このように背景が大きく変動する場合、実際の観測画像と大きく異なった背景画像が生成されてしまう。

2.3 固有空間法に基づく手法

固有空間法は、大量のベクトルの集合に対し、その中で限られた変動パターンを主成分分析によってその変化の大きいものから順に主成分として見つけ出し、その線形和によって各ベクトルを表現する手法である。これを固定カメラを用いて長期間観測した画像セットに適用することを考える。固有空間法を用いることで、背景の変動の主成分を求め、人物を含みうる現在の観測画像から、人物を含まない背景画像を主成分の線形和によって生成することができる[7]。これは、次のようにして実現される。

まず、ある固定カメラで観測した大量の画像に対し、各画像をそれぞれひとつのベクトルとして主成分分析を行うことで次元圧縮が行われ、 k 次元の固有空間（寄与率の大きい順にソートされている固有ベクトル（主成分） e_1, e_2, \dots, e_k によって張られる空間）を得る。こうして得られた固有空間の基底は、観測画像セットの変動傾向を反映しているので、背景画像 $x^b(t)$ は観測画像 $x(t)$ から式 (1) によって得られる。

$$x^b(t) \approx E p(t) + x^{\text{ave}} = E E^T (x(t) - x^{\text{ave}}) + x^{\text{ave}} \quad (1)$$

ただし、 $p(t)$ は e_1, e_2, \dots, e_k によって張られる固有空間における観測画像 $x(t)$ に対応する点、 $E = [e_1 \ e_2 \ \dots \ e_k]$ は、それらを結びつける正規直交行列であり、 x^{ave} は平均画像である。

この手法では、背景の変動を固有空間でモデル化することを意図しているため、人物などを含まない背景のみの画像セットを主成分分析に使用するべきである。しかし、人物が含まれているかどうかを確実に判定することはできないため、人物を含まない画像セットを抽出しそのみを主成分分析に利用することはできない。ここで、人物は画像中において比較的小さく、任意に移動し、また瞬間的にしか現れないため、人物の存在は背景の変化による変動と比較すると非常に小さい。そのため、主成分分析による次元圧縮により k を十分小さくとれば人物の影響を受けない背景変動の主成分を抽出できる。

しかし、このように背景画像をいくつかの主成分の線形和で表現すると、昼や夜といった大きく異なる変化は表現できるものの、移動する影や、置き場所が変化しうる看板などの多くのパリエーションを持つ変化は次元圧縮の結果失われ、うまく再現することは出来ない。例えば、観測画像中で鋭いエッジを持つ影が存在し、またそのような影は時間とともに徐々に移動する場合、その

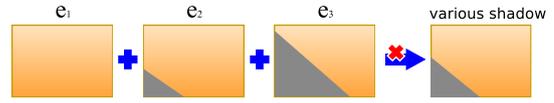


図 2 少数の固有空間の線形和では様々な位置に出る影を表現できない。

シーンにおける様々な影の出方を含む画像セットから固有空間法を用いて背景画像を生成するには、様々な位置に出る影に応じた様々な固有ベクトルが必要となり、固有空間は非常に高次元なものとなる。寄与率の高い上位 j 個の固有ベクトルからなる固有空間を用いて背景画像を生成する場合、このような様々な位置に出る影に応じた固有ベクトルは無視されてしまい、様々な影の出方に応じた背景画像を固有ベクトルの線形和で生成することができない(図2)。したがって、実際の観測画像と大きく異なった背景画像が生成されてしまい、違和感のある画像になる。これに対し、 j 個以上の固有ベクトルからなる非常に高次元の固有空間を用いて背景画像を生成することで再現性を上げることができるが、固有空間の次元が増えるに従い、画像中の人物による影響が出てしまい、生成した背景画像中に人物など不要なものが現れてしまう。そのため、固有空間の次元数を増やすことによって再現性を向上するアプローチは採れない。

2.4 日照を保存した背景画像生成法

前節の最後に述べたような影の移動に伴う背景変化に対し、Kawanishi ら [8] は屋外シーンにおいて、長時間観測・蓄積した画像セットに対し、影の出方に応じた分類を行い、同じ位置に影が出る画像ごとに固有空間法を適用することで、日照の様子を保った背景画像生成を実現する手法を提案している。この手法では、観測画像セットに対して影の出方に応じた分類を行うことにより、各固有空間内での影の出る位置の違いによる背景の変動を吸収し、類似した画像セットごとの処理をすることにより、背景の変動に対処している。以下でこれについて簡単に手法を説明し、その問題点について述べる。

