

夜間における景観の複合現実感型提示

清水 諒[†] 北原 格[‡] 亀田 能成[‡] 大田 友一[‡]

[†] [‡] 筑波大学大学院システム情報工学研究科 〒305-8573 茨城県つくば市天王台 1-1-1

E-mail: [‡] {kitahara,kameda,ohta} @iit.tsukuba.ac.jp

あらまし 本研究では、カメラ付きモバイル端末を用いて、周囲を見渡すような映像に合わせた複合現実型の夜間映像を提示する。複合現実型の夜間映像の合成にあたっては、夜近景（夜間映像中の近景部分）と昼近景（事前撮影映像中の近景部分）を手がかりに、夜近景と昼間に予め撮影した事前撮影映像中の遠方の景観を幾何的な位置合わせをして合成する。本研究では、事前撮影映像を用意する際、GPS センサ、方位センサ、加速度センサなどモバイル端末に搭載されているセンサを用いて、撮影時のカメラの位置、姿勢情報を映像と合わせて記録し、それらの情報を参照することで最も近い条件で撮影された事前撮影映像を選べる状況であることを仮定する。夜間映像と事前撮影映像の撮影条件にある程度の未知の位置と方位のズレがあるという下で、画像処理に基づく幾何的な位置合わせによって複合現実型の夜間映像の合成を実現する。そのために、映像中に存在する自然特徴点を用いて、夜近景と昼近景の対応点探索を行う。

キーワード 拡張現実, 対応点探索, 夜景, 画像合成, 屋外

Mixed Reality Presentation of Scenery View at Nighttime

Makoto SHIMIZU[†] Itaru KITAHARA[‡] Yoshinari KAMEDA[‡] and Yuichi KITAHARA[‡]

[†] [‡] Graduate School of Systems and Information Engineering, University of Tsukuba 1-1-1 Tennoudai, Tsukuba, Ibaraki, 305-8573, Japan.

E-mail: [‡] {kitahara,kameda,ohta} @iit.tsukuba.ac.jp

Abstract We propose a new method to synthesize a night view in Mixed Reality fashion in which distant scenery of a daytime view is integrated. When it is dark at night, it is difficult to see distant scenery. On the other hand, near-range view scenery could be visible at night because of some light nearby. In our approach distant scenery of daytime view is aligned to near-range scenery of night view by precise geometric registration. When the night view and the daytime view are taken at same direction but at slightly different location, or when the two views are taken at slightly different direction at the same location, two views could be geometrically aligned by certain amount of planar transition on images. Our method estimates the amount of the planar transition by feature point matching between the two views. For preparing daytime views, we shoot the same scene for several times at different sunlight conditions to improve the accuracy of the synthesis.

Keyword Augmented Reality, Image feature matching, Night view, Image synthesis, Outdoor

1. はじめに

屋外において、遠方の風景や建築物などの景観は、昼間では視認することができても、夜間では照明等の光源がなければその景観を視認することが難しい。そこで、利用者から見て遠方の風景がどのようになっているかを夜間でも確認することができれば、屋外における新しい視覚支援となりうる。この時、利用者のすぐ近くには、街灯などの光源が存在することがあり、近傍の景観は夜間でも視認可能な場合が考えられる。

近景が昼間でも夜間でも見えることを手がかりに、夜間において近景と遠方の景観を同時に把握可能な複

合現実型の夜間映像の提示方法を提案する。夜間に撮影した映像に対して、近景としては夜間映像をそのまま用い、遠方の景観としては昼間に撮影した事前撮影映像中にある遠方の景観に相当する部分を合成する。

本研究では、カメラ付きモバイル端末を用いて、周囲を見渡すような映像に合わせた複合現実型の夜間映像を提示する。複合現実型の夜間映像の合成にあたっては、夜近景（夜間映像中の近景部分）と昼近景（事前撮影映像中の近景部分）を手がかりに、夜近景と昼間に予め撮影した事前撮影映像中の遠方の景観を幾何的な位置合わせをして合成する。一般に夜間映像に対

