

画像の局所的特徴を考慮した固有値勾配法による動き推定の精度改善

金子 豊 鹿喰 善明 金次 保明 田中 豊
NHK放送技術研究所

〒157-8510 東京都世田谷区砧 1-10-11

TEL: 03-5494-2231

E-mail: kaneko@strl.nhk.or.jp

あらまし 本報告は、画像の局所的な特徴を考慮した固有値勾配法による動き推定手法を提案する。提案する手法は、始めに固有値勾配法による動き推定を行った後、信頼性の低い推定箇所を、推定に用いたブロック内の画像の特徴で分類する。推定精度を改善するために本報告では2つの固有値勾配法に基づく手法を提案する。複数の動きが含まれるブロックに対しては、ロバスト推定を適用した動き推定を行う。一方、ブロック内に1種類の勾配のみ存在すると分類されたブロックに対しては、同一エッジの画像の異なるブロック位置から推定対象の固有ベクトルを補完することで動き推定を行う。実験の結果、提案手法が動き推定の精度改善に有効であることを確認した。

キーワード 勾配法, 動き推定, 固有値, ロバスト推定

Improvement of Motion Estimation with Eigenvalue Algorithm Matching Local Image Feature

Yutaka KANEKO, Yoshiaki SHISHIKUI, Yasuaki KANATSUGU, and Yutaka TANAKA
NHK Science and Technical Research Laboratories

1-10-11, Kinuta, Setagaya-ku, Tokyo 157-8510

TEL: 03-5494-2231

E-mail: kaneko@strl.nhk.or.jp

Abstract This paper describes a motion estimation method with eigenvalue algorithm matching local image feature. The proposed method first formulates the optical flow using the block gradient method with eigenvalue algorithm, and then low reliability optical flow is classified according to the characteristics inside the block. We introduce two estimation methods based on eigenvalue algorithm for improvement in accuracy. The first method is based on robust estimation. We apply this method to the blocks which include multiple movements. The second method is applied to the blocks in which one typical gradient alone exists. In the method, we complement the eigenvector from the object with the same edge in the subjective block, and then motion estimation is carried out with the complementary eigenvector. Experimental results show that the proposed method improves the estimation accuracy.

key words gradient method, motion estimation, eigenvalue, robust estimation

1. はじめに

動画処理において、動き推定は重要な技術の1つである。特に動画の領域分割では、動き量は重要な特徴量であり、各画素ごとに精度の高い動き量が必要とされる。

動き推定の手法は大きく分けると、勾配法[1-4]と、ブロックマッチング法[5]に大別される。勾配法は推定値に精度の制限がないこと、計算量がブロックマッチング法に比べ少ないことが特徴である。

動き推定値の信頼度を利用することで推定精度を向上させる手法が報告されている[6-8]。画像ブロック内の時空間輝度勾配から動き推定を行う固有値勾配法では、計算過程で導出される固有値、および固有ベクトルを使うことで、ブロック内の画像の特徴による分類や、推定値の信頼度を求めることができる[7,9]。

本報告では局所的な画像の特徴を考慮することで固有値勾配法の推定精度を向上させる手法を提案する。提案する手法は、推定値の信頼性が低いブロックのうち、単一の勾配のみを含むブロックおよび複数の動きが含まれるブロックに対して、それぞれ固有値勾配法に基づく異なる推定手法を用いる。単一勾配のみのブロックに対しては、推定に用いる第2固有ベクトルを、同一エッジを有する画像の異なるブロック位置の固有ベクトルで補完することで動き推定を行う。一方、複数の動きが含まれるブロックに対しては、ロバスト推定を適用することで、ブロック内の主要な動きに対応する輝度勾配ベクトルのみを用いて推定を行う。

以下、2章では提案する動き推定手法を説明する。3章では実験結果を示すとともに考察を行い、4章でまとめる。

2. 提案する動き推定手法

2.1 固有値勾配法

画像ブロック内(画素数 N)の i 番目の画素位置の輝度を I_i 、時空間輝度勾配ベクトルを $\mathbf{g}_i = (\frac{\partial I_i}{\partial x}, \frac{\partial I_i}{\partial y}, \frac{\partial I_i}{\partial t})$ 、移動ベクトルを $\mathbf{m}_i = (m_x, m_y, 1)$ とする。このとき、勾配法の基本式[3]より以下の近似式が成り立つ。

$$\mathbf{g}_i \cdot \mathbf{m}_i = 0 \quad (1)$$

ここで、画像ブロック内の全ての画素が同一の動き \mathbf{m} であると仮定すると、(2)式で表されるブロック内の各画素位置の誤差 r_i の自乗和 $(\sum_{i=1}^N r_i^2)$ を最小とすることで、移動ベクトル \mathbf{m} を推定できる。