影の出方に応じた分類を行う際、画像の類似度からそれらを分類することは困難である。そのためこの手法では、日が異なっても同時刻に観測された画像ならば影の出る位置は同じであるということに着目し、長時間の観測によって得られた画像を、その観測時刻によって分類する。図3を見ると、ある一日中における観測画像は、異なる位置に影が出ているが、15時に着目すると、異なる日の15時であっても同じ位置に影が出ていることがわかる。

したがって、この手法は次のように書ける。ここでは、 d 日前の時刻 t における観測画像を $x(d, t)$ と書くことにする。まず、長時間の観測によって画像セット $I = \{x(d, t) | \forall d, t\}$ を得る。次に、画像セット I を観測時刻 t によって分類し、時刻 t に観測した画像セッ

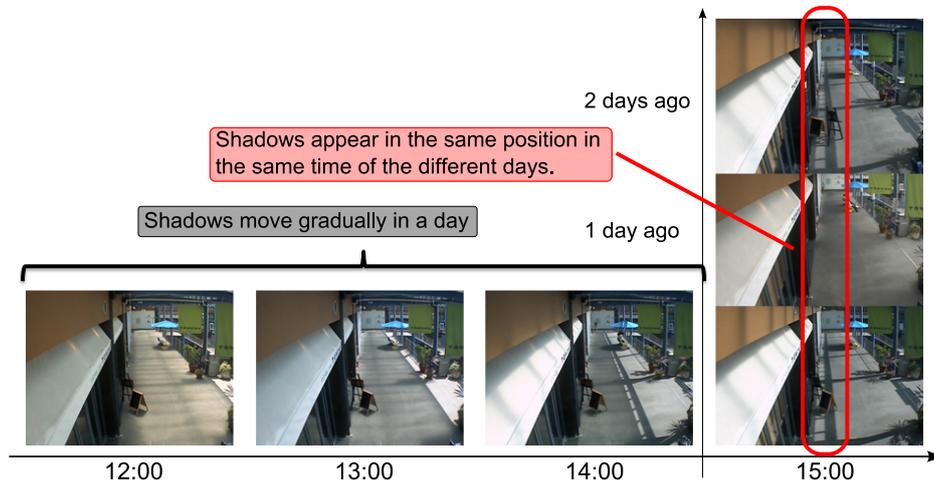


図 3 観測時刻が同じ画像ならば影の出方は同じ。

ト $I(t) = \{x(d, t) | \forall d\}$ を得る。こうして得られた全ての画像セット $I(t)$ に対し、主成分分析を行うことで任意の時刻 t において、 $I(t)$ から得られる固有ベクトル $e_1(t), e_2(t), \dots, e_j(t)$ によって張られる背景固有空間 $S(t)$ を得る。ここでは、前述の単純な固有空間法に基づく手法と同様、上位の k 次元のみを用いる。最後に、現在の観測画像 $x(0, t_c)$ に対し、それを観測した時刻 t_c に対応する固有空間 $S(t_c)$ を用いた固有空間法を適用する。これにより、 $x(0, t_c)$ から照明を保存し、かつ人物を確実に除去した背景画像 $x^b(0, t_c)$ を得ることができる(式(2))。

$$x^b(0, t_c) \approx E(t)E^T(t)(x(0, t_c) - x^{ave}(t)) + x^{ave}(t) \quad (2)$$

ただし、 $E(t) = [e_1(t) \ e_2(t) \ \dots \ e_j(t)]$ の正規直行行列、 $x^{ave}(t)$ は $I(t)$ の平均画像である。

この手法により、背景の日照の変化を再現した背景画像を生成することができる。しかし、この手法では背景の変動は日照の変動しかないものとしており、日ごとに置かれる場所が変化する看板のような物体の移動に伴うシーンの構造の変動については考慮されていない。また、固有空間の上位の k 次元のみを用いて背景画像を生成しているため、画像に与える影響の小さい背景の変動による影響は失われてしまっている。そのため、看板のような物体がシーン中に存在する場合、うまくそれらを背景画像中に再現できない。