して事前撮影映像の位置，姿勢などの撮影条件が同じであるとは限らない．そのため，映像を合成するには夜間映像とできる限り近い条件で撮影された事前撮影映像を用いることが望ましい．本研究では，事前撮影映像を用意する際，GPS センサ，方位センサ，加速度センサなどモバイル端末に搭載されているセンサを用いて，撮影時のカメラの位置，姿勢情報を映像と合わせて記録し，それらの情報を参照することで最も近い条件で撮影された事前撮影映像を選べる状況であることを仮定する．

夜間映像と事前撮影映像は上に述べた通り，撮影時の条件が完全に一致しているとは限らない．GPS センサ，方位センサ，加速度センサの値から，二つの映像の撮影時におけるカメラの位置と姿勢のズレの大きさを求めることができるが，それらのセンサの精度は映像合成に十分な精度があるとは言えない可能性がある．

そこで，本研究では夜間映像と事前撮影映像の撮影条件にある程度の未知の位置と方位のズレがあるという下で，幾何的な位置合わせによって複合現実型の夜間映像の合成を実現する．そのために，映像中に存在する自然特徴点を用いて，夜近景と昼近景の対応点探索を行う．

2. 関連研究

風景や建築物などの景観は，昼間に視認することはできても，夜になって街灯や照明などの光源が存在しない場合，それらを視認することは難しい．

Raskar ら [1] や Cai ら [6] は固定カメラを用いて，夜間映像で視認が難しい部分を事前撮影映像から合成した映像の生成手法を提案している．しかし，この手法で

はカメラの移動は全く考慮されていない．

モバイル端末上で夜間に撮影された映像と予め昼間に撮影された事前撮影映像から複合現実型の夜間映像を合成する際，2 つの映像が撮影された際のカメラの位置や姿勢などの撮影条件は完全に一致しているとは限らない．そのため，夜近景と昼近景の対応関係を手がかりに，夜間映像と事前撮影映像を幾何的な位置合わせをして合成する必要がある．異なる時間帯で撮影された映像間の対応関係を自然特徴点の対応点探索により算出する手法を Yazawa ら [2] は提案している．しかし，夜間映像と事前撮影映像では日照変動により特徴量が大きく変化してしまうため，正しい対応関係を得るには困難が予想される．

そこで，本研究では昼間，午後，日没など複数の異なる時間帯の事前撮影映像が用意されているものとし，夜間映像と事前撮影映像の間で対応点探索を行うことで夜近景と昼近景の正しい対応関係を取得し，幾何的な位置合わせによる複合現実型の夜間映像の合成を行う．

3. 複合現実型の夜間映像提示

屋外において，夜間の遠方の景観は街灯や照明などの光源が存在しなければ視認することが難しい．その一方で近景は夜間でも光源が存在することがあり，視認することは容易である．

そこで，夜間にカメラ付きモバイル端末で撮影を行い，夜近景はそのままにして遠方の景観のみを事前撮影映像からのものとするすることで，夜近景と事前撮影映像による遠方の景観を同時に見ることが可能な複合現実型の夜間映像提示を実現する．図 1 に提案手法の概

