

自転車位置推定のための車載全方位カメラ映像と 通常カメラ映像のフレーム間対応付け

内山 寛之[†] 高橋 友和^{†,††} 井手 一郎[†] 村瀬 洋[†]

[†] 名古屋大学大学院 情報科学研究科 〒464-8601 愛知県名古屋市千種区不老町

^{††} 日本学術振興会

E-mail: †{uchiyama,ttakahashi,ide,murase}@murase.m.is.nagoya-u.ac.jp

あらまし 本報告では、車載カメラを使用した自転車位置推定を目的とし、通常カメラ映像と全方位カメラ映像のフレーム間を対応付ける手法を提案する。車載通常カメラからの入力画像と、事前に作成された位置情報が付随する全方位カメラ映像データベースの画像を比較することにより、画像取得時に車両の位置を推定することを想定する。両カメラの特性を利用して、各映像をそれぞれ円筒面上に投影してマッチング処理を行う。通常カメラは全方位カメラよりも撮影範囲が狭いため、フレーム方向の探索に加え、フレーム中の縦列位置の探索が必要となり、探索空間は3次元となる。本手法では、単一フレーム同士の対応付けを行う代わりに時系列の複数のフレーム同士をDPマッチングをベースとした手法により対応付ける。時系列情報を利用することで対応付けの精度向上と探索コストの削減を図る。実際に走行して撮影した映像を用いて評価実験を行い、提案手法の有効性を確認した。

キーワード ITS, 全方位画像, DP マッチング

Frame Registration between In-vehicle Omni-directional and Normal Cameras for Self-position Estimation

Hiroyuki UCHIYAMA[†], Tomokazu TAKAHASHI^{†,††}, Ichiro IDE[†], and Hiroshi MURASE[†]

[†] Graduate School of Information Science, Nagoya University Furo-cho, Chikusa-ku, Nagoya-shi, Aichi
464-8601 Japan

^{††} Japan Society for the Promotion of Science

E-mail: †{uchiyama,ttakahashi,ide,murase}@murase.m.is.nagoya-u.ac.jp

Abstract We propose a method for frame registration of in-vehicle normal camera with omni-directional camera for self-position estimation. In this paper, we assume that the position is estimated by aligning the input in-vehicle normal camera video frames with the frames of the omni-directional camera video database which has position information. The images are projected to a cylindrical surface and the matching process is conducted on it. Because the view angle of the normal camera is smaller than that of the omni-directional camera, we need to align vertical and horizontal position in addition to frame sequences. A DTW based method is applied in order to achieve high accuracy and low computing cost. The experimental results demonstrate the effectiveness of the proposed method.

Key words ITS, omni-directional image, DP matching

1. ま え が き

カーナビゲーションシステムの発展に伴って、高精度に自転車位置推定を行うことが求められている。現在のシステムは通常、GPS (Global Positioning System) および車速パルス、ジャイロなどを利用した自律航法により自転車位置を推定している。しかし、谷間や高架下など GPS 衛星の電波が受信困難な場所で

は、自律航法を併用しても位置推定誤差が徐々に蓄積されていくという問題がある。また、GPS の測定値自体にも数メートルから数 10 メートルの誤差が含まれる。その解決方法の一つとして、車載カメラからの映像を利用して自転車位置推定を行うアプローチが研究されている [1]。車載カメラはすでに視界補助を目的として一般に普及しており、レーザレーダやミリ波レーダなどの他のセンサに比べ安価であるという利点がある。一方で、

施設情報などのテキスト情報や風景などの画像情報など、地図上に様々な情報を埋め込む研究が行われている。Google Maps [2] の「Street View」など、地図上に埋め込まれた全方位カメラ映像により街並み風景を提供する GIS サービスが実際に稼働している。今後、自車位置推定のデータベースの基となる風景の実映像を大規模に収集することは可能になると思われる。