3. 日照と構造の分離・統合による背景画像生成

3.1 日照の変動と構造の変動

2.4 節の最後に述べたように、屋外シーンにおいては、太陽の移動に伴う日照の変動と、看板など日ごとに置き場所が変わるような物体の移動による構造の変動が存在する。2.4 節では、観測して得た画像セットに対し日照が同一であるものだけを選び、それらのみに対して処理を

することで、日照の変動の影響を受けずに背景画像を生成していた。選び出した画像セット中に依然背景の変動が含まれているため、人物を確実に観測画像中から消すこれまでの手法ではうまく再現することができなかった。

ここで、背景の変動について整理する。まず、日照の変動は、太陽の動きに伴うため、観測時刻に応じて変化する。一方で、シーンの構造は日ごとに置き場を変えられる看板などに起因する。看板などが異なる日に画像上で全く同じ場所に置かれることは想定できないため、シーンの構造が同じである画像はその 1 日中の画像に限られる。そのため、大量の観測画像において、シーンの構造と日照が共に同一である画像はほとんど存在しない。ここで、これらの変動は互いに独立であるため、長期間観測して獲得した大量の画像セットにおいても、これらの変動は独立であると考えられる。

背景の変動において、曇りの時など日照の影響がない場合、各時刻の画像は非常に類似したものになる。そこでもし長時間観測して得た画像セットに対し、日照の変動を補正して何らかの値に固定できれば、日照の変動を含まず、シーンの構造の変動のみを含む画像セットが得られる。これに対し、シーンの構造の変動が同一である画像セットをさらに選び出して処理をすることで、背景の変動によらない背景画像生成が実現できると考えられる。

3.2 背景画像生成手法

3.2.1 概要

2.4 節で述べたように $I(t)$ は影の出る位置が同一である画像セットである。主成分分析により、この画像セット $I(t)$ の中で最も分散が大きな変動が、固有空間 $S(t)$ の上位の固有ベクトルに対応づけられる。屋外シーンにおいては、 $I(t)$ に晴れの日、曇りの日の画像が含まれている。天候の変化は、画像の広い範囲に大きな影響を与え、またその範囲は $I(t)$ の中では一定である。一方、人物が移動したり、看板などの物体が移動することによる画像への影響は、狭い範囲であり、またその範囲はランダムで



図 4 日照の影響と重み $p_1(t)$ の関係

ある．そのため，画像セット $I(t)$ の中で最も分散の大きな変動は，そのシーンにおける日照の強弱となると考えられる．

そこで，観測画像 $x(0, t)$ を固有空間 $S(t)$ に投影し，得られるベクトル $\mathbf{p}(0, t) = (p_1(t), p_2(t), \dots)$ の上位の固有ベクトルの重み $p_1(t)$ を日照の影響がない時の値に置き換えることで，日照の影響が除去された画像を得ることができる．こうして得た日照の影響がない画像セットのうち，シーンの構造が同一である同日の画像セットを選び出して平均化することで人物が除去された画像を得る．最後に，こうして得た人物が除去された画像に対し，固有空間中の照明に対応する重み $p_1(t)$ を，現在の観測画像と同じものに置き換えることにより，人物が除去され，シーンの構造が保たれ，日照の様子が正しく反映された背景画像を得ることができる．

以下の節で，日照と構造の分離および統合の各手順について詳述する．

3.2.2 固有空間を用いた日照と構造の分離

ここでは，観測画像 $x(0, t)$ から日照の影響を除去することを考える．先に述べたように，最も分散が大きな変動に対応するベクトル $\mathbf{e}_1(t)$ には日照の影響が対応している．これにより，その重み $p_1(t)$ を変化させることで，画像中の日照を変化させることができる（図 4）． $\mathbf{p}(0, t)$ の $p_1(t)$ を置き換えたベクトルを画像に逆投影すると，固有空間の次元 k より高次元の成分が失われてしまい，背景画像の再現性が低下してしまう．そこで固有空間中で $p_1(t)$ の値を置き換えるのではなく，置き換え前後の差ベクトルを画像空間に逆投影し，それを元の観測画像から引くことで，観測画像の高次元の成分を失うことなく背景画像から日照成分を除去する．

