

見回し行動時の個人視点映像を対象とした広視野貼り合わせのための画像選択*

近藤 一晃^{†a)} 松井 研太[†] 小泉 敬寛^{†b)} 中村 裕一^{†c)}

Image Selection for Mosaicing First-Person-View Videos on Looking-Around Behaviors*

Kazuaki KONDO^{†a)}, Kenta MATSUI[†], Takahiro KOIZUMI^{†b)},
and Yuichi NAKAMURA^{†c)}

あらまし 個人視点映像から貼り合わせに適した画像を自動的に選び出して仮想的に広視野・高解像度の画像を合成する方法について述べる。個人視点映像は貼り合わせに不適な画像を多く含むため入力画像全てを合成すると却って質の悪い画像となってしまう。提案手法では貼り合わせの良さを定量化する評価関数を用いて画像を選択することで、できるだけ質の良い貼り合わせ画像を合成する。評価関数は貼り合わせによって画素値の確からしさが増大することに基づいて設計した。加えて評価値を計算する画像の組の数を減らすことで効率的に画像選択を行う手法を合わせて提案する。実際の個人視点映像を入力として実験を行い、提案手法の有効性及び特性を確認した。

キーワード イメージモザイクング, 個人視点映像, 貼り合わせの良さ

1. ま え が き

複数枚の画像を貼り合わせることで仮想的に広視野・高解像度の画像を生成する手法はイメージモザイクング若しくはイメージスティッチングと呼ばれている。この技術は、衛星写真や航空写真を貼り合わせて広大な地図を作成する [5], [6], 小型の手持ちカメラで文書冊子を画像スキャンする [3], [4], 全周撮影画像からパノラマ画像を生成する [1], [2], など様々な用途に用いられている。本研究ではそれらとは異なる用途として個人視点映像を構成する画像列の貼り合わせを扱う。個人視点映像とは人物の頭部や胸部に小型カメラを装着することで撮影される一人称視点の映像のことで、主に映像メディアとしてのライフログ [7], [8] や体験活動の記録 [11]~[13] として幅広く用いられている。

ただし、視野が狭い・画質が低い・映像揺れを多く含む・長時間記録になりがちである、といった理由から閲覧する際の負担が大きい。我々は個人視点映像中の画像を貼り合わせてこの問題を緩和することを目指している。すなわち、見回すような区間の個人視点映像を貼り合わせることで静止画ではあるものの広視野・高解像度の画像が得られるので、それを利用すれば広視野の揺れ補正映像や映像全体を一覧するための要約的な画像を作成できると考える。

しかし、個人視点映像を構成する全ての画像、あるいは無作為に選んだ画像列から必ずしも良い貼り合わせ画像が合成できるとは限らず、むしろ所望の質が得られることの方が少ない。なぜなら個人視点映像は人物の移動・頭部運動・環境変化に起因した「貼り合わせに不適な画像」を含むことが多いからである。つまり良い貼り合わせ画像を得るためにはそれに適した画像の組を選んで合成する必要がある (図 1)。長時間の個人視点映像から逐一人手で選択することは実用的でないため、本論文ではこの作業を自動化することを目標に、貼り合わせの良さを定量化するような評価関数 [9] を提案する。これにより貼り合わせの観点から点数付けを行うことができるので「どの組み合わせの方がよ

[†] 京都大学学術情報メディアセンター, 京都市
Academic Center for Computing and Media Studies, Kyoto University, Yoshida Honmachi, Sakyo-ku, Kyoto-shi, 606-8501 Japan

a) E-mail: kondo@ccm.media.kyoto-u.ac.jp

b) E-mail: koizumi@ccm.media.kyoto-u.ac.jp

c) E-mail: yuichi@ccm.media.kyoto-u.ac.jp

* 本論文はマルチメディア・仮想環境基礎研究専門委員会推薦論文である。

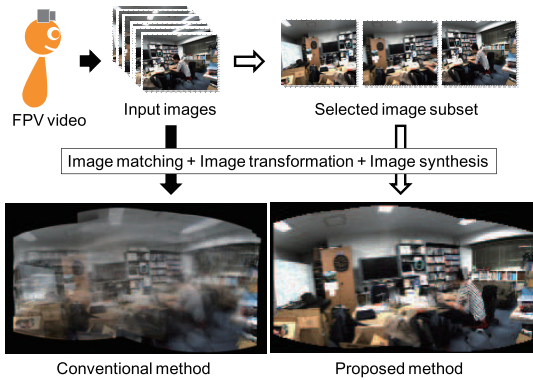


図1 質の良い貼り合わせ画像を合成するための入力画像選択

Fig. 1 Image selection for better stitching.

り質の良い貼り合わせ結果となるか」や「これ以上貼り合わせる画像を増やすと却って悪い結果となる」といった判断が可能となる。

貼り合わせに最も適した画像の組を選ぶことは組み合わせ最適化問題である。単純に考えれば全ての画像の組について貼り合わせ評価値を計算してその最大値を取り出せばよい。しかし、原理的には入力画像列のそれぞれについて貼り合わせに用いるかどうかの2通りが考えられるため容易に組み合わせ爆発が発生してしまう。貼り合わせ評価値は画像の組み合わせに対して非単調であり、例えば二枚の貼り合わせでは良い評価値だが三枚目を追加すると悪くなってしまう場合やその逆の場合も十分にありうるので、一般的な組み合わせ最適化手法をそのまま用いることは難しい。本論文では局所的な少数画像の貼り合わせ評価値に基づいて探索範囲を近似的に削減する手法[10]を併せて提案する。

2. 関連研究

画像貼り合わせに関する従来手法の多くは、画像間に共通する領域があるかどうかの判定・画像の変形と位置合わせ・共通領域における画素の合成、といった問題を、(1)射影幾何に基づいた変形モデルで記述できること、(2)入力画像の質が良いこと、という仮定の下に精度良く解いている。一般によく用いられているホモグラフィ変換は、入力画像撮影時のカメラ運動が回転のみに近似できる、若しくは撮影対象シーンが平面とみなせることを前提とした変形モデルである[1],[2]。文書冊子を対象とした貼り合わせには、ホモグラフィ変換に紙面の曲率等のパラメータを加えた変形モデル

が用いられている[3],[4]。しかし、どのような変形モデルを用いたとしても、任意の三次元空間を一枚の二次元画像で矛盾なく表現することには限界がある。テレビ番組映像やハンディカム・スマートフォン映像などの一般の動画は「興味対象を記録に残す」「後で閲覧する」ことを意識して撮影されている。そのため自ずと良い見回しカメラワークとなる傾向にあり、そのときの画像位置合わせがホモグラフィ変換から大きく逸脱することは考えにくい。一方で個人視点映像は、撮影することを意識していない人間が装着カメラにより自動的かつ無自覚に記録した映像である。そのためたとえ見回し行動中であっても頭部移動が頻繁に発生するので運動視差に起因した不整合が多く発生する。加えて画質の良い個人視点映像がいつも得られるとも限らない。カメラ装着者が素早く見回したり速く動いたりすると容易にモーションブラーが現れる。室内外を移動した場合には急激な照明変化に対応しきれずに暗すぎたり明るすぎたりした映像になってしまう。