$$r_i = \mathbf{g}_i \cdot \mathbf{m} \quad (2)$$

ここで、ブロック内の時空間輝度勾配ベクトルのモーメント行列を $\mathbf{G} = \sum_{i=0}^{N-1} \mathbf{g}_i \cdot \mathbf{g}_i^T$ とし、その固有値を大きい順に $\lambda_1, \lambda_2, \lambda_3$ 、対応する固有ベクトルを $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ とする。ブロック内の自乗誤差 r_i^2 の和を最小にするのは、推定移動ベクトル $\tilde{\mathbf{m}}$ を、第3固有ベクトル $\mathbf{u}_3 = (u_{31}, u_{32}, u_{33})$ の方向としたときである[7,10]。したがって、移動ベクトル $\tilde{\mathbf{m}}$ は以下の式から求めることができる。

$$\tilde{\mathbf{m}} = \frac{1}{u_{33}} \mathbf{u}_3 \quad (3)$$

固有値勾配法は、図1に示すように、3次元空間内に分布する時空間輝度勾配から、平面を推定することと等価である。このとき、固有ベクトル \mathbf{u}_3 は推定平面の法線ベクトルを表す。また、第1主成分である固有ベクトル \mathbf{u}_1 の方向は、ブロック内の時空間輝度勾配ベクトルのうち、主要な時空間輝度勾配ベクトルの方向を表す。

2.2 局所領域の画像特徴による分類

固有値勾配法による動き推定を平面の当てはめ問題と考えると、時空間輝度勾配が平面上に分布していれば推定値は信頼性が高く、逆に平面を推定するのが困難な分布状態であれば、推定された動き量の信頼性は低くなる。そこで、文献[9]では図2に示すように、時空間輝度勾配

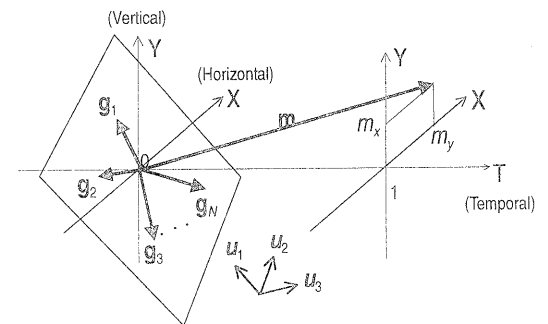


図1 輝度勾配ベクトルと移動ベクトル

ベクトルの分布状態を、信頼性の高い分布 (a)、信頼性の低い分布 (b および c) の 3 つの状態に分類した。これらの状態は、モーメント行列 G の固有値 $\lambda_1, \lambda_2, \lambda_3$ を使った以下の 2 つの指標 a_s, a_l から判別できる。

$$a_s = \frac{\lambda_3^2}{\lambda_1^2 + \lambda_2^2 + \lambda_3^2} \quad (4)$$

$$a_l = \frac{\lambda_2^2}{\lambda_1^2 + \lambda_2^2} \quad (5)$$

ここで、図 3 に示すように、勾配法による動き推定で、推定が困難なブロック内の状態を Type1~3 の 3 種類に分類する。これらの Type と図 2 で示した信頼性の低い分布状態との間には対応関係がある。Type1 の場合、特徴的な時空間輝度勾配ベクトルがないため、図 2.b の分布であり、かつ、そのノルムは小さな値となる。一方、Type2 は複数の動きが含まれるため、複数の平面上に分布する輝度勾配ベクトルが混在することになり、同じく図 2.b の分布となるが、そのノルムは大きな値となる。したがって、Type1 と Type2 との判別は、以下の指標を使って分類できる。

$$a_n = \lambda_1^2 + \lambda_2^2 + \lambda_3^2 \quad (6)$$

また、Type3 は特徴的な勾配が 1 種類であるため、分布状態は図 2.c になる。

以上から、指標 a_s, a_l, a_n を用いて、信頼性の高いブロックおよび、信頼性の低い Type1~3 のブロックの 4 種類に分類できる。

2.3 提案手法の概要

提案手法の概要を図 4 に示す。本手法は、始めに画像フレーム内の全画素に対する動き量を (3) 式の固有値勾配法を用いて推定する。次に、

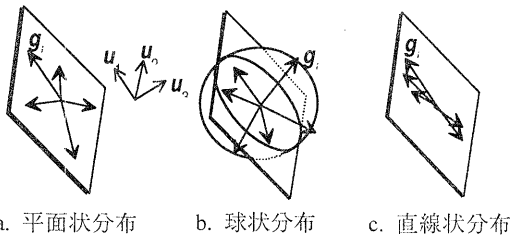


図 2 輝度勾配ベクトルの分布

動き推定の過程で導出された固有値を用いた (4)~(6) 式の指標を閾値 Th_s, Th_l, Th_n を使って、各画素位置の演算に用いたブロックを分類する。分類結果に基づき、信頼性の低い Type2 および Type3 に対して、2.4 節および 2.5 節で述べる手法を適用する。なお、Type1 に関しては、動き量を 0 として扱うことにする。