このとき，日照の影響を除去するため単純に $p_1(t)$ を 0（平均画像と同じ）としても，完全に日照の影響を除去することはできない（図 4(ii)）．なぜなら，固有空間の平均画像を求める際，晴れの日（日照がある時）や曇りの日（日照の影響を受けていない時）の平均をとっているため平均画像 $\mathbf{x}^{\text{ave}}(t)$ （このとき $p_1^{\text{ave}}(t) = 0$ ）は日照の変動の影響を受けているからである．そのため，平均画像 $\mathbf{x}^{\text{ave}}(t)$ は，日照の影響を受けていない画像よりも明るい画像になっている（図 4）．日照の変動の影響を完全に除去するためには，上位の固有ベクトルの重み $p_1(t)$ を 0 に置き換えるだけでなく，さらに平均画像が受けている日照の変動によるオフセット $p_1^{\text{offset}}(t)$ 分の補正

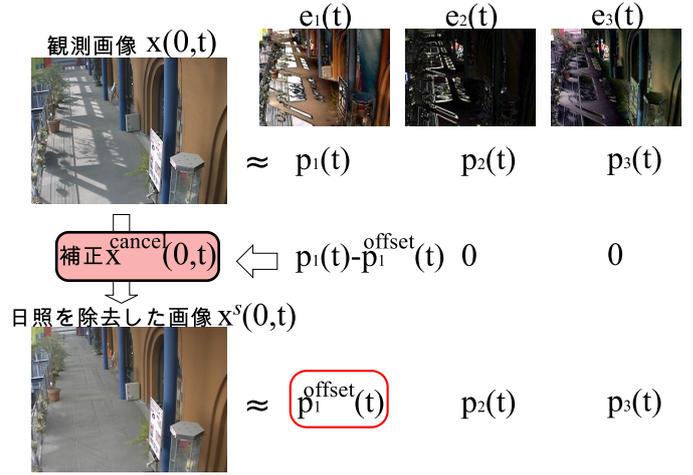


図 5 日照成分の除去

が必要となる．そのためには，まず $p_1^{\text{offset}}(t)$ を推定する必要がある．

この平均画像 $\mathbf{x}^{\text{ave}}(t)$ が受けている日照の変動の影響 $p_1^{\text{offset}}(t)$ は，背景の固有空間 $S(t)$ を生成するとき用いた画像セット $I(t)$ のうち最も曇っているとき，つまり背景が全く日照の影響を受けていないときの $p_1(t)$ の値である．日照の変動の影響 $p_1^{\text{offset}}(t)$ の正負は固有ベクトル $\mathbf{e}_1(t)$ の方向に依存する．そこで， $\mathbf{e}_1(t)$ を，各要素の総和が正となるように方向を定めたとき，各画像 $x(0, t)$ の重み $p_1(t)$ のうち，小さいものは日照の影響が少ない画像となる．頑健に推定を行うため，複数枚の $p_1(t)$ の小さい画像を用い，複数の $p_1(t)$ の平均処理によってこの $p_1^{\text{offset}}(t)$ を推定する．

次に，観測画像から日照の影響を除去する方法について説明する（図 5）．まず，同一時刻の観測画像セット $I(t)$ から得られる固有空間 $S(t)$ に対し時刻 t における観測画像 $x(0, t)$ を投影することで，固有空間 $S(t)$ 中の点 $\mathbf{p}(0, t) = (p_1(t), p_2(t), p_3(t), \dots)$ を得る（式 (3)）．

$$\mathbf{p}(0, t) \approx E(t)(\mathbf{x}(0, t) - \mathbf{x}^{\text{ave}}(t)) \quad (3)$$

この $p_1(t)$ は，現在の観測画像 $x(0, t)$ が受けている日照の影響である．そこで，この $p_1(t)$ と，3.2.2 で求めたオフセット $p_1^{\text{offset}}(t)$ を用い，ベクトル $\mathbf{p}^{\text{cancel}}(0, t) = (p_1(t) - p_1^{\text{offset}}(t), 0, 0, \dots)$ を画像へと逆投影し，照明成分を除去するための画像 $\mathbf{x}^{\text{cancel}}(0, t)$ を得る（式 (4)）．

$$\mathbf{x}^{\text{cancel}}(0, t) = E\mathbf{p}^{\text{cancel}}(0, t) \quad (4)$$

最後に，式 (5) により，観測画像 $x(0, t)$ から照明成分を除去した画像 $x^s(0, t)$ を得る（以下，上付きの s は日照の影響を除去したこと，つまりシーンの構造を意味する）．

$$\mathbf{x}^s(0, t) = \mathbf{x}(0, t) - \mathbf{x}^{\text{cancel}}(0, t) \quad (5)$$