このように、個人視点映像を貼り合わせ対象とするとこれまで満たされていた前提が成り立たなくなることが多い。提案手法はそれが満たされている度合いを定量化することで、できるだけ良い貼り合わせ画像を生成するものである。実際このような問題はこれまでほとんど扱われてこなかった。入力画像群を互いに貼り合わせられる画像群にグループ分けする手法[2]は提案されているが、結局全ての画像を貼り合わせに用いることに変わりはない。またグループ分けには画像間の共通領域の有無を用いており貼り合わせの良さを扱ってはいない。

映像を貼り合わせ対象とした場合、どの区間を一枚に貼り合わせるのかという時分割問題も重要である。映像時分割の手法にはコンテンツがもつ意味に注目したものが多く、編集された映像番組においてはショットやカットを検出することが内容の切り替えに対応した時分割をもたらす[14]~[16]。個人視点映像を対象とした時分割手法には、装着者の行動を認識するもの[18]、シーンの変化に基づいたもの[19]、人物の検出を用いたもの[17]などが提案されている。本論文では主にカメラ装着者の行動に基づいた手法により見回し行動区間が切り出されていることを前提とし、その区間を貼り合わせ対象としている。

3. 貼り合わせ評価関数の設計

3.1 貼り合わせの良さを定量化するキーアイデア

画像の組に対して貼り合わせの良さを与える評価関数を設計する。理想的な貼り合わせを考えると、その効果は (1) 複数枚の画像が重なり合った共通領域において撮影過程で生じるノイズや揺らぎが抑制される、(2) 複数枚の画像を繋ぎ合わせることで仮想的に視野が拡大される、の大きく二つである。本提案の目的はそれらの度合いを定量化することである。従来の指標、例えば、輝度差の絶対値の総和や正規化相互相関などは (1) を評価する方法として利用可能であるが、二枚の画像の共通領域のみしか計算できず、加えて別途 (2) も評価しなければならない。複数枚の共通領域における画素値の整合性と貼り合わせることによる視野の拡大の両方を同時に評価する仕組みが求められる。

我々は貼り合わせによるどちらのメリットも「観測値から本来の画素値をどの程度確からしく推定できるか」言い換えれば「推定画素値の確からしさ」が増大することに対応するのではないかと考えた。共通領域では同一の対象に対して複数の画素値が観測されるため本来の画素値をより確からしく推定できる。一方、視野が広がることは、未観測であった領域に対して情報が与えられて確からしい状態に変化したとみなせる。そこで提案手法では入力画像から理想的な貼り合わせ画像を確率的に推定することを考え、その過程における推定輝度値の確からしさを用いて貼り合わせ評価値とする。カメラによる撮影を光学系と光電変換を通した確率過程とみなせば、その逆過程を考えることにより入力画像群から理想的な画素値の確率分布を導出することができる (図 2)。このときの推定画素値が確からしいことは分布が特定の値に偏っていることに対応するため、分布の形を評価することで確からしさ、すなわち貼り合わせの良さを定量化する (図 3)。

3.2 画素値の確からしさに基づいた評価関数

理想的な貼り合わせ画像及びその観測画像列を \mathbf{x} , $\{\mathbf{y}_t\}$ のように画像の輝度値を一次元に並べたベクトルで表記する。提案手法では位置合わせ・変形済みの $\{\mathbf{y}_t\}$ を入力画像列とみなし、 $\{\mathbf{y}_t\}$ から \mathbf{x} を推定したときの画素値の確からしさを基に貼り合わせ評価値 $Q(\{\mathbf{y}_t\})$ を導出する。位置合わせ・変形の過程で探索量削減 (4.) に必要な情報を得るために、 $\{\mathbf{y}_t\}$ の作成には局所特徴量の対応関係に基づいた手法 [2] を用いる。

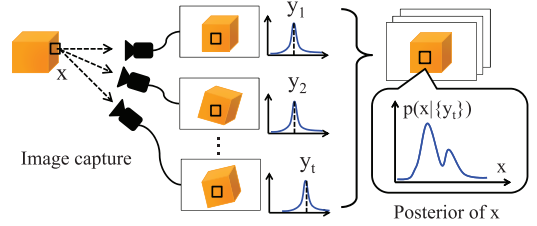


図 2 観測モデルに基づいた画素値の事後分布推定
Fig. 2 Scheme of estimating the posterior based on the image capture model.

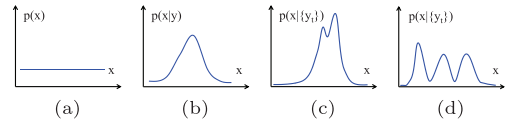


図 3 事後分布の例。(a) 無観測状態。(b) 一枚の画像でのみ観測された場合。(c) 整合性の高い複数画素値が観測された場合。(d) 整合性の低い複数画素が観測された場合。

Fig. 3 Examples of the posteriors.

ベイズの定理より事後確率 $p(\mathbf{x}|\{\mathbf{y}_t\})$ は、 \mathbf{x} が与えられたときの $\{\mathbf{y}_t\}$ の確率分布すなわちゆう度 $p(\{\mathbf{y}_t\}|\mathbf{x})$ を用いて、

$$p(\mathbf{x}|\{\mathbf{y}_t\}) = \frac{p(\{\mathbf{y}_t\}|\mathbf{x})}{p(\{\mathbf{y}_t\})} p(\mathbf{x}) \quad (1)$$

と表され、複数枚の入力画像 $\{\mathbf{y}_t\}$ が互いに独立である仮定の下で

$$p(\mathbf{x}|\{\mathbf{y}_t\}) = \frac{\prod_{t=1}^T p(\mathbf{y}_t|\mathbf{x})}{\prod_{t=1}^T p(\mathbf{y}_t)} p(\mathbf{x}) \quad (2)$$

のように、任意の入力画像枚数 T を扱うことができる一般形に変形される。これは複数枚の低解像度画像から一枚の高解像度画像を生成する超解像の手法 [20], [21] にも見られる導出法である。超解像では入力画像全ての共通領域のみを扱っているため T は位置によらず常に入力画像の枚数に等しいが、画像の貼り合わせでは重なり合う画像の枚数やその組み合わせは共通領域ごとに様々である (図 4)。言い換えれば、領域全体を一括で扱うような式 (2) では画像の貼り合わせを表現することは難しい。そこで同一画像上の画素同士も互いに独立であるという仮定を新たに設けることで式 (2) を画素単位 i で考えた

$$p(x(i)|\{y_t(i)\}) = p(x(i)) \prod_{t=1}^T r(x(i), y_t(i)) \quad (3)$$

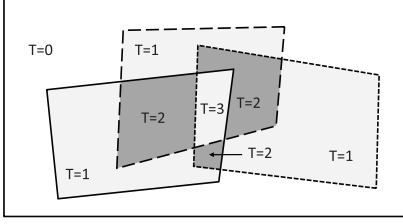


図4 多数枚画像の貼り合わせ関係の例 (図中では3枚). $T = 0, 1, \geq 2$ がそれぞれ無観測の領域, 一枚の画像でのみ観測された領域, 複数画像の共通領域に対応する. 貼り合わせ領域全体 S_A はこれら複数画像の和領域として定義される

Fig. 4 Example of the geometric relation on multiple image stitching.

$$r(x(i), y_t(i)) = \begin{cases} 1.0 & y_t(i) = \phi \\ \frac{p(y_t(i)|x(i))}{\int p(y_t(i)|x(i))p(x(i))dx} & (4) \\ \text{otherwise} & \end{cases}$$

を得る. ここで $y_t(i) = \phi$ は「位置 i は画像 t の視野外である」ことを示す. 式 (3) は式 (2) と同様に任意の観測数を扱うことができる. すなわち $T = 0$ のとき無観測の領域, $T = 1$ のとき一枚の入力画像でのみ観測された領域, $T \geq 2$ のとき複数入力画像の共通領域での事後確率導出式となる. 画像が与えられていなければ本来の画素値が全くわからないので事前確率は一様分布 $p(x) = \frac{1}{X}$ で与える. X は画素値のとりうる範囲を示す. ゆう度 $p(\{y_t\}|\mathbf{x})$ は「位置ずれ・ぼけがない場合に画素値はガウス分布に従った揺らぎをもち, それらがある場合にはどの値も等しくとりうる」という観測過程モデルに基づいて

$$p(y_t|x) = \alpha G(x - y_t, \sigma_1) + (1 - \alpha) \frac{1}{X} \quad (5)$$

と設計した. $G(\cdot)$ は揺らぎを表すガウス分布, σ_1 はその分散, $0 \leq \alpha \leq 1$ は位置ずれやぼけがない確率を各々示す.