佐藤らは GPS 座標が付随する複数の全方位カメラ映像の組を DP マッチングにより対応付け、地点ごとに GPS 座標を平均化することにより位置推定精度を向上させる手法を提案している [3]。これに対し、通常カメラを用いた自己位置推定手法は数多く提案されている。W. Lee らはキャリブレーションを行っていないカメラを用いてロボットの位置を推定する手法を提案している [4]。近年、一般車には全方位カメラではなく通常カメラが多く搭載されている。しかし、通常カメラは画角が小さいため、カメラの姿勢が正確にわかっていない場合、通常カメラ映像同士のマッチングは困難である。池田らは全方位カメラを用いて特徴点を風景から抽出し、3次元のランドマークデータベースを構築する手法を提案しており、それを用いて通常カメラの位置・姿勢を推定する実験を行っている [5]。複数の通常カメラ間のレジストレーション手法として、小野らは時空間画像を球として表現し、進行方向に対して任意姿勢のカメラを一列に配置することにより、ゆがみのないレジストレーションを行う手法を提案している [6]。川崎らは EPI (Epipolar Plane Image) を利用し、カメラ間の対応付けを行う手法を提案している [7]。これらの 2 つの手法は同じ走行で撮影された映像を対象としている。

本研究では、車載カメラから取得した映像と、位置情報があらかじめ付加された車載カメラ映像データベースを比較することにより、自車位置推定を行うシステムを考える。通常、データベース作成時のカメラと自車位置推定時のカメラは異なる。その極端な例として、本報告では全方位カメラ映像と通常カメラ映像のフレーム間を対応付ける手法を提案する。その際、DP マッチングを用いる。通常カメラの姿勢は水平であり、光軸の方向はわからないものとする。単純な画像比較によるマッチング手法においても、時系列情報を用いることで精度の高いマッチングが行えることを示す。

2. 提案手法

2.1 概要

図 1 に提案手法の概要を示す。はじめに、全方位カメラと通常カメラでは撮影機器の違いにより明るさや色相が異なるため、正規化処理としてそれぞれの映像の RGB 各成分についてヒストグラムを平坦化する。次に、それぞれの映像を円筒面上に投影する。そして、円筒面上の全方位カメラフレーム中に複数の照合窓を配置する。DP マッチングをベースとした手法により、2 つの映像間の共通する区間を抽出し、フレームの対応付けを行う。以下の節で提案手法の詳細について述べる。

2.2 円筒面上における照合窓による画像切り出し

ヒストグラム平坦化の後、全方位カメラ映像および通常カメラ映像を円筒面上に投影する。本研究では、全方位カメラとし

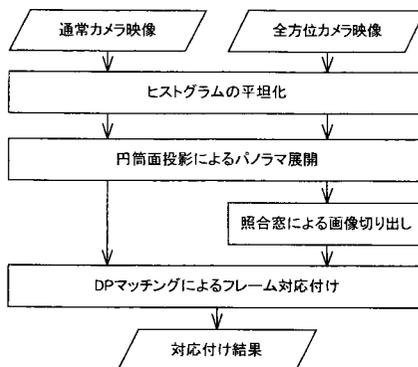
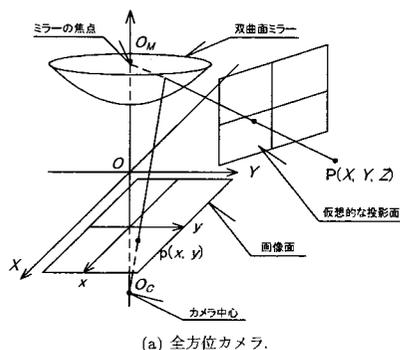
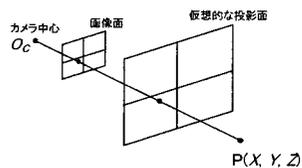


図 1 フレーム間対応付け処理の流れ。



(a) 全方位カメラ。



(b) 通常カメラ。

図 2 カメラモデル。

て Hyper Omni Vision [8] を使用する。このセンサは双曲面ミラーで構成されており、中心射影の特性を持つ。そのため、投影面を仮定することで通常カメラと同様にピンホールカメラモデルで扱うことができる (図 2)。撮影時、全方位カメラの双曲面ミラーの焦点の軌跡と通常カメラの光学中心の軌跡がほぼ同じであると仮定する場合、両フレーム画像を円筒面上に投影することで、画像の比較を平行移動のみで行うことができる。次に、円筒面に投影した全方位カメラ映像上に通常カメラ映像と同じ大きさの矩形の照合窓を複数配置する (図 3)。照合窓の数は水平方向 X 、鉛直方向 Y の計 $X \times Y$ 個とする。RGB カラー画像を使用し、照合窓により切り取られる全方位カメラ映像の 1 フレームと通常カメラ映像の 1 フレームを画素の要素の輝度値の SAD (Sum of Absolute Difference) により比較する。