現在の観測画像 $x(0, t_c)$ と同日に観測された観測画

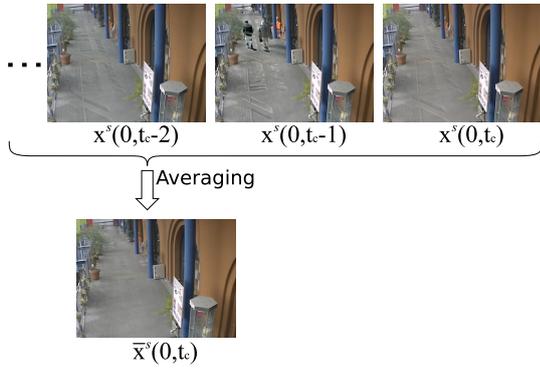


図 6 構造を保存した背景画像生成

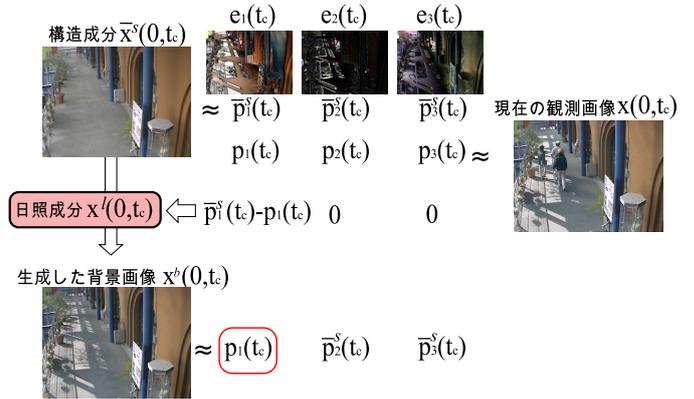


図 7 構造成と照明成分の統合

像セット $\{x(0, t) | t < t_c\}$ から日照の影響を除去することにより、日照の影響を受けていない画像セット $I^s(t_c) = \{x^s(0, t) | t < t_c\}$ を得る。この画像セット $I^s(t_c)$ は日照の影響を受けていないので、画素毎の平均化処理をすることにより、シーンの構造を保ちつつ人物を除去した画像 $\bar{x}^s(0, t_c)$ を獲得することができる (図 6)。

3.2.3 固有空間を用いた日照と構造の統合

構造は保存しているが日照の様子を反映していない背景画像が生成されるが、これに対し、固有空間法を用いて現在の日照の様子を反映させることを考える (図 7)。

背景画像生成の対象である現在の観測画像 $x(0, t_c)$ に対応する固有空間 $S(t_c)$ に対し、画素毎の平均化処理によって得られた背景画像 $\bar{x}^s(0, t_c)$ を投影し、対応するベクトル $\bar{p}^s(0, t_c) = \{\bar{p}_1^s(t_c), \bar{p}_2^s(t_c), \dots\}$ を得る (式 (6))。

$$\bar{p}^s(0, t_c) = E^T(t_c) \bar{x}^s(0, t_c) \quad (6)$$

一方で、現在の観測画像 $x(0, t_c)$ も同様に固有空間 $S(t_c)$ へと投影し、対応するベクトル $p(0, t_c) = \{p_1(t_c), p_2(t_c), \dots\}$ を得る (式 (7))。

$$p(0, t_c) \approx E(t_c)^T x(0, t_c) \quad (7)$$

前節で生成した背景画像 $\bar{x}^s(t_c)$ の日照の影響を現在の観測画像の日照の影響で置き換えるためのベクトル $p^1(t_c) = \{\bar{p}_1^s(t_c) - p_1(t_c), 0, 0, \dots\}$ を画像へと逆投影することで、 $x^1(0, t_c)$ を得る (式 (8))。

$$x^1(0, t_c) = E(t_c) p^1(t_c) \quad (8)$$

この $x^1(0, t_c)$ を用いて、 $\bar{x}^s(0, t_c)$ を式 (9) によって補正することにより、背景画像 $x^b(0, t_c)$ を得る。