こうして得られた事後分布がどれだけ特定の値付近に集中しているかを定量化して位置 i における貼り合わせ評価値とする. 提案手法では最ゆう推定で求めた偏りの中心 x_c に位置する参照ガウス分布との類似度を用いた. この手続きは

$$q(\{y_t\}) = \int_X p(x|\{y_t\})G(x - x_c, \sigma_2)dx \quad (6)$$

$$x_c = \operatorname{argmax} p(x|\{y_t\})$$

のように表される. σ_2 は参照ガウス分布の分散を示

す. 最終的な貼り合わせ評価値 Q は式 (6) で導出された画素ごとの評価値を貼り合わせ領域全体 S_A で統合することで得られる. 提案手法における貼り合わせ評価値は確からしさをなわち確率である. そこで領域全体での貼り合わせ評価値は各画素における貼り合わせ評価値の積で表すことにした. ただし確率を画素数分かけ算するととても小さな値になってしまい計算機で扱うことが困難であるため対数をとることで現実的な値として算出する. 積の対数は対数の和で表すことができるので

$$Q(\{y_t\}) = \frac{1}{|S_A|} \int_{S_A} \log q(\{y_t(i)\})di \quad (7)$$

のように定式化される. ただし貼り合わせ領域全体で積算すると膨大な値になりうるため面積で正規化して実用的な値としている.

3.3 評価関数のパラメータ設定

貼り合わせ評価値を制御するパラメータ $\alpha, \sigma_1, \sigma_2$ が満たすべき条件について考えよう. α は重み係数, σ_1, σ_2 は分散を示すパラメータなので基本的には $0 < \alpha < 1$, $0 < \sigma_1, \sigma_2$ である. 一方, 貼り合わせることで視野が広がる効果を評価するには, 一枚の画像だけで観測された領域の貼り合わせの良さ q_1 が無観測の領域の貼り合わせの良さ q_0 よりも大きい必要がある. q_0 は常に一定で $q_0 = \frac{1}{X}$ である. また q_1 は画素値によらず $\alpha, \sigma_1, \sigma_2$ のみから決定される. これは

$$q_1 = \int p(x|y)G(x - x_c, \sigma_2)dx$$

$$= \int \left\{ \alpha G(x - y, \sigma_1) + (1 - \alpha) \frac{1}{X} \right\} \times G(x - y, \sigma_2)dx$$

$$= \alpha \int G(x, \sigma_1)G(x, \sigma_2)dx + (1 - \alpha) \frac{1}{X}$$

$$= \alpha H(\sigma_1, \sigma_2) + (1 - \alpha) \frac{1}{X}. \quad (8)$$

なる変形から導き出される. これより条件 $q_0 < q_1$ は,

$$q_0 < q_1 \iff \frac{1}{X} < \alpha H(\sigma_1, \sigma_2) + (1 - \alpha) \frac{1}{X}$$

$$\iff 0 < \alpha \left(H(\sigma_1, \sigma_2) - \frac{1}{X} \right)$$

$$\iff H(\sigma_1, \sigma_2) > \frac{1}{X} \quad (9)$$

と表される. すなわち σ_1, σ_2 に対して $H(\sigma_1, \sigma_2) > \frac{1}{X}$ なる条件が与えられる. 二枚以上が観測された領域で

はそれらを貼り合わせない方が良いこともあるので評価値の大小に係る条件は与えられない。すなわち $q_t, t \geq 2$ を含む場合について上記のような条件を考慮する必要はない。

これらのパラメータは観測過程を特徴づける値であるが、どのような貼り合わせを良いとみなすかを定める値でもある。つまりパラメータ設定により、貼り合わせによって広がる効果と共通領域における整合性のどちらを重く評価するかを制御できる。 σ_1 は事後確率の分布形状に影響を与えるパラメータである。この値をあまり大きく設定しすぎると、ほぼ同じ値の複数画素が観測されたときと明らかに異なる値の複数画素が観測されたときとで、事後分布にほとんど違いが現れなくなってしまう。そのため本論文では画素値の取り得る範囲 $X = 256$ に対して $\sigma_1 = 5$ を固定値として採用している。一方 σ_2 は分布の偏りの評価方法に関わるパラメータであるためユーザの目的に応じて設定できる。この値を大きくするほど 3.2 で述べた参照ガウス分布が幅広くなるため、偏りの少ない事後分布すなわち整合性の低い複数画像共通領域にもある程度高い評価値が与えられる。結果として多少整合性が低くても貼り合わせ領域が広がるような画像の組が相対的に高い評価値を得る。逆に σ_2 を小さくすると視野の広がりよりも整合性を重視した貼り合わせが高い評価値を得ることになる。ただし σ_2 をあまりに大きく設定すると整合性の低い方がむしろ高い評価値を得てしまうため注意が必要である。この限度について詳細な検討は行っていないが、画素値の取り得る範囲 $X = 256$ に対しておおよそ $\sigma_2 = 30$ 程度まではほぼ所望の挙動となることが確認できている。

4. 画像選択における探索量の削減

個人視点映像を連番画像に分解した入力画像列 $\{\mathbf{I}\}$ から最も貼り合わせ評価値が高い画像の組 $\{\mathbf{I}_p\}$ を取り出して合成することで広視野画像を得る。この処理は

$$\{\mathbf{I}_p\} = \operatorname{argmax} Q(\{\mathbf{I}_s\}), \forall \{\mathbf{I}_s\} \subset \{\mathbf{I}\} \quad (10)$$

と表すことができる。ここで $\{\mathbf{I}_s\}$ は入力画像列 $\{\mathbf{I}\}$ の部分集合を示しており、貼り合わせ評価値の最大値を探索する上での媒介変数である。 $Q(\{\mathbf{I}_s\})$ の導出には式 (7) を用いる。ただし評価値同士の比較を行うために $\{\mathbf{I}\}$ の貼り合わせ領域全体を S_A として共通に用いる。この手続きは考える画像の組全てについて貼

り合わせ評価値を求めるのでその計算量は膨大となる。本章では近似的に探索量を削減するような方法について述べる。

4.1 代表画像による類似画像列の置き換え

個人視点映像の中には、シーンやカメラワークがよく変化する部分があれば内容がほとんど変化しない場面もある。後者のような画像列はその中のどれを用いても同等の貼り合わせ結果となると予想されるため、互いに類似した画像列は一枚の代表画像で置き換えることにする。これにより画像列全てを入力として用いた場合と同等の結果を少ない枚数の入力画像で導くことができると考えた。提案手法では以下の条件を満たす区間を類似画像列とみなす。