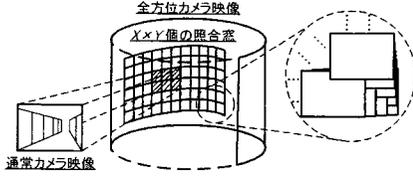


図3 映像は円筒面上に投影され、多数の照合窓が少しずつずれながら重なり合って設置される。

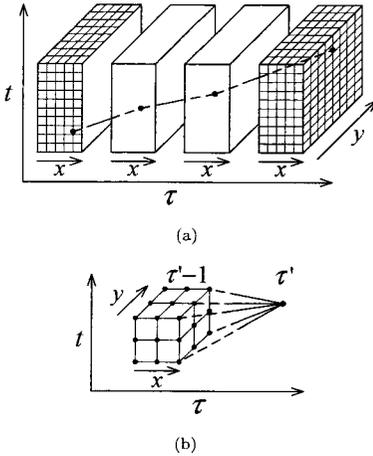


図4 DP マッチングの概要図。(a) 計算空間は4次元で表現される。(b) DP パス：局所的に最適な経路が27経路から選択される。

$$SAD = \sum_x \sum_y \sum_c |I_1(x, y, c) - I_2(x, y, c)| \quad (1)$$

ここで、 $I_N(x, y, c)$ は座標 (x, y) の画素の c 番目の要素 ($c = \{r, g, b\}$) の輝度値である。本研究では、それぞれのカメラの内部パラメータは既知であるものとする。

2.3 DP マッチングによるフレーム間対応付け

通常カメラ映像 $V_{normal}(\tau)$ および座標 (x, y) の照合窓により切り出される全方位カメラ映像 $V_{omni}(x, y, t)$ はそれぞれ1次元および3次元のデータ系列と見ることができる。迫江らはRSM (Rubber String Matching) 法とよばれるDP マッチングを基にした1次元—2次元系列間の非線形パターンマッチング手法を提案している [9]。この手法では3次元の計算空間において、局所的にコストが低い経路を求めていくことにより、大域的に最小距離 (あるいは最大類似度) の経路を求める。本研究では、通常カメラ映像と共通する区間を全方位カメラ映像から抽出するために、RSM 法を1次元—3次元間のマッチングに拡張する (図4)。この場合、計算空間は4次元となる。以下の実験では図4 (b) のようなDP パスを用い、27経路から最小のコストを持つ経路を選択する。 T フレームからなる全方位カメラ映像データベースから、 τ 番目の通常カメラ映像のフレームに対応する最適な照合窓の位置 (x_{est}, y_{est}) および全方位カメラ映像のフレーム t_{est} を推定するアルゴリズムは以下の

ようになる。

Local distance

$$d(x, y, t, \tau) = SAD\{V_{omni}(x, y, t), V_{normal}(\tau)\} \quad (2)$$

Initial condition

$$\begin{cases} g(-1, y, t, \tau) = g(x, -1, t, \tau) = \infty \\ g(X, y, t, \tau) = g(x, Y, t, \tau) = \infty \\ g(x, y, -1, \tau) = g(x, y, -2, \tau) = \infty \end{cases} \quad (3)$$

Iteration ($\tau = 0, 1, \dots$)

for $\tau = 0$

$$g(x, y, t, 0) = d(x, y, t, 0) \quad (4)$$

for $\tau \geq 1$

$$g(x, y, t, \tau) = \min_{\substack{-1 \leq \Delta x \leq 1 \\ -1 \leq \Delta y \leq 1 \\ 0 \leq \Delta t \leq 2}} \{g(x - \Delta x, y - \Delta y, t - \Delta t, \tau - 1) + d(x, y, t, \tau)\} \quad (5)$$

Output ($\tau = 0, 1, \dots$)

推定パラメータ x_{est} , y_{est} and t_{est} . ここで、

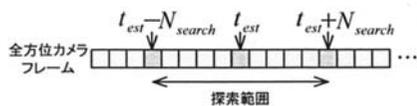
$$g(x_{est}, y_{est}, t_{est}, \tau) = \min_{\substack{0 \leq x \leq X-1 \\ 0 \leq y \leq Y-1 \\ 0 \leq t \leq T-1}} g(x, y, t, \tau) \quad (6)$$