$$x^b(0, t_c) = \bar{x}^s(0, t_c) - x^1(0, t_c) \quad (9)$$

こうして得られた背景画像 $x^b(0, t_c)$ は、人物を含まず、背景の構造をできるだけ保ち、かつ現在の日照をできるだけ再現した背景画像となっている。

4. 実験

4.1 背景画像生成

提案手法の有効性を示すため、複数の屋外シーンにお

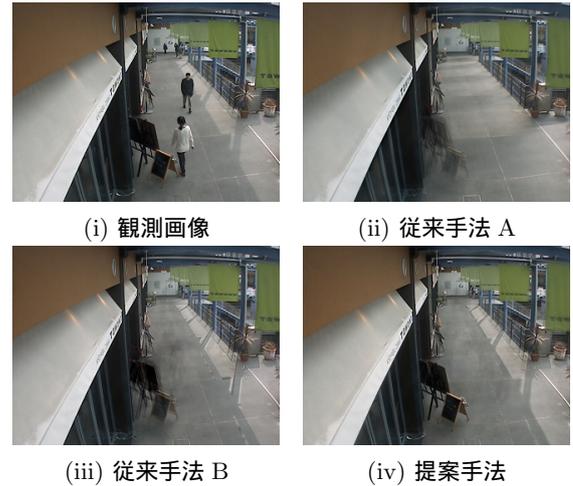


図 8 背景画像生成結果 (シーン 1)

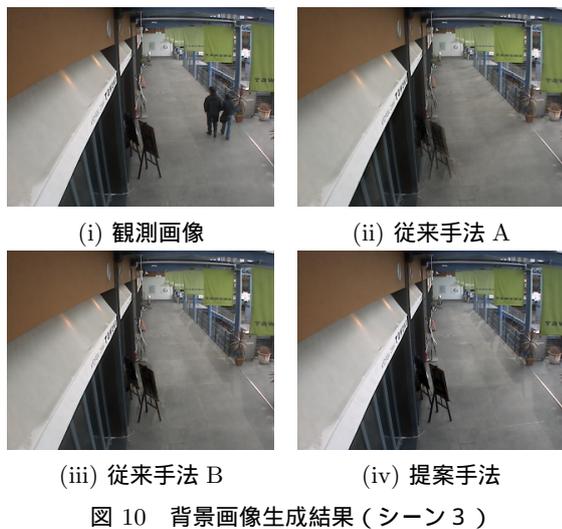
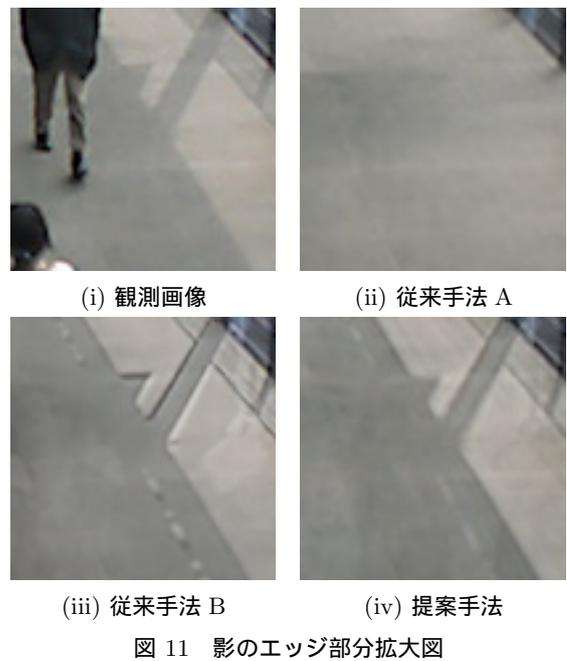
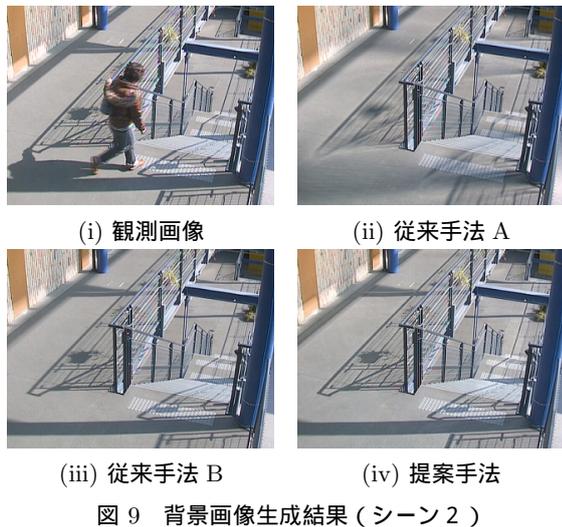
いて、次の 3 つの手法により背景画像を生成した。