- 任意の二枚の画像が十分な共通領域をもつ
- 共通領域の画素値が十分に整合する

まず入力画像列全体を対象に画像のペアを列挙・位置合わせし、その共通領域の面積と画素値の整合性を計算する。このときの位置合わせ処理は 3.2 で述べた $\{\mathbf{y}_t\}$ の作成と共通化できる。時間的に遠い画像ペアには、共通領域がない場合が多い・共通領域の有無を誤推定しやすい・計算量が増える、といった性質があるため、時刻差がしきい値 F_{th} フレーム以下のペアだけを対象として計算する。誤推定された位置合わせの除去には、位置合わせを行う際に用いた局所特徴量の対応点数に基づいた指標 [2] とともに下記に示す画像変形後の幾何関係に基づいた条件を併用している。具体的には二枚の画像間の位置合わせが以下の条件を満たさない場合には後述の r_{ab}, w_{ab} の値を 0 とすることで対象の二枚が類似画像列として扱われないようにする。変形後の面積比 近い時刻で撮影された二枚の画像であれば位置合わせ・変形後の大きさが極端に異なることはないという仮定に基づいて、面積比に関する条件

$$1 - s_{min} < \left| \frac{S(I_a)}{S(I_b)} \right| < 1 + s_{max} \quad (11)$$

を設ける。 I_a, I_b は変形後の 2 枚の画像、 $S(\cdot)$ は面積を出力するオペレータ、 s_{min}, s_{max} は許容する面積変化のしきい値である。

4 端点の位置関係 画像変形前後で 4 端点の上下関係が入れ替わるような変形は誤った対応点に基づいた誤推定とみなす。

以上の条件を満たした画像ペア (I_a, I_b) に対して、共通領域の面積 r_{ab} を計算する。これは二枚の画像の大きさを正規化した値

$$r_{ab} = \frac{2S(I_a \cap I_b)}{S(I_a) + S(I_b)} \quad (12)$$

で定義する．共通領域における画素値の整合性 w_{ij} は画素値の差から定義される指標

$$w_{ij} = \frac{1}{S(I_a \cap I_b)} \sum_i^{I_a \cap I_b} \frac{X - |I_a(i) - I_b(i)|}{X} \quad (13)$$

を用いた．

なお式 (13) の代わりに共通領域に対して提案手法を用いることもできる．その場合，画素値の確かさに基づいた評価値が得られ，しきい値 r_{th}, w_{th} もそれに合わせた値に設定する必要がある．どちらの方法を用いても本質的な違いはほとんどないため，本論文ではより計算量が少ない式 (13) を採用した． $\mathbf{R} = \{r_{ab}\}, \mathbf{W} = \{w_{ab}\}$ は二枚の画像の関係を行列の形で表現しているのので，対角成分 (i, i) と (j, j) を頂点とする $(j - i + 1) \times (j - i + 1)$ の正方部分行列 $\mathbf{R}_i^j = \{r_{ab}\}, \mathbf{W}_i^j = \{w_{ab}\}, i \leq a, b \leq j$ を取り出せばそれは \mathbf{I}_i から \mathbf{I}_j までの連続する $j - i + 1$ 枚の画像間の共通領域情報を取り出すことに対応する．提案手法ではこの正方部分行列に含まれる要素が全てしきい値 r_{th}, w_{th} 以上であることを節頭で述べた類似画像列の条件として定義し，該当した正方部分行列に対応する画像列を類似画像列として抽出する．図 5 はそうして取り出された類似画像列に対応する正方部分行列の例を示している．類似画像列から代表画像を選ぶ方法は様々に考えられるが本論文では単純に区間の時間中心に位置する画像とした．

この方法によってどの程度枚数が削減されるかは入力映像の内容変化に依存する．カメラがほとんど動いておらずシーンも静的であるような区間は代表画像が少数となるため削減率が高く，カメラが動いたりシーンが動的である場合にはあまり削減されない

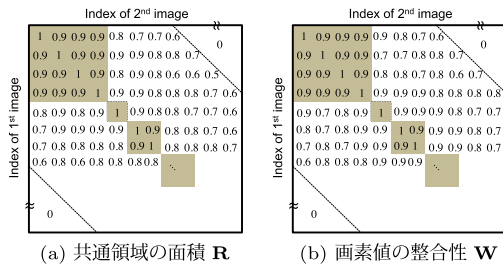


図 5 しきい値パラメータに $r_{th} = 0.9, w_{th} = 0.9$ を用いた場合の類似画像列の抽出結果例 (図中網掛部)

Fig. 5 Detecting similar image sequences.

い．加えて本手法はどの程度まで類似した画像を 1 枚とみなすかという近似であり，その度合いはしきい値 r_{th}, w_{th} によって変化することに注意したい．

4.2 不整合な三つ組画像の同時不使用

提案手法は領域全体を対象とした貼り合わせの良さを導出するので，部分的には不整合な画像の組に対してもよい評価値を与える場合がある．この問題を解決するために局所的にみたときの貼り合わせ評価値が低い場合はたとえ領域全体についての評価値が高くても棄却するような条件を加えよう．ここで「貼り合わせの評価値が低い領域に新たに画像を加えてもやはり評価値は低い」という仮定をおくと，少数枚画像の貼り合わせ評価値に基づいて領域全体に対する計算を省くことができる．この方法は原理的に言えば任意の画像枚数の共通領域について適用できるが，あまりに少ない枚数では上述の前提が成り立たないことが多い．一方で多数枚の共通領域を対象とすれば，そのような組の数は少ないので探索量の削減効果が低くなってしまふ．そこで 3 枚の画像の組に対して上記手法を適用する．本論文では省略するが 2 枚，4 枚，5 枚…といった枚数の画像の不整合を利用することも同様に考えることができる．

まず共通領域をもつ 3 枚の画像の組み合わせを前節で導出されている \mathbf{r}_{ab} を活用して効率的に列挙する． \mathbf{r}_{ab} の各要素は 2 枚の画像が共通領域をもっているかどうかを示しているのので，その情報から任意の 2 枚が互いに共通領域をもつような三つ組画像 $\mathbf{I}_{klm} = (I_k, I_l, I_m)$, s.t. $r_{kl} > 0, r_{lm} > 0, r_{mk} > 0$ を列挙する．ただし，この条件を満たす \mathbf{I}_{klm} でも 3 枚全てに関する共通領域をもつとは限らない．またその共通領域が狭い場合は影響が少ないので除外したい．そこでしきい値以下の共通領域しかもたないような組は破棄する．こうして列挙された三つ組画像について共通領域 S_{ov} における貼り合わせ評価値 $Q_3(\mathbf{I}_{klm})$ を，

$$Q_3(\mathbf{I}_{klm}) = \frac{\sum_i^{S_{ov}} q(\{I_k(i), I_l(i), I_m(i)\})}{|S_{ov}|} \quad (14)$$

のように計算する．式中 q は三枚の画像の共通領域に属する画素を対象とした貼り合わせ評価値である．すなわち $T = 3$ とした式 (3), (4) を用いて導出される． $|S_{ov}|$ は共通領域の面積を示す． $Q_3(\mathbf{I}_{klm})$ がしきい値 Q_{th} 以下であったとき， \mathbf{I}_{klm} を同時には使用してはいけない三つ組画像として登録し，貼り合わせ評価値を計算する画像の組に対するフィルタとして用いる．す