まず、ローカルディスタンスは各画像間のSADで定義される。初期条件として、計算範囲外のグローバルディスタンスを無限大とする。通常カメラ映像のフレームが入力されるごとに式(4)、(5)の漸化式を全ての x, y, t について計算する。そして、提案手法ではリアルタイム処理を想定しているため、計算空間に対してバックトレースは行わず、入力された通常カメラ映像のその時点における最終フレームに対して推定された t_{est} を対応するフレームとみなす。

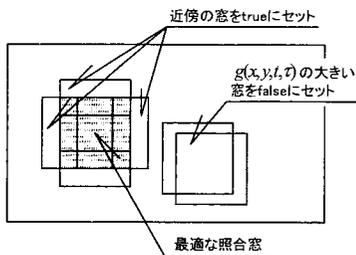
2.4 照合範囲制限による計算量削減

式(5)は多くの計算量を必要とするため、いくつかの制限を加えることで高速化を図る。まず、実際の車両の位置は滑らかに変化すると考えられるため、全方位カメラ映像データベース内において1フレーム前の位置の周辺のみを探索する。従って、 $g(x, y, t, \tau)$ は $t_{est} - N_{search} \leq t \leq t_{est} + N_{search}$ の範囲のみを計算する (図5 (a))。ここで、 N_{search} は閾値である。

次に、各照合窓に対しフラグを用意し、フラグが true であったときのみその照合窓に対応する $g(x, y, t, \tau)$ を計算するものとする (図5 (b))。値の小さい $g(x, y, t, \tau)$ に対応する照合窓は次の反復において最適な照合窓になりやすい。そこで、 $g(x, y, t, \tau)$ を昇順にソートし、 N_{sort} 番目以内に入らなかった $g(x, y, t, \tau)$ に対応する照合窓のフラグを false とする。さらに、最適な照合窓の位置が滑らかに変化すると仮定すると、最適な照合窓の近傍の照合窓が次の反復において最適な照合窓になりやすくなる。そこで、最適な照合窓の近傍の照合窓のフラグを true とする。このような処理を行うことで次第に照合を行う範囲が限



(a) 探索範囲の制限.



(b) 照合範囲の制限.

図5 計算量削減方法の概略図.

表1 カメラの仕様.

全方位カメラ	
外形寸法	41 Φ × 86.5H
ミラー径	30 Φ
撮像範囲※	上方約 10 度, 下方約 55 度
撮像素子	25 万画素カラー CCD
映像出力	NTSC
通常カメラ	
レンズ	(ワイド端) f=2.45mm
	35mm 換算: 44.7mm
撮像素子	80 万画素カラー CCD
信号方式	NTSC

※ただし, 天地逆転させて使用.

定されていき, 最適な照合窓の周辺だけに照合窓が残ることになる. また, 通常カメラは固定されているため最適な照合窓の位置が大きく変化することはないと考えられるが, 道路の窪みによる上下運動や車両の軌道の変化などのために画像の見かけが変化することが想定される. そのため, 計算を行う照合窓が完全に1点に収束しないよう, 最適な照合窓の近傍の照合窓のフラグを true にする操作を行う.

3. 評価実験

提案手法の有効性を確認するために, 実際に路上を走行して撮影した映像を用いた評価実験を行った.

3.1 実験条件

全方位カメラとして Viston VS-C14N を使用し, 車両の屋根に鉛直上向きに取り付けた. 通常カメラは民生用のデジタルビデオカメラ Panasonic NV-GS150-S を使用し, 車両のフロントガラスの左隅に水平に固定した. フレームレートはいずれも 30 [frames/sec] である. 実験に使用したカメラの取り付け位置を図6に, 仕様を表1に示す.