- 従来手法 A: 単純に全ての画像から生成した固有空間を用いる背景画像生成手法。
- 従来手法 B: 2.4 節で説明した、観測時刻毎に画像を分類した上で、観測時間毎に生成した固有空間を用いる従来手法
- 提案手法: 固有空間を用いてシーンの日照と構造を分離・統合する提案手法

また、対象のシーンとして、次の 2 つを用意した。

- シーン 1: 商業施設の廊下。レストランの前の看板は毎日移動される。背景画像の生成は、2月22日13時15分 ($t_c = 13$ 時 15 分) に観測した画像について行った。
- シーン 2: シーン 1 と同じ商業施設内の異なる廊下。背景画像の生成は、2月22日11時05分 ($t_c = 11$ 時 05 分) に観測した画像について行った。
- シーン 3: シーン 1 と同じ場所。背景画像の生成は、シーン 1 とは天候が異なる 3月17日13時05分 ($t_c = 13$ 時 05 分) に観測した画像について行った。

それぞれの結果を図 8,9,10 に示す。従来手法 B および提案手法では、背景固有空間 $S(t)$ の生成に、シーン 1 では 13 時 15 分の、シーン 2 では 11 時 05 分の、シーン 3 では 13 時 05 分の画像を 7 日分 (およそ 1000 枚) 用



いた。また、提案手法では、背景固有空間 $S(t)$ を 1 分ごとに生成し、各分 100 枚ずつの画像に対し日照の影響を除去し、5 分おきに合計およそ 1000 枚の画像を平均化処理に用いた。背景の固有空間 $S(t)$ の次元数 k は [8] で最適とされている 4 次元を用いた。

結果、日照の再現性については、従来手法 A では再現できず、影のエッジがぼけてしまっていた日照の影響が、従来手法 B と提案手法では再現出来ていることがわかる。強い日照によって影が出ている場合 (図 8,9) でも、影のない場合 (図 10) でもうまく日照の影響を反映した画像が生成出来ている。図 11 に、時間の経過とともに連続的に移動していく日照のエッジ部分を拡大したものを示す。

一方、構造の再現性については、従来手法 B では半透明な看板が現れてしまっているのに対し、提案手法では比較的うまく看板が再現されていることがわかる。図 12 に、毎日おかれる場所が変わる看板部分を拡大したものを示す。

提案手法ではいずれの場合も、現在の観測画像に非常に類似した画像を生成出来ていることが確認できる。

次に、本手法の失敗例を示す (図 13)。本手法では、シーンの構造は 1 日ごとに変化し、ある 1 日の中では不変であると仮定しているが、図 13 の例では、1 日の間に看板がスタンドを残して撤去されており、本手法の仮定から外れている。それにも関わらず撤去前の画像と撤去後の画像の両方を含む画像群に対して平均化処理が行われているため、シーンに存在しないはずの看板が生成画像に現れており、うまく再現出来ていない。

4.2 背景画像生成精度の評価

背景画像生成精度の数値的な比較を行うため、図 8 の生成結果のうち、人物や旗などを含まない領域に対して観測画像と各手法で生成した背景画像との差をとることによって誤差を評価した。図 14 は、観測画像との差分