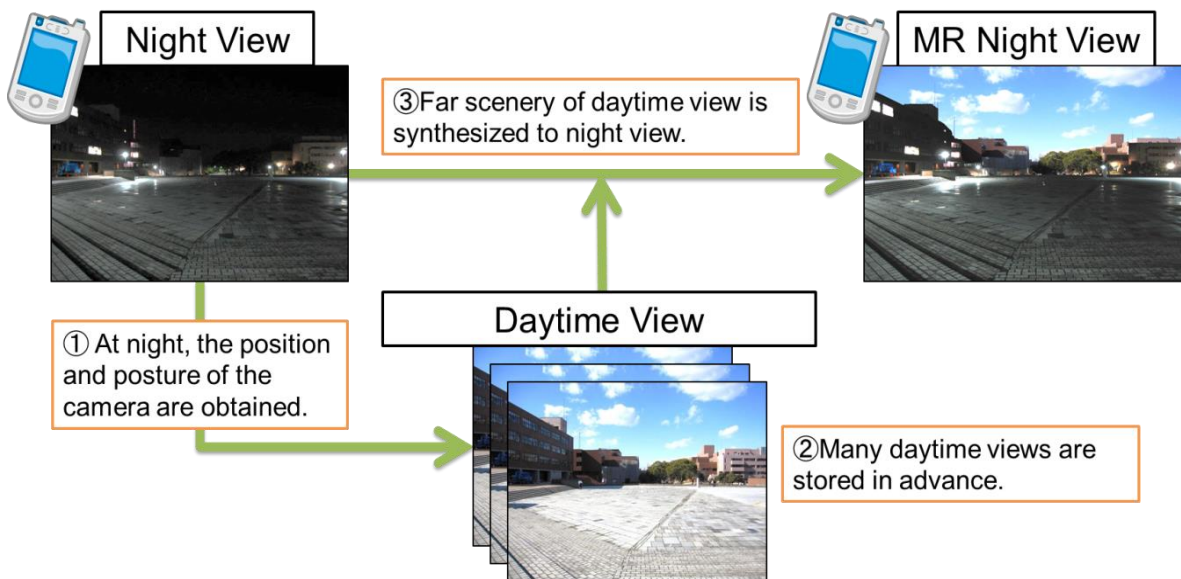


図 1: 提案手法の概要

要を示す。

夜間映像における遠方の景観を合成するために、昼間のうちに事前撮影映像を予め用意しておく。想定される区域内では多くの位置、姿勢で撮影が行われているものとする。幾何的な位置合わせによる複合現実型の夜間映像の合成を行うために、カメラの位置、姿勢などの撮影条件は事前撮影映像と夜間映像の間で限りなく近いことが望ましい。そこで、事前撮影映像の撮影状況を参照するために、撮影時のカメラの位置、姿勢、内部パラメータなどを映像と合わせて記録しておく。

本研究では夜間映像と事前撮影映像の撮影条件に位置と方位のズレのみが生じた状況を想定し、この状況下で夜近景と昼近景の対応関係を画像処理によって求め、幾何的な位置合わせによる複合現実型の夜間映像の合成を行う。以降本研究では、位置と方位以外の撮影条件は一致しているものとする。

4. 幾何的な位置合わせによる複合現実型の夜間映像の合成

本節では位置と方位のズレが夜近景と事前撮影映像の遠方の景観を合成する際の幾何的な位置合わせに与える影響を検証する。また、夜近景と昼近景の対応関係を画像処理によって求め、幾何的な位置合わせによって複合現実型の夜間映像の合成を実現する方法について説明する。

4.1 位置と方位のズレが位置合わせに与える影響

撮影対象となる物体がカメラから十分遠くに存在する場合、遠方の平面上に存在していると近似して考えることができる。本研究では近景と遠方の景観は平面への近似が成立する距離に存在する物体で構成されているものとする。位置と方位のズレがあった場合、映像中の近景と遠方の景観の水平方向の移動を平面の平行移動に近似する。本節では位置と方位のズレのそれぞれに関してどの程度平行移動が生じるのか検証する。

4.1.1 位置ズレによる影響

図2に示すように、遠方の景観を構成する物体AがカメラC₀から距離L_Aの位置に存在する場合を考える。この時、物体AとカメラC₀の光軸のなす角をθ_Aとする。光軸に対して垂直横方向に位置ズレがあった場合のカメラをC₁、カメラC₀、C₁の焦点距離をfとする。位置ズレがΔpあった場合、物体Aの映像中の水平移動量d(L_A, Δp)は式(1)で表される。今、水平画角θ[deg]、映像の水平幅1の正規化カメラを考えると、d(L_A, Δp)に対応する映像中の水平移動量T_Hは式(2)で表される。すな

わち、映像中の物体はΔpによって一様に水平移動する。

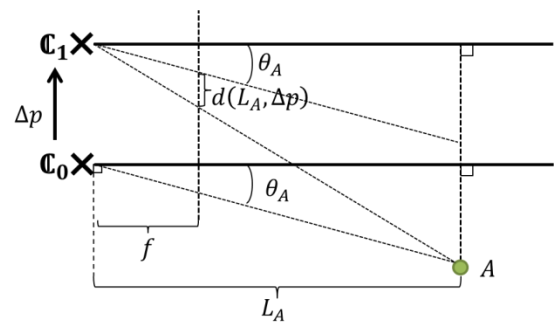


図 2: 左右の並行移動

$$d(L_A, \Delta p) = f \left(\tan \theta_A + \frac{\Delta p}{L_A} - \tan \theta_A \right) = \frac{f}{L_A} \Delta p \tag{1}$$

$$T_H = \frac{d(L_A, \Delta p)}{2f \tan \frac{\theta}{2}} = \frac{\Delta p}{2L_A \tan \frac{\theta}{2}} \tag{2}$$