評価のための位置情報付き全方位カメラ映像データベースを作成するため, まず全方位カメラ映像と衛星写真を比較し, 建

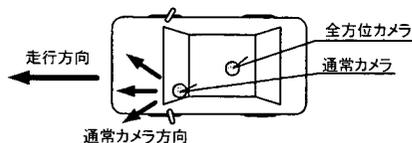
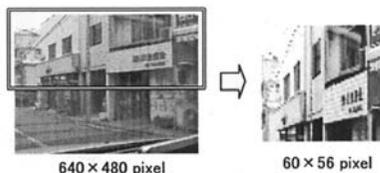


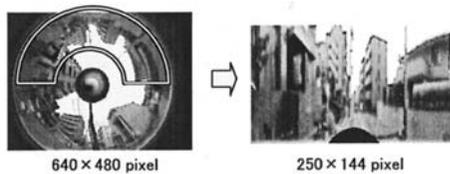
図6 カメラの取り付け位置.

表2 実験に用いたパラメータの値.

パラメータ	値
照合窓の水平方向の数 X	55
照合窓の鉛直方向の数 Y	28
探索幅 N_{search}	100
照合窓閾値 N_{sort}	200



(a) 通常カメラ映像.



(b) 全方位カメラ映像.

図7 画像の使用領域.

物の角や電柱など特徴となる地点を手で対応付けた. そして, 位置座標を線形補完することにより映像のフレームと位置情報とを対応付けた. 実際には GPS を搭載して位置座標を測定したが, 手で位置座標を入力の方がより精度が高いためにこのような方法を用いた. 全方位カメラ映像データベースの区間は通常カメラ映像の区間より長く, フレーム数は 1600 であった. 通常カメラの方向は 1) 前方, 2) 前方より左へ約 15 [deg], 3) 前方より右へ約 35 [deg] の 3 つについて実験を行う. 通常カメラ撮影時の経路長は約 400 [m] で, フレーム数は各 1046, 1195, 1316 である. 全方位カメラ映像と通常カメラ映像をそれぞれ異なる走行で撮影した.

通常カメラ映像の下部は主にボンネットおよび道路が映っており, マッチングに寄与しないと考えられるため映像の上半分を用いた. また, 本実験は車両の正面方向のみを対象としたため, 全方位カメラ映像は前方の 180 [deg] の視界のみを用いた. 図7に示すように, 通常カメラ映像および全方位カメラ映像のサイズを円筒面投影によりそれぞれ 60×56 pixel, 250×144 pixel に変換して処理を行った. 初期値 t_{est} を人手で与え, 各閾値は事前の実験により決定した. 照合窓は 2 pixel ずつずらして配置した. 実験に用いたパラメータの値を表2に示す.

表 3 推定誤差.

カメラの方向	推定誤差
前方	5.1 frames (1.78 m)
左前方	1.3 frames (0.47 m)
右前方	3.1 frames (1.08 m)

3.2 結果と考察

入力した通常カメラ映像のフレームに対し、提案手法で推定された全方位カメラ映像のフレームの真値との誤差の平均は表 3 のようになった。ただし、誤差を評価する際、車両が停止中のフレームは除外する。精度はカメラが左前方を向いているときが最も良く、続いて右前方、前方の順番となった。普及型 GPS による単独測位には 5~30 m 程度の誤差があるといわれているため、いずれの結果もカーナビゲーションシステムにおける位置推定に使用するための十分な精度があるといえる。図 8 に正しく対応付けられた画像対の例をそれぞれの方向について示す。全方位カメラ画像上に表示された矩形は推定された通常カメラ画像との対応部分である。これらの画像から正しく対応がとれていることがわかる。

通常カメラ映像の最初のフレームからの全方位カメラ映像上の有効な照合窓の変化の過程を図 9 に示す。0 フレーム目では映像全体に照合窓が配置されているが、フレーム数が増えるに従って最適な照合窓に向かって急速に収束していることがわかる。また、車両の挙動による通常カメラ映像の見かけ上の変化に応じ、有効な照合窓の配置も変化していることがわかる。

図 10 に各カメラの方向における誤差の分布を示す。誤差はほぼ平均を中心として散らばっていることがわかる。しかし、前方および右前方の場合、誤差が 10 [frames] 以上の割合が突出して大きくなっている。これは対応付けができなかった場合に大きくずれた値を推定フレームとしたために発生した外れ値によるものである。前方および右前方は左前方に比べて誤差の平均が大きくなっている。カメラが前方を向いている場合、通常カメラの視野角が小さいため遠方のみを撮影することになる。遠方は映像の見かけ上の変化が乏しいため、前後のフレームとの類似性が大きくなる。したがって、対応付けの精度のばらつきが大きくなると考えられる。右前方の場合では、対向車によって対応部分が遮蔽されたため対応付けを誤ることがあった (図 11)。しかし、対向車によるオクルージョン発生の直後は対応付けを誤っているが、数フレーム後には正しい対応付けを再開していることがわかる。