なわち対象の三枚の画像を含むような画像の組み合わせは評価値の計算・比較から除外することで効率的な探索を行う。

代表画像による類似画像列の置き換えと同様にこの手法も近似的な探索量の削減であることに注意したい。仮定が満たされない、すなわち、追加の画像によっては貼り合わせ評価値が増加する場合もある。探索量削減と近似性能はトレードオフの関係にあり、そのバランスはしきい値 Q_{th} によって制御される。

5. 実験

5.1 貼り合わせ評価値に基づいた広視野画像生成

提案手法を用いて画像選択を行った場合にどのような広視野画像が合成されるのかを実際の個人視点映像を入力として確認する。そのために入力画像全てを貼り合わせる従来手法と選択された画像のみを貼り合わせる提案手法とで最終的に出力される広視野貼り合わせ画像の質がどのように異なるのかを比較・評価した。データセット 屋内・屋外における日常行動の個人視点映像を頭部に装着した小型カメラにより撮影した。対象のシーンは研究室内の個人スペース・共有スペース・実験室、及び建物周囲や駐輪場などである。それらを十数秒から数十秒程度の短時間映像に時分割したものを提案手法の入力画像とした。図6に使用した画像の例を示す。カメラの解像度及びフレームレートは各々 640×480 pixel, 15fps とした。映像は 8bit カラー画像で撮影したため輝度値のとりうる範囲は $X = 256$ である。

パラメータ設定 貼り合わせの評価関数には $\alpha = 0.5$, $\sigma_1 = 5$, $\sigma_2 = 5$, ホモグラフィ行列の誤推定を判断するパラメータには $s_{min} = 0.2$, $s_{max} = 0.1$ を用いた。一回の貼り合わせ対象とする区間長の上限は

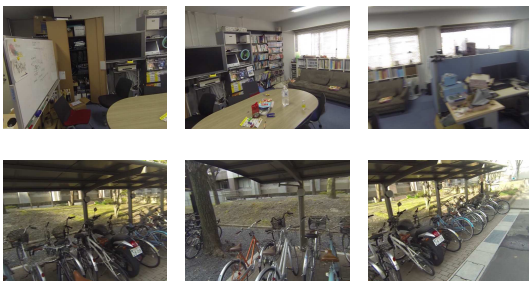


図6 実験に用いた入力個人視点映像の例。(上段) 屋内シーン。(下段) 屋外シーン

Fig. 6 Examples of input FPV images.

$F_{th} = 100$ frame である。また類似画像列を検出する際のしきい値は $r_{th} = 0.85$, $w_{th} = 0.80$ とした。

実験結果と考察 合成結果の例を図7に示す。提案手法を用いて画像の選択を行った上で合成した方が入力画像全てを用いるよりも画質の良い広視野画像が得られている。特に図7のPeriod-A, B, F, Gにおいてその効果が顕著に見られ、従来手法では全体的なコントラストが大きく低下してしまっている。これは位置合わせ誤差を含む多くの画像を貼り合わせたため結果的に平均的な輝度値に近づいてしまったからだと考えられる。図7のPeriod-C, D, Eではそれほど大きな差は見られないが、各々、モニタ領域・左や正面の戸棚領域・左端のホワイトボード領域でやはり同様の傾向が現れている。提案手法はできるだけコントラストを維持したままで広視野を構成するように画像選択することが確認された。

ただし、幾つかの入力映像では移動に起因する運動視差の影響を十分に取り除くことはできなかった。提案手法は入力画像の中で最適な画像の組を選択するものなので、そもそも貼り合わせに不適な画像ばかりが与えられたときに所望の質を満たすのは難しいことに注意したい。

5.2 パラメータ設定の違いによる影響

評価関数に用いるパラメータを異なる値に設定した場合、広視野貼り合わせ画像の結果にどのような影響を与えるかを比較・検証する。





パラメータ設定 貼り合わせの良さの評価関数に対して $PS_a = \{\alpha = 0.5, \sigma_1 = 5.0, \sigma_2 = 5.0\}$ と $PS_b = \{\alpha = 0.5, \sigma_1 = 5.0, \sigma_2 = 30.0\}$ の二種のパラメータセットを用いた。 PS_a に比べて PS_b は σ_2 の値が大きいため画像共通領域における画素値の整合性よりも貼り合わせることによって視野が広がることを優先するような設定である。入力に用いたデータセット及びその他のパラメータ設定は前節と同一とした。

実験結果と考察 合成結果の比較例を図8に示す。全体的に見てパラメータセット PS_b の方が多くの画像を選択することでより広い貼り合わせ画像を生成している。それに比べてパラメータセット PS_a では視野の広がりこそ少ないものの貼り合わせの共通領域があまり矛盾しない画像の組が選ばれている。特に図8のPeriod-Gではその効果が顕著に見られており、整合性を重要視してこれ以上貼り合わせないか、それとも整合性を多少犠牲にして視野を広げるか、といった判断がパラメータ設定で制御されていることが分かる。

	Period-A		Period-B	
(a)			5 images, $Q = -4.241$	4 images, $Q = -4.302$
(b)			41 images, $Q = -4.297$	52 images, $Q = -4.367$
	Period-C	Period-D	Period-E	
(a)				
	8 images, $Q = -4.382$	12 images, $Q = -4.651$	13 images, $Q = -4.489$	
(b)				
	93 images, $Q = -4.424$	94 images, $Q = -4.670$	100 images, $Q = -4.500$	
	Period-F		Period-G	
(a)				
	7 images, $Q = -4.678$	10 images, $Q = -4.515$		
(b)				
	92 images, $Q = -4.740$	77 images, $Q = -4.691$		

図 7 貼り合わせ合成結果の比較. (a) 提案手法: 貼り合わせ評価値が最大となる画像の組を選択して合成した場合. (b) 従来手法: 画像選択を行わずに全てを貼り合わせ合成した場合

Fig. 7 The synthesized results with the proposed and conventional methods.

	Period-A	Period-B
(a)	 $Q^{PS_a}(\{\mathbf{I}_a\}) = -4.241, Q^{PS_b}(\{\mathbf{I}_a\}) = -5.231$	 $Q^{PS_a}(\{\mathbf{I}_a\}) = -4.302, Q^{PS_b}(\{\mathbf{I}_a\}) = -5.268$
(b)	 $Q^{PS_a}(\{\mathbf{I}_b\}) = -4.258, Q^{PS_b}(\{\mathbf{I}_b\}) = -5.211$	 $Q^{PS_a}(\{\mathbf{I}_b\}) = -4.337, Q^{PS_b}(\{\mathbf{I}_b\}) = -5.246$
(c)	$\Delta Q = 0.037$	$\Delta Q = 0.057$





	Period-F	Period-G
(a)	 $Q^{PS_a}(\{\mathbf{I}_a\}) = -4.678, Q^{PS_b}(\{\mathbf{I}_a\}) = -5.334$	 $Q^{PS_a}(\{\mathbf{I}_a\}) = -4.515, Q^{PS_b}(\{\mathbf{I}_a\}) = -5.309$
(b)	 $Q^{PS_a}(\{\mathbf{I}_b\}) = -4.680, Q^{PS_b}(\{\mathbf{I}_b\}) = -5.327$	 $Q^{PS_a}(\{\mathbf{I}_b\}) = -4.522, Q^{PS_b}(\{\mathbf{I}_b\}) = -5.273$
(c)	$\Delta Q = 0.009$	$\Delta Q = 0.043$

図 8 異なるパラメータ設定 $PS_a = \{\alpha = 0.5, \sigma_1 = 5.0, \sigma_2 = 5.0\}$ 及び $PS_b = \{\alpha = 0.5, \sigma_1 = 5.0, \sigma_2 = 30.0\}$ を用いた場合の貼り合わせ合成結果比較。(a) パラメータセット PS_a を用いた場合に貼り合わせ評価値が最大となった画像の組 $\{\mathbf{I}_a\}$ についての結果。(b) パラメータセット PS_b を用いた場合に貼り合わせ評価値が最大となった画像の組 $\{\mathbf{I}_b\}$ についての結果。(c) パラメータを変更することで (a) と (b) がもつ評価値の差がどれだけ変動したかを表す ΔQ

Fig. 8 The synthesized results with the different parameter configurations.