位置を測定する GPS センサでは一般的に 5[m]から 15[m]の誤差が生じることが知られている[3][4]。そこで L_A = 1000[m]、θ = 52[deg]とし、Δpを 5[m]から 15[m]まで変化させた場合の水平移動量T_Hを図3(上)に示す。また、T_HはL_Aによって平面へ近似する際の精度が変化するため、L_Aを 200[m]から 1000[m]まで変化させた場合のT_Hを図3(下)に示す。

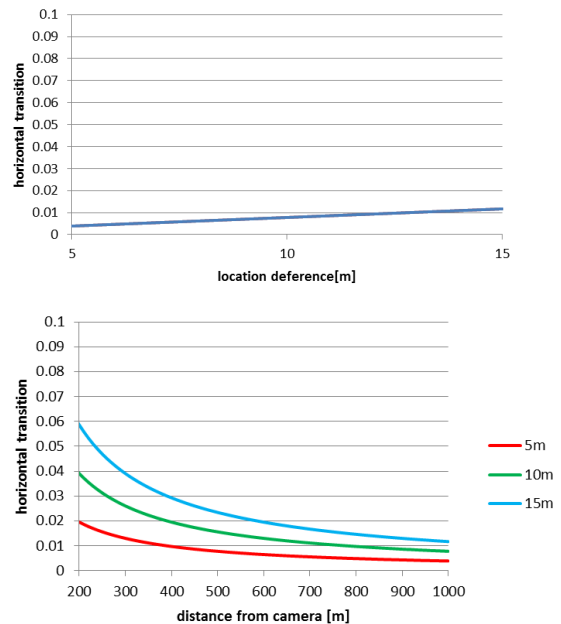


図 3: 左右の平行移動による影響

カメラからの距離が遠い地点を撮影するほど、水平位置ズレによって生じる T_H は小さくなることが分

かる.そのため,遠方の景観が十分遠くに存在すれば,位置ズレによる T_H への影響は小さいと考えられる.

次に,カメラが光軸に対して鉛直方向に位置ズレ Δp があつた場合を考える. この場合も映像中での鉛直移動量 T_v は T_H と同様に求めることができる. ただし, 実際には緯度, 経度から GIS によって高さを得られるものとする, この Δp は小さい値と想定されるため, T_v を 0 とみなす.

前後位置ズレは映像中のスケール変化に影響を及ぼす. 図 4 に示すようにカメラ C_0 の光軸に沿って前後位置ズレ Δp があつた場合のカメラを C_1 とする. カメラ C_0, C_1 の焦点距離を f とした場合, 物体 A から物体 B までの距離の映像中のスケール変化 $s(L_A, \Delta p)$ を式(3)で表す. ただし, $L'_A = L_A + \Delta p$ とする. $|\Delta p| \ll L_A$ より, 式(3)は Δp について定数と見なせるため, スケールはほぼ変化しないと考えることができる.

以降, 本稿では垂直位置ズレと前後位置ズレによる影響はいずれも小さいと考えて無視し, 水平位置ズレのみを考慮する.

$$s(L_A, \Delta p) = f \tan \theta_A \left(1 - \frac{L_A}{L_A + \Delta p}\right) \quad (3)$$