処理速度は、左前方のデータを使用したときメモリに画像を読み込んだ状態から開始した場合で平均 10.6 [frames/sec] であった。なお、全方位カメラ映像データベースの探索範囲を制限し、照合窓の制限を行わなかった場合では平均 0.36 [frames/sec] であった。本実験では実時間処理が可能な速度に達しなかったものの、実装の工夫やパラメータの調整により実時間処理も可能であると考えられる。環境は CPU Core2Quad Q6600 2.40GHz, Memory 4.0GB である。

4. 結 論

本報告では、カメラ画像を利用した自車位置推定を目指し、全方位カメラ映像と通常カメラ映像のフレーム間の対応付け手法を提案した。全方位カメラ映像における双曲面ミラーの焦点の軌道と通常カメラの光学中心の軌道がほぼ等しいと仮定することで、映像を円筒面上に投影し、対応付けた。通常カメラ映像を 1 次元の系列、照合窓により切り出される全方位カメラ映像を 3 次元の系列とみなし、DP マッチングを基にした手法により、系列間の対応区間を抽出することでフレーム間の対応付けを行った。実際の車載カメラ映像を用いた評価実験により、提案手法は車両の自己位置推定のために用いるのに十分な精度があることを示した。今後の課題として、対向車によるオクルージョンや新改築による街並変化への対応など、ロバスト性を向上させることが挙げられる。

謝 辞

日頃より熱心にご討論いただく名古屋大学村瀬研究室諸氏ならびに原稿作成について御指導をいただいた石田皓之氏に感謝する。本研究の一部は日本学術振興会科学研究費補助金による。また、本研究では画像処理に MIST ライブラリ (<http://mist.suenaga.m.is.nagoya-u.ac.jp/>) を使用した。

文 献

- [1] T. K. Xia, M. Yang, R. Q. Yang, "Vision Based Global Localization for Intelligent Vehicles," Proc. 2006 IEEE Intelligent Vehicles Symposium (IV2006), pp. 571-576, June 2006.
- [2] Google Maps, <http://maps.google.com/>.
- [3] J. Sato, T. Takahashi, I. Ide, H. Murase, "Change Detection in Streetscapes from GPS Coordinated Omni-Directional Image Sequences," Proc. 18th Int. Conf. on Pattern Recognition (ICPR'06), Vol. 1, pp. 975-979, Aug. 2006.
- [4] W. Lee, K. Roh, I. Kweon, "Self-localization of a Mobile Robot without Camera Calibration using Projective Invariants," Patten Recognition Letters, Vol. 21, Issue 1, pp. 45-60, Jan. 2000.
- [5] S. Ikeda, T. Sato, K. Yamaguchi, N. Yokoya, "Construction of Feature Landmark Database using Omnidirectional Videos and GPS Positions," Proc. Int. Conf. on 3-D Digital Imaging and Modeling (3DIM), pp. 249-256, Aug. 2007.
- [6] S. Ono, T. Mikami, H. Kawasaki, K. Ikeuchi, "Space-time Analysis of Spherical Projection Image," Proc. 18th Int. Conf. on Pattern Recognition (ICPR'06), Vol. 1, pp. 975-979, Aug. 2006.
- [7] H. Kawasaki, A. Miyamoto, Y. Ohsawa, S. Ono, K. Ikeuchi, "Multiple Video Camera Calibration using EPI for City Modeling," Proc. Asian Conf. on Computer Vision, Vol. 1, pp. 569-574, Jan. 2004.
- [8] 山澤一誠, 八木康史, 谷内田正彦, "移動ロボットのナビゲーションのための全方位視覚系 HyperOmni Vision の提案," 信学論 (D-II), Vol. J79-D-II, No. 5, pp. 698-707, May 1996.
- [9] H. Sakoe, Y. Katayama, "One Dimensional-Two Dimensional Dynamic Programming Matching Algorithm for Character Recognition", IEICE Trans. Information & Systems, Vol. E77-D, No. 9, pp. 1047-1054 Sep. 1994.