定量的な影響を検証するため、異なるパラメータで貼り合わせ評価値を導出したときに特定の2組の画像群についての大小関係がどの程度変化するかを検証した。具体的には PS_a, PS_b の各設定で最大の評価値

となる画像の組み合わせ $\{\mathbf{I}_a\}, \{\mathbf{I}_b\}$ について、その最大評価値 $Q^{PS_a}(\{\mathbf{I}_a\}), Q^{PS_b}(\{\mathbf{I}_b\})$ と他方のパラメータにおける評価値 $Q^{PS_b}(\{\mathbf{I}_a\}), Q^{PS_a}(\{\mathbf{I}_b\})$ を取り出しその差の変化 $\Delta Q = |Q^{PS_a}(\{\mathbf{I}_a\}) - Q^{PS_a}(\{\mathbf{I}_b\}) -$

$(Q^{PS_b}(\{\mathbf{I}_a\}) - Q^{PS_b}(\{\mathbf{I}_b\}))$ を算出した(図8(c)). 変化量は入力画像列により様々であったが多くの場合で0.05ポイント以下であった。これは評価値の絶対値に対してかなり小さな値である。この結果から現状の評価関数の仕組みでは「貼り合わせによる視野の広がり」と「共通領域での整合性」のバランスを高感度に制御することは難しいことが確認された。実際、図8のPeriod-Fなどではパラメータセット PS_b を用いた方が広視野画像を出力しているが、その差はごく少量であったため目視で確認することは難しく、見た目にはパラメータの違いがほとんど現れていない。提案手法を様々な用途に適用するには、共通領域の整合性に対してより厳しくできる仕組みや逆にもっと視野を広げられるような仕組みも必要と考える。

5.3 代表画像化におけるパラメータ設定

代表画像化を制御するしきい値パラメータ r_{th}, w_{th} と探索量との関係について検証し、これらの値をどのように設定すればよいかについて考察する。ただししきい値パラメータの効果は入力映像の内容に大きく依存し絶対的な指標を探すことは難しい。そこで特定のシーンを対象として、実用的な枚数の入力画像と探索量(計算時間)を与えるしきい値パラメータについて調査することで、一般的な入力画像におけるパラメータ設定の指標を明らかにする。

データセット 5.1での入力画像列からPeriod-E(屋内:100frames)及びPeriod-G(屋外:77frames)に対応する個人視点画像列を選んで用いた。これらの元映像はおおよそ5から7秒で一般的な見直し動作時間に

相当する。

パラメータ設定の範囲 どのような値が実用的であるのかを確認するためある程度網羅的に与えた。ただしあまりに小さい値では代表画像が少数枚となってしまうので、少なくとも10枚程度の代表画像が抽出されるような範囲とした。具体的には r_{th} は0.80から0.88まで、 w_{th} は0.94から0.97まで、どちらも0.1刻みとした。

実験結果と考察 実験により得られたパラメータと探索量の関係を表1に示す。パラメータ値が小さいほど少ない探索量で済む相対的な傾向は確認されたが、探索量の絶対数はやはり入力画像に依存してまちまちであった。その一方で抽出された代表画像の数と探索量にはある程度の相関がみられた(図9)。貼り合わせ領域の大きさや実装方法にもよるが、1組画像群に対する貼り合わせ評価値の計算には現状で平均して30ミリ秒程度要している。この値を元に仮に数十秒から数分を実用的な計算時間とすると許容される代表画像の枚数は18枚以下となる。すなわち、事前情報なしにパラメータ値のみから直接計算時間を見積もることは難しいが、抽出される代表画像の数からはある程度予測可能であることが確認された。代表画像数の計算方法は4.1で述べたとおりであり、この処理に必要な時間はごく短い。そのため複数のパラメータを用いて代表画像の数のみを計算し、その結果に基づいてどのパラメータ値を採用するかを決める方法が有効と考えられる。

表1 類似画像列の代表画像化におけるしきい値パラメータ r_{th} (横軸)・ w_{th} (縦軸)と探索量の関係。各セル中の左項は抽出された代表画像の枚数、右項は探索量を示す。また“-”は組み合わせ爆発により計算が終了しなかったことを表す

Table 1 # of extracted representative images and search amounts with various thresholds r_{th} and w_{th} .

Period-E シーンを入力画像としたとき。

	0.80	0.81	0.82	0.83	0.84	0.85	0.86	0.87	0.88
0.94	17, 5.5×10^2	18, 1.0×10^3	18, 1.0×10^3	19, 5.7×10^3	20, 1.5×10^4	21, 4.8×10^4	24, 3.0×10^5	24, 4.2×10^5	27, 3.3×10^6
0.95	17, 5.5×10^2	18, 1.0×10^3	18, 1.0×10^3	19, 5.7×10^3	20, 1.5×10^4	21, 4.8×10^4	24, 3.0×10^5	24, 4.2×10^5	27, 3.3×10^6
0.96	18, 2.4×10^2	18, 1.6×10^3	19, 3.8×10^3	19, 5.7×10^3	20, 1.5×10^4	21, 4.8×10^4	24, 3.0×10^5	42, 4.2×10^5	27, 3.3×10^6
0.97	20, 2.1×10^3	21, 5.7×10^4	21, 5.4×10^4	22, 1.1×10^5	23, 2.1×10^5	24, 9.4×10^5	26, 3.1×10^6	26, 4.5×10^6	29, 3.1×10^7

Period-G シーンを入力画像としたとき。

	0.80	0.81	0.82	0.83	0.84	0.85	0.86	0.87	0.88
0.94	14, 1.4×10^2	16, 1.2×10^3	17, 1.2×10^3	17, 3.1×10^3	17, 3.6×10^3	17, 3.6×10^3	19, 3.2×10^4	19, 3.3×10^4	20, 8.9×10^4
0.95	17, 3.1×10^3	18, 5.6×10^3	19, 5.6×10^3	19, 1.5×10^4	19, 1.5×10^4	19, 1.5×10^4	20, 6.6×10^4	21, 1.4×10^5	21, 1.9×10^5
0.96	22, 4.1×10^5	23, 5.6×10^5	23, 5.6×10^5	23, 5.6×10^5	23, 5.6×10^5	23, 5.6×10^5	23, 5.6×10^5	23, 5.6×10^5	24, 1.5×10^6
0.97	-	-	-	-	-	-	-	-	-

表 2 不整合な三つ組画像の同時不使用におけるしきい値パラメータ Q_{th} と探索量の関係 (ただし $r_{th} = 0.84, w_{th} = 0.95$). 各セル中の左項は同時不使用とされた三つ組画像の数, 右項は探索量を示す

Table 2 # of search amounts with various thresholds Q_{th} .