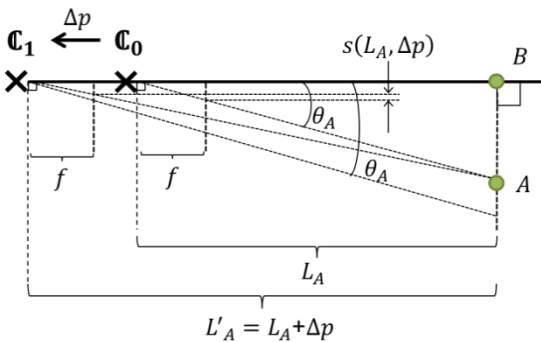


図 4: 前後の並行移動

4.1.2 方位ズレによる影響

図 5 に示すように, 遠方の景観を構成する物体 A がカメラ C から距離 L_A の位置に存在する場合を考える. 物体 A とカメラ C の光軸のなす角を θ_A とする. 方位ズレが Δr あつた場合の物体 A とカメラの距離を $[\mathit{L}'_A]$ とする. 物体 A とカメラ C の光軸のなす角は $\theta_A + \Delta r$ で表される. カメラ C の焦点距離を f とした場合, Δr が十分小さい条件の下, 物体 A の映像中の水平移動量 $e(\Delta r)$ は式(4)で表される. 式(4)より水平面角 θ [deg] の正規化カメラにおける映像中の水平移動量 T_R は θ_A が 0 に近いと仮定すると式(5)で表される. 式(5)より映像中の物体は Δr によって一様に水平移動する.

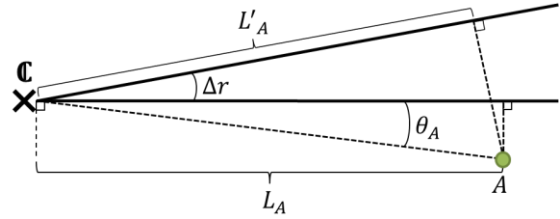


図 5: 水平回転

方位を測定する方位センサでは一般的に 5° 程度の誤差が生じることが知られている [6]. そこで, $\theta = 52[\text{deg}]$ とし, Δr を $0[\text{deg}]$ から $5[\text{deg}]$ まで変化させた場合の T_R を図 6 に示す.

$$e(\Delta r) = f \tan(\theta_A + \Delta r) - f \tan \theta_A \approx f \Delta r \tan' \theta_A \quad (4)$$

$$\begin{aligned} T_R &= \frac{e(\Delta r)}{2f \tan \frac{\theta}{2}} \\ &= \frac{f \Delta r \frac{1}{\cos^2 \theta_A}}{2f \tan \frac{\theta}{2}} \\ &= \frac{\Delta r}{2 \tan \frac{\theta}{2}} \end{aligned} \quad (5)$$

$\theta_A \approx 0$ の場合, $\cos^2 \theta_A \approx 1$ となるため

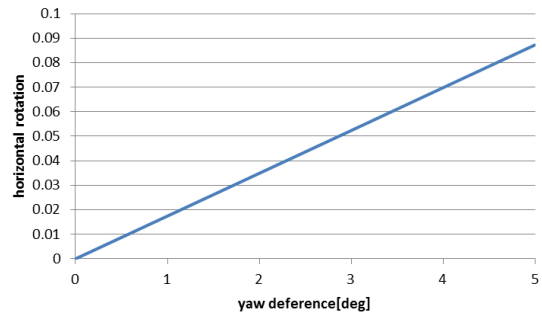


図 6: 水平回転による影響

方位ズレが 5° 生じた場合, R は 0.09 となる. 位置ズレの時と異なり, 方位ズレによる影響は大きいと考えられる. これらの位置ズレ, 方位ズレによる映像中の水平移動量を調べることができれば, 2 つの映像の位置合わせが可能となる.

4.2 夜近景と昼近景の対応関係の算出

4.1 節で述べた通り, 夜間映像と事前撮影映像の間に位置ズレと方位ズレがある場合, これらを平行移動として考え, 事前撮影映像から夜間映像への 2 次元変換行列の平行移動成分で水平移動量を表す. 2 次元変

換行列は映像間の自然特徴点の対応点探索により求めることが可能である。

そこで、以下の手順で対応点を求める。まず、高速に特徴点の検出が可能なアルゴリズムである FAST[7] を用いて、昼間の映像と夜間映像に存在する特徴点を検出する。そして、検出した特徴点に対して、SURF[8] による特徴量の記述を行う。夜間映像で検出された L 個の特徴点、事前撮影映像で検出された M 個の特徴点について、その特徴量ベクトルの距離が最小となる特徴点の組み合わせを求める。