図 8 対応付けの成功例。左の画像は入力された通常カメラ映像のフレームを、右の画像は推定された全方位カメラ映像のフレームを示す。推定された通常カメラ映像のフレームに対応する部分を矩形で示す。

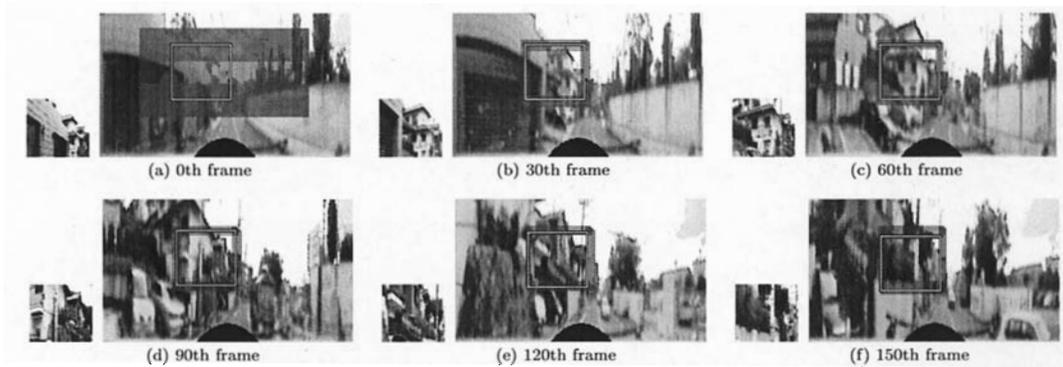


図 9 照合窓の変化。照合窓を矩形で示す。太枠の矩形は最適な照合窓を示す。各照合窓は重なり合い、少しずつずれて配置されている。

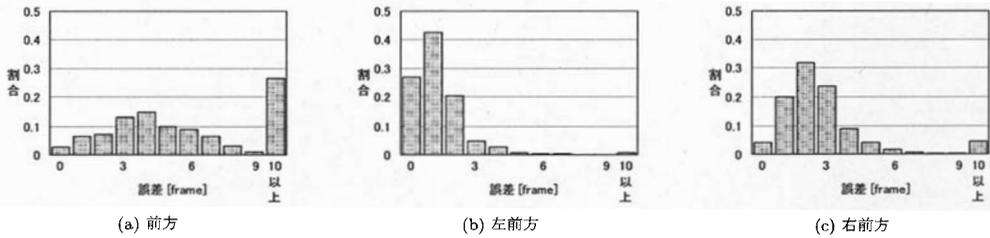


図 10 カメラ方向ごとの誤差の分布。

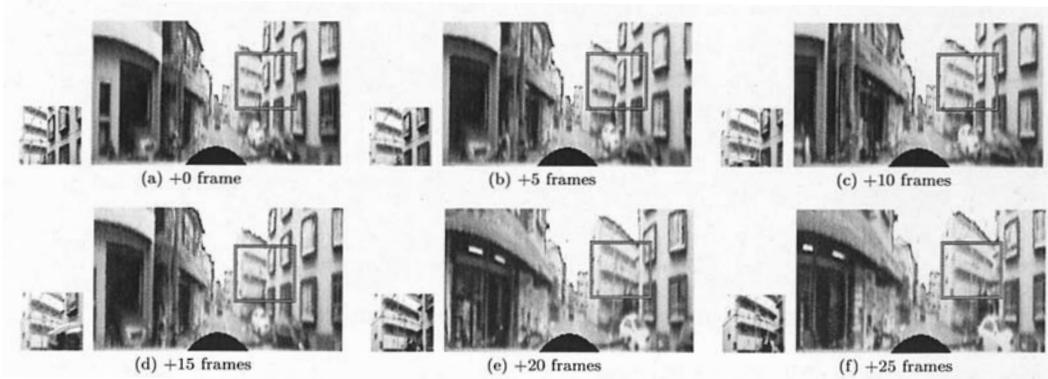


図 11 対向車両による対応付け誤りの例。(a)(b)(c)(e)(f)は正しく対応付けられているが、(d)は通常カメラ映像の下部を対向車に遮蔽されたため、対応付けに失敗した。