(i) 観測画像 (ii) 生成した背景画像
図 13 背景画像の失敗例

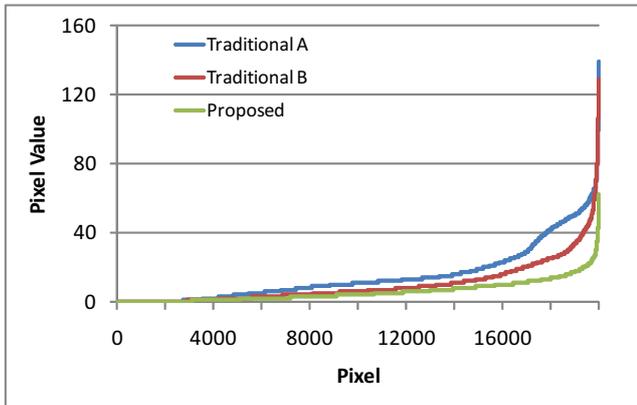


図 14 生成した画像と観測画像の画素値の差

結果である画像の画素を、画素値の小さい順にならべたものである。

従来手法 A に対し、従来手法 B と提案手法は全体的に差分画像の画素値が小さいことがわかる。これは、広い領域における日照変化を良く再現出来ていることを表していると考えられる。一方、図 14 のグラフ右側（16000 画素を超えたあたりから）では、従来手法 A と従来手法 B において誤差が大きくなっているのに対し、提案手法では比較的小さなものとなっている。このように一部の画素で大きな誤差が起きているのは、看板などの再現性により、従来手法では小さい領域での大きな誤差が起きているのに対し、提案手法ではそれを小さく出来ていることに対応していると考えられる。

5. おわりに

本研究では、そのシーンの日照と、シーン中に置かれている物体によって決まるシーンの構造を固有空間法を用いて分離し、同一時刻の画像群の日照成分を用いて日照を保った画像生成を行い、また同日の画像群の構造成分を用いてシーンの構造を保った画像生成をした上で、それらの結果を固有空間法を用いて統合することにより、シーンの日照と構造を保存した背景画像生成を実現した。いくつかのシーンにおいて、提案手法と従来手法によって背景画像を生成し、従来手法との比較により本手法の有効性を確認した。

提案手法では、日照成分は日照の様子が同一な同時刻の画像を用いて生成し、構造成分は構造が同一な直近の観測画像を用いて生成することにより、日照成分と構

造成分をそれぞれ再現した。

本研究では、シーンの構造の変化は看板など日によって置き場を変えられるものであるとし、1 日の間にはシーンの構造は変化しないと仮定していた。こうした物体の移動を検出し、平均化処理に移動前の画像をできるだけ用いないようにすることによって、シーンの構造が 1 日の間で変化したとしても背景画像をうまく再現することが今後の課題である。また、風になびく木や旗など、常に動いているような背景物体を、現在の観測画像の情報を有効活用して再現し、よりリアルタイムな情報を伝えることができる画像を生成することも今後の課題である。

謝 辞

本研究は、文部科学省の科学技術振興調整費（科学技術連携施策群の効果的・効率的な推進）による「センサ情報の社会利用のためのコンテンツ化」の一環として実施したものである。

文 献

- [1] 美濃：“センシングウェブ：概念と課題（<特集>センシングウェブ）”，人工知能学会誌，**24**，2，pp. 179–184 (20090301).
- [2] M. Minoh, K. Kakusho, N. Babaguchi and T. Ajsjaka: “Sensing Web Project-How to handle privacy information in sensor data”, 12th International Conference on Information Processing and Management of Uncertainty in Knowledge-based Systems (2008).
- [3] M. Bertalmio, G. Sapiro, V. Caselles and C. Ballester: “Image inpainting”, SIGGRAPH '00: Proceedings of the 27th annual conference on Computer graphics and interactive techniques, New York, NY, USA, ACM Press/Addison-Wesley Publishing Co., pp. 417–424 (2000).
- [4] C. R. Wren, A. Azarbayejani, T. Darrell and A. P. Pentland: “Pfinder: Real-time tracking of the human body”, IEEE Transactions on Pattern Analysis and Machine Intelligence, **19**, pp. 780–785 (1997).
- [5] B. Lo and S. Velastin: “Automatic congestion detection system for underground platforms”, Intelligent Multimedia, Video and Speech Processing, 2001. Proceedings of 2001 International Symposium on, pp. 158–161 (2001).
- [6] R. Cucchiara, C. Grana, M. Piccardi and A. Prati: “Detecting moving objects, ghosts, and shadows in video streams”, IEEE Transactions on Pattern Analysis and Machine Intelligence, **25**, pp. 1337–1342 (2003).
- [7] N. Oliver, B. Rosario and A. Pentland: “A bayesian computer vision system for modeling human interactions”, IEEE Transactions on Pattern Analysis and Machine Intelligence, **22**, pp. 831–843 (1999).
- [8] Y. Kawanishi, I. Mitsugami, M. Mukunoki and M. Minoh: “Background image generation by preserving lighting condition of outdoor scenes”, Procedia - Social and Behavioral Sciences, **2**，1，pp. 129 – 136 (2010).