Q_{th}	-4.50	-4.45	-4.40	-4.35	-4.30	-4.25	-4.20	-4.15	-4.10	-4.05	-4.00	-3.95
Period-E	0, 15217	0, 15217	0, 15217	0, 15217	1, 12049	4, 9409	5, 8881	6, 8683	6, 8683	9, 5779	11, 4525	22, 1003
Period-G	2, 10255	6, 7951	11, 5647	12, 5263	16, 3871	18, 2845	20, 2509	20, 2509	22, 1673	22, 1673	22, 1673	22, 1673

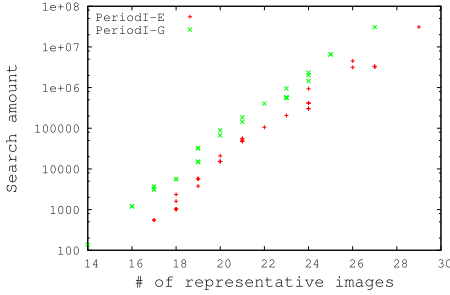


図 9 代表画像の枚数と探索量の関係

Fig. 9 The relationship between # of the representative images and # of search for image selection.

5.4 不整合な三つ組画像の同時不使用におけるパラメータ設定

不整合な三つ組画像を取り出す際のパラメータ Q_{th} についても検証を行った。 Q_{th} による探索量の変化は入力画像の内容だけでなく組み合わせる r_{th}, w_{th} の値にも依存するため、網羅的に調査することは難しい。そこで前節と同様に計算時間を現実的な量に収める観点から検討した。具体的には、計算時間がやや多いとされる r_{th}, w_{th} に対して様々な Q_{th} を設定して探索量を導出し、実用的な計算時間となるような値について確認を行った。

データセット 5.3 で用いたものと同一の個人視点映像 (Period-E, Period-G)。

パラメータ設定の範囲 5.3 で現実的とされた代表画像枚数である 18 枚からやや多い 19 や 20 枚となるパラメータ $r_{th} = 0.84, w_{th} = 0.95$ を対象とした。また Q_{th} の範囲は列挙された三つ組画像の共通領域における貼り合わせ評価値 $Q_3(\mathbf{I}_{klm})$ の上限と下限を含むように設定した。

実験結果と考察 パラメータ Q_{th} と探索量の関係を表 2 に示す。代表画像化に比べて効果は小さいが、10 から 20 組程度の三つ組画像を同時不使用とすれば 18 枚の代表画像のときと同等の計算量まで削減できることが確認できた。しかしパラメータ値と探索量の間に一般的な関係を見いだすことはできなかった。これは

映像の内容に大きく依存する性質が予想通りに現れた結果と考えられる。以上の結果から Q_{th} の決め方には次のような方法が有効と考えられる。まず本実験のように網羅的な Q_{th} を用いて探索量のみについて調査を行う。探索量のみを計算すること、 Q_3 の導出は限定された画像の組と領域が対象であること、 Q_3 の取り得る範囲はそれほど大きくないことから、この手続きに要する時間はほぼ問題とならない。本実験においても数秒程度の計算時間であった。その上で現実的な時間で計算が終了するようなパラメータ Q_{th} を選んで貼り合わせ評価値の計算を行い最適な画像の組を選択する、という手続きである。

6. 位置合わせ手法との関連性

6.1 選択された画像群の再位置合わせ・変形の効果

本研究では入力画像全てを用いた場合の位置合わせ・変形に対して貼り合わせ評価値の導出・最適な画像群の選択を行っている。画像選択後の広視野画像合成にもこのときの変形画像を用いているが、より整合性の高い貼り合わせを得るためには選択された画像群のみを用いて再度位置合わせ・変形する方法も考えられる。そこで選択後の画像群を再位置合わせ・変形して合成すると最終的な画像の品質がどの程度向上するのかを評価した。その結果、図 10 に示す事例を含めた多くの場合で、再位置合わせした方がより良い画質の貼り合わせ画像となることが確認された。提案手法で貼り合わせに適した画像群が選択されているため、入力画像全てを対象とするよりも整合性の高い位置合わせ・変形が得られたからと考える。ただし十分な共通領域をもたない画像ペアがしばしば現れ、位置合わせに失敗することもあった。そのようなペアには入力画像全体を対象としたときの位置合わせ結果を利用するなどの工夫が必要である。

6.2 位置合わせに要する処理時間

入力画像列の位置合わせ処理 [2] では画像ペア間の位置合わせの連結で表された初期値に対して全体最適化を行う。一般に入力画像枚数が増えるほどに最適化



図 10 選択された画像のみを用いた再位置合わせ・変形の効果. (a) 全入力画像を対象とした位置合わせ・変形を用いた場合 (図 7(a) の一部を再掲載したもの). (b) 選択された画像を対象として再度位置合わせ・変形した画像を用いた場合. 再位置合わせにより多少貼り合わせ画像全体の形が変化しているが, 右端領域 (Period-E) や中央左よりの自転車領域 (Period-G) における画質が向上しているのが確認できる

Fig. 10 Effect of re-registrating the selected images.

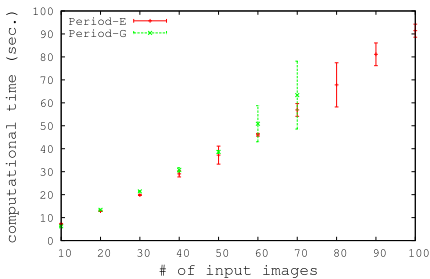


図 11 入力画像枚数の増加に伴う位置合わせ計算時間の変化

Fig. 11 The computational times for various # of input images.

パラメータの数が増大し収束に時間を要すると考えられる. そこで位置合わせ処理における入力画像枚数と計算時間の関係を検証することで, 実用的な入力画像の枚数に対して現実的な計算時間となっているか確認した. 入力画像には 5.3 と同じ画像列を, 位置合わせ処理の実装には Hugin [22] を用いた. その結果 100 枚程度の入力画像ではやや非線形ではあるもののほぼ線形とみなしてもよい計算時間であることが確認された (図 11). 最適化時間は非線形に増加しているがあまり支配的ではなく, 局所特徴量の抽出や画像ペア間の位置合わせ推定に多くの時間が割かれていることが

原因と考えられる.

7. む す び

本論文では個人視点映像から貼り合わせに適した画像の組み合わせを自動的に選び出して広視野画像を合成する方法について述べた. 画像の選択は貼り合わせの良さを定量化した評価値に基づいて行う. 我々は貼り合わせることによる効果を視野が広がることと共通領域における画素値の信頼性が增大することとみなし, それらを一括して扱うことができる仕組みとして画素値の確からしさに注目した. 加えて貼り合わせに適した画像の組み合わせを近似的に高速探索するために, 類似した画像列を一枚の画像に置き換える方法と貼り合わせ評価値の悪い三つ組画像を同時には使用しない方法を提案した. その一方で「複数画像の共通領域における整合性」と「貼り合わせによる視野の広がり」に対する重み調整の方法や計算量削減について更なる検討が必要であることも確認された.