対応点組の中には誤対応も多く存在している。幾何的な位置合わせによる合成のために、対応点組の中から誤対応を除去して 2 次元変換行列 A^* の推定を行う。対応点組の誤対応の除去には RANSAC[9] によるロバスト推定を用い、対応点組に対して 2 次元平行移動拘束を与える。これにより昼間の映像から夜間映像への 2 次元変換行列 A^* を推定する。

4.3 複合現実型の夜間映像の合成

周囲を見渡すようにして夜間の景観を撮影する際、夜近景は映像の下部、遠方の景観は映像の上部に存在していると考えられる。このような状況で撮影された夜間映像を想定し、映像の下部は夜間映像、映像の上部は昼間の映像で表現されるような重みを用いて各画素の画素値をブレンドすることで、夜間映像と事前撮影映像の遠方の景観の合成を行う。

高さ 1 ($0 \leq y \leq 1$, ただし映像の上端を 0 , 下端を 1 とする) の映像において、 $y < 1/4$ は昼間の映像のみ、 $y \geq 3/4$ は夜間映像のみで表現されることを想定し、ある点における夜間映像の重み ω_N と事前撮影映像の重み ω_D を式(6)(7)で定義する。

高さ 1 の映像を想定した場合の ω_N, ω_D を図 7 に示す。

$$\omega_N = \begin{cases} 0 & (y < \frac{1}{4}) \\ \sqrt[3]{2y - \frac{1}{2}} & (\frac{1}{4} \leq y < \frac{3}{4}) \\ 1 & (y \geq \frac{3}{4}) \end{cases} \quad (6)$$

$$\omega_D = 1 - \omega_N \quad (7)$$

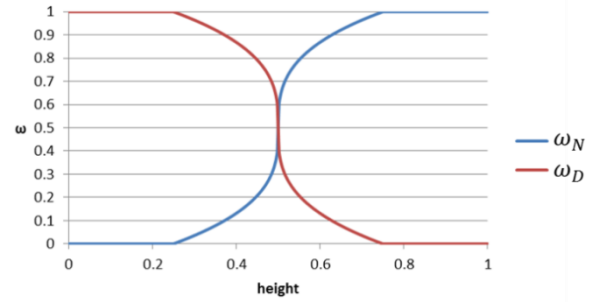


図 7: 画像合成時の重み付け

夜間映像の座標 (x,y) における画素値を $N_{(x,y)}$, 事前撮影映像の座標 (x,y) における画素値を $D_{(x,y)}$ で表し、以下のように画素値をブレンドすることで、複合現実型の夜間映像の画素値 $C_{(x,y)}$ を求める。

$$C_{(x,y)} = \omega_D D_{(x,y)} + \omega_N N_{(x,y)} \quad (8)$$

5. 複数の異なる時間帯に撮影された事前撮影映像と夜間映像の対応点探索

4 節では、夜間映像と事前撮影映像で SURF による対応点探索を行い、2 次元変換行列 A^* を求めることで幾何的な位置合わせをした合成手法を提案した。

SURF は照明変動にロバストなアルゴリズムではあるが、昼間と夜間では日照変動による特徴量の変化が大きく、対応点探索の精度が著しく低下する。合成をする際、幾何的な位置合わせが可能な精度の A^* を推定するために十分な対応点組を得ることができない場合がある。

そこで、事前撮影映像には撮影条件が固定された、昼間や午後、日没など複数の異なる時間帯で撮影した映像が用意されているものとする。複数の異なる時間帯で撮影された事前撮影映像と夜間映像で対応点探索を行うことで、対応点探索の精度を向上させ、 A^* を求めるのに十分な数の対応点組を得る。

6. 実験

6.1. 複数の異なる時間帯に撮影された事前撮影映像と夜間映像の対応点探索

5 章で述べた日照変動による特徴量の変化が大きいと考えられる昼間、午後、日没の 3 つの時間帯における事前撮影映像と夜間の映像との対応点探索精度の検証を行う。

昼間の映像と同じ条件で撮影した午後の映像及び日没の映像を図 8 に示す。



図 8: 昼間, 午後, 日没時の映像

表 1 で示すように, 昼間の映像は RANSAC によって誤対応を除去した後, 5 個の対応点組を得ているが, いずれも正しい対応点組ではない. 午後の映像は誤対応を除去した 9 個の内, 正しい対応点組が 2 個存在した. 日没の映像は他の 2 つの映像に比べ, 正しい対応の対応点組が大幅に増加し, 全て正しい対応点組だった. 正しい対応点組が増加した原因は, 街灯などの照明が点灯したことにより夜間映像の撮影時と照明環境が近くなり, 特徴点と同じ傾向の特徴量が現れた為だと考えられる.