本研究では画像選択の問題のみに焦点をあてるために, 入力画像間の位置合わせや変形は適用済みであるという前提をおき, 合成には単純に画素値の平均を用いた. また提案評価値と人間が見たときの画質との関係性については厳密には考慮していない. より良い貼

り合わせ画像を得るためにはこれらの問題も合わせて扱う必要がある。最終的には「人物が撮影されていたときには必ず貼り合わせ画像に登場させる」といったように意味的な側面も考慮してカメラ装着者の体験をうまく取り出すような仕組みに向けて検討を進めたいと考えている。

文 献

- [1] A. Sibiryakov and M. Bober, "Graph-based multiple panorama extraction from unordered image sets," Proc. SPIE 6498, Computational Imaging V, 2007. doi:10.1117/12.704025
- [2] M. Brown and D. Lowe, "Automatic panoramic image stitching using invariant features," Int. J. Comput. Vis., vol.74, no.1, pp.59–73, 2007.
- [3] J. Hannuksela, P. Sangi, J. Heikkila, L. Xu, and D. Doermann, "Document image mosaicing with mobile phones," Proc. 14th International Conference on Image Analysis and Processing (ICIAP2007), pp.575–582, 2007.
- [4] M. Ligang and Y. Yongjuan, "Automatic document image mosaicing algorithm with hand-held camera," Proc. 2nd International Conference on Intelligent Control and Information Processing (ICICIP2011), pp.1094–1097, 2011.
- [5] T. Coito, J.R.C. Pinto, and J. Azinheira, "Building and evaluation of a mosaic of images using aerial photographs," Proc. 10th International Conference on Image Analysis and Recognition (ICIAR2013), pp.798–805, 2013.
- [6] T. Suzuki, Y. Amano, T. Hashizume, S. Suzuki, and A. Yamaba, "Generation of large mosaic images for vegetation monitoring using a small unmanned aerial vehicle," J. Robotics and Mechatronics, vol.22, no.2, pp.212–220, 2010.
- [7] K. Aizawa, S. Kawasaki, T. Ishikawa, and T. Yamasaki, "Capture and retrieval of life log," Proc. Int. Conf. on Artificial Reality and Telexistence (ICAT), pp.49–55, 2004.
- [8] B.H. Prananto, I. Kim, and H. Kim, "Multi-level experience retrieval for the personal lifelog media system," Proc. Third Int. IEEE Conf. on Signal-Image Technologies and Internet-Based System (SITIS), pp.175–182, 2007.
- [9] 松井研太, 近藤一見, 小泉敬寛, 中村裕一, "個人視点映像からの広視野画像の自動生成—輝度値の確率分布に基づいた貼り合わせに適した画像群の選択—," 信学技報, MVE2013-21, 2013.
- [10] 松井研太, 近藤一見, 小泉敬寛, 中村裕一, "個人視点映像を一覧するための広視野貼り合わせ画像群の自動生成—貼り合わせの良さに基づいた画像の選択とグループ化—," HCG シンポジウム 2013, pp.54–61, 2013.
- [11] Y. Sumi, S. Ito, T. Matsuguchi, S. Fels, S. Iwasawa, K. Mase, K. Kogure, and N. Hagita, "Collaborative capturing, interpreting, and sharing of experiences," Personal and Ubiquitous Computing, vol.11, no.4, pp.213–328, 2007.
- [12] 近藤一見, 高瀬恵三郎, 小泉敬寛, 中村裕一, 森 幹彦, 喜多 一, "個人視点映像を用いた気づき体験の回想と整理支援—フィールド調査における問題発見を通じて—," 信学技報, PRMU2010-128, 2010.
- [13] 岡田昌也, 鳥山朋二, 多田昌裕, 角 康之, 間瀬健二, 小暮潔, 萩田紀博, "実世界重要体験の抽出・再現に基づく事後学習支援手法の提案," 信学論 (D-II), vol. J91-D-II, no.1, pp.65–77, Jan. 2008.
- [14] Z. Cernekova, I. Pitas, and C. Nikou, "Information theory-based shot cut/fade detection and video summarization," IEEE Trans. Circuits Syst. Video Technol., vol.16, no.1, pp.82–91, 2006.
- [15] A. Amiri and M. Fathy, "Video shot boundary detection using generalized eigenvalue decomposition and Gaussian transition detection," Computing and Informatics, vol.30, no.3, pp.595–619, 2011.
- [16] Z.M. Lu and Y. Shi, "Fast video shot boundary detection based on SVD and pattern matching," IEEE Trans. Image Process., vol.22, no.12, pp.5136–5145, 2013.
- [17] Y.J. Lee, J. Ghosh, and K. Grauman, "Discovering important people and objects for egocentric video summarization," Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR2012), pp.1346–1353, 2012.
- [18] E.H. Spriggs, F.T. Frade, and M. Hebert, "Temporal segmentation and activity classification from first-person sensing," IEEE Workshop on Egocentric Vision, pp.17–24, 2009.
- [19] A.R. Doherty, A.F. Smeaton, K. Lee, and D.P. Ellis, "Multimodal segmentation of lifelog data," Proc. RIAO2007 — Large-Scale Semantic Access to Content, pp.21–38, 2007.
- [20] M.E. Tipping and C.M. Bishop, "Bayesian image super-resolution," NIPS 15, pp.1279–1286, MIT Press, 2003.
- [21] G.K. Chantas, N.P. Galatsanos, and N.A. Woods, "Super-resolution based on fast registration and maximum a posteriori reconstruction," IEEE Trans. Image Process., vol.16, no.7, pp.1821–1830, 2007.
- [22] Hugin — Panorama photo stitcher,
<http://hugin.sourceforge.net/>
(平成 26 年 4 月 8 日受付, 8 月 15 日再受付)



近藤 一晃 (正員)

2002年大阪大学基礎工学部システム科学科卒業。2004年同大学大学院基礎工学研究科システム人間系専攻博士前期課程修了。2007年同大学大学院情報科学研究科コンピュータサイエンス専攻博士後期課程修了。同年同大学産業科学研究所特任研究員。2009年京都大学学術情報メディアセンター助教就任後現在に至る。反射屈折光学系，知能ロボット，マンマシンインタラクション，知的行動支援に関する研究に従事。博士(情報科学)。情報処理学会会員。



松井 研太

2012京大・工・電気電子卒。2014同大大学院工学研究科電気工学専攻博士前期課程了。装着型カメラの映像処理に関する研究に従事。



小泉 敬寛 (正員)

2003年筑波大学第三学群工学システム学類卒業。2005年同大学大学院工学研究科知能機能システム専攻博士課程修士号取得・退学。2007年京都大学大学院工学研究科電気工学専攻博士課程中退。同年京都大学大学院工学研究科助教。個人視点映像を用いた個人行動記録の意味的検索手法についての研究等に従事。修士(工学)。情報処理学会会員。



中村 裕一 (正員)

1985年京都大学工学部電気工学第二学科卒。1990年同大学院博士課程修了。同年京都大学工学部助手。1993年筑波大学電子・情報工学系講師。1999年機能工学系助教授。2004年京都大学学術情報メディアセンター教授。現在に至る。博士(工学)。画像理解，映像処理，自然言語処理などの研究に従事。1996年カーネギーメロン大学ロボティクス研究所客員研究員。1998～2001年科学技術振興事業団さきがけ21研究「情報と知」領域研究員(兼任)。情報処理学会，人工知能学会，ACM，IEEE各会員。