表 1: 撮影時間帯と特徴点对応数との関係

	daytime	afternoon	evening
Detected keypoint	459	276	316
Matching after RANSAC	5	9	24
Correct Matching	0	2	24

6.2. 夜間景観の複合現実型提示の実験

本手法により, 撮影地点の異なる 3 つのシーンで複合現実型の夜間映像を合成し, 幾何的な位置合わせによる合成が行えていることを検証する.

6.2.1. 実験環境

事前撮影映像と夜間映像の撮影に Acer ICONIA TAB A100 を用いる. 事前撮影映像は方位ズレが 5° 生じた状況を想定し, 3 次元測量機(Nikon GF-EX11)に A100 を固定し, 夜間映像の撮影条件から方位方向に 5° 右回りに回転するように測量した状態で撮影を行った. 映像の解像度は $640[\text{pixel}] \times 480[\text{pixel}]$ である.

6.2.2 実験方法

図 9 で示すように 3 つのシーンで事前撮影映像と夜間映像を用意した. 事前撮影映像は昼間, 午後, 日没の時間帯で撮影を行った. 各行左から昼, 午後, 日没, 夜間(入力)映像である. (a) シーン 1 は夜近景がカメラから比較的遠方に存在するシーン, (b) シーン 2 は夜近景がカメラから比較的近傍に存在するシーン, (c) シーン 3 は夜近景の他に, 遠くの景観にも明るい部分が含まれるシーンである. 各時間帯の事前撮影映像と

夜間映像で対応点探索を行い, 最も正しい対応の多い時間帯の事前撮影映像と夜間映像から求めた 2 次元変換行列 A^* を用いて, 複合現実型の夜間映像を合成する.

6.2.3. 実験結果と考察

シーン 1 から 3 において夜間映像と事前撮影映像から複合現実型の夜間映像を合成した結果を図 10 に示す. (a) は図 9 (a) 右の夜間映像と同じ撮影条件で予め撮影された昼間映像, (d) は結果である(以下同様).

シーン 1 からシーン 3 はいずれも夜近景と事前撮影映像の遠方の景観が幾何的に位置合わせをして合成されていることが分かる.

本提案手法では夜近景と事前撮影映像の遠方の景観を合成する際, 遠方の景観を平面の平行移動で近似することで幾何的な位置合わせを実現している. そのため, 2 次元変換行列を推定するのに用いる対応点組がカメラから近い位置に集中していると, 夜近景と昼近景で正しい対応が取れていたとしても, 幾何的な位置合わせを行うことが難しくなる.

そのため, 本提案手法を用いて幾何的な位置合わせによる複合現実型の夜間映像の合成を行うには, 夜近景にカメラから比較的遠方にある対応点組が得られるようなシーンで夜間映像を撮影する必要がある.

7. おわりに

本研究では夜近景と事前撮影映像の遠方の景観を幾何的な位置合わせによる複合現実型の夜間映像の合成手法を提案した.

複合現実型の夜間映像を合成する際, 夜間映像に対する事前撮影映像の撮影位置と方位のズレが位置合わせに大きな影響を及ぼすが, 自然特徴点による対応点探索を用いて夜近景と昼近景の対応関係を求め, 幾何的な位置合わせによる合成を実現した.

さらに, 事前撮影映像が複数の異なる時間帯で撮影されているものとして夜間映像と対応点探索を行うことで, 対応点探索の精度の向上ができることを示した.

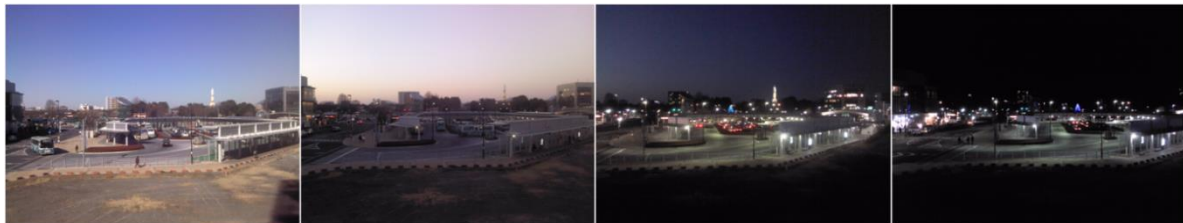
今後の課題として, 夜間映像中の特徴点の中でカメラの近傍にある特徴点と遠方にある特徴点を分類し, より精度の高い位置合わせをした複合現実型の夜間映像を合成することなどが挙げられる.

文 献

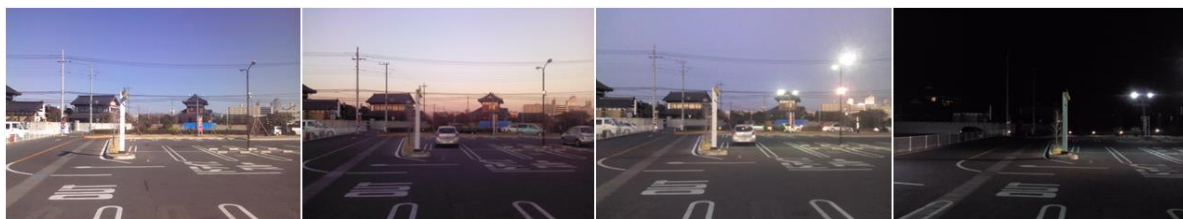
- [1] R. Raskar, A. Ilie, J. Yu, "Image fusion for context enhancement and video surrealism," Proc. NPAR, pp. 85-95, 2004.
- [2] N. Yazawa, H. Uchiyama, H. Saito, M. Servieres, and G. Moreau, "Image based view localization system retrieving from a panorama database by surf," Proc.

of the IAPR Conference on Machine Vision Applications, vol. 1, pp. 118-121, 2009.

- [3] 佐藤准嗣, 高橋友和, 井手一郎, 村瀬洋, “GPS 座標付き全方位映像群からの市街地映像マップの構築と街並変化の検出,” 信学論(D-II), vol. J90-D, no. 4, pp. 1085-1095, 2007.
- [4] 横地裕次, 池田聖, 佐藤智和, 横矢直和, “特徴点追跡と GPS 測位に基づくカメラ外部パラメータの推定,” 情報処理学会論文誌: CVIM, vol. 47, no. 5, pp. 69-79, 2003.
- [5] 天目隆平, 平岡貴志, 海川真佑, 柴田史久, 木村朝子, 田村秀行, “環境の三次元モデルと魚眼カメラを利用した屋外複合現実感システムのためのユーザ位置推定,” MIRU, pp. 1011-1016, 2007.
- [6] Y. Cai, K. Huang, T. Tan, Y. Wang, “Context enhancement of nighttime surveillance by image fusion,” Proc. of International Conference on Production Research, pp. 980-983, 2006.
- [7] E. Rosten, T. Drummond, “Machine learning for high-speed corner detection,” Proc. of the 9th European Conference on Computer Vision, vol. 1, pp. 430-443, 2006.
- [8] H. Bay, T. Tuytelaars, L.V. Gool, “SURF: Speeded Up Robust Features,” Proceedings of the 9th European Conference on Computer Vision, Springer LNCS, vol. 3951, part 1, pp. 404-417, 2006.
- [9] M. Fischler, R. Bolles, “Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography,” Communications of the ACM, vol. 24, no. 6, pp. 381-395, 1981.



(a) Scene 1



(b) Scene 2



(c) Scene 3

図 9: 使用した昼間, 午後, 夕方, 夜間の映像



(a)



(b)



(c)



(d)



(e)



(f)

図 10: 夜間における複合現実感型提示の結果 (下段)