2.4 Type2 に対する推定法 (ロバスト推定)

統計的手法であるロバスト推定[12]を動き推定に適用することの有効性が報告されている[13-17]。本手法は対象ブロックに対してロバスト推定を適用し、ブロック内の主要な動きに関する時空間輝度勾配ベクトルだけを選択的に用いた固有値勾配法により動き推定を行う。ロバスト推定の手法には、筆者らが先に報告した文献[15]の手法を用いる。ただし、この手法は M-推定[12,13]に基づく動き推定手法であり、初期推定の影響を受けやすい欠点がある。そのため、初期推定としてロバスト推定の手法の一つである LMedS (Least Median of Square) 推定[12,14]を用いた。

文献[15]の手法は (8) 式で示す重み関数 w を用いて主要な動きに関連する M^k 個の時空間輝度勾配ベクトルを選択し、固有値勾配法を繰り返し演算 (繰り返し数 l) することで推定を行う手法である。

$$w_i^{(k+1)} = \begin{cases} 1: & \text{if } r_i^{(k)} < \frac{c\lambda_3^{(k)}}{N^{(k)}} \\ 0: & \text{otherwise} \end{cases} \quad (8)$$

ここで、 $*(k)$ は k 回目の演算、 c は定数を表す。

一方、初期推定に用いた LMedS 法は自乗誤差 r_i^2 のメディアン値を最小にする手法である[12,14]。

$$\min_i (\text{med } r_i^2) \quad (9)$$

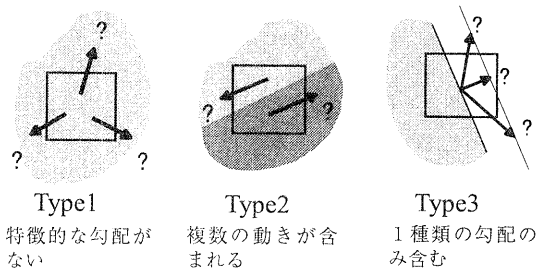


図 3 動き推定が困難な画像の特徴

- (1) 全画素位置の動きを固有値勾配法により推定する。
 (2) 各画素毎に以下の判定を行い、信頼性の低い画素位置の動き推定を行う。

```

if  $a_s < Th_s$ 
  if  $a_n < Th_n$ 
    (Type1 ブロック)
    動き推定値を 0 とする
  else
    (Type2 ブロック)
    2.4 節の手法を適用
else
  if  $a_l < Th_l$ 
    (Type3 ブロック)
    2.5 節の手法を適用
  else
    (信頼性の高いブロック)
    (1)の推定値をそのまま使う
  
```

図4 提案手法の概要

2.5 Type3 に対する推定法 (固有ベクトル補完)

Type3 ではブロック内に単一の輝度勾配成分しか存在しないため、エッジ方向に対して垂直な成分の動きのみ正確な推定ができる。これは、 $\lambda_1 \gg \lambda_2 = \lambda_3 \approx 0$ という関係から、第1固有ベクトル \mathbf{u}_1 の方向のみが信頼でき、第2固有ベクトル \mathbf{u}_2 は信頼できないベクトルとなり、平面推定が正しく行えないためである。

そこで、本手法では同一エッジをもつ画像は同一の動きであると仮定し、同一エッジを有する異なるブロック位置の固有ベクトルを、推定対象ブロックの第2固有ベクトルとして補完することにより動き推定を行う。これは、推定対象ブロックに含まれる輝度勾配と異なる種類の輝度勾配を、同一動物体の他の位置のブロック内の時空間輝度勾配ベクトルを用いて推定を行うことに相当する。

推定対象ブロックの第1固有ベクトルと異なる方向の固有ベクトルをもつブロックの探索は、図5に示すように、推定対象ブロックを中心としたエッジ方向に沿った両側に対して行う。探索は以下のいずれかの条件を満たすブロックを検出した時点で終了する。

- ① 信頼性の高いブロック (平面状分布)
- ② 推定対象ブロックの第1固有ベクトルとの角度が Th_θ よりも大きな方向成分の固有ベクトルをもつ Type 2 ブロック

①のブロックが検出された場合は、第1および第2固有ベクトルのうち、推定対象ブロックの第1固有ベクトルとのなす角が大きなベクトルを用いる。また、②のブロックが検出された場合はその第1固有ベクトルを用いる。

なお、探索するエッジの方向 θ は、時空間輝度勾配が3次元エッジ検出と等価である[11]ことから、探索ブロックの第1固有ベクトル $\mathbf{u}_1 = (u_{11}, u_{12}, u_{13})$ を $t=0$ の平面に投影したときの2次元ベクトル (u_{11}, u_{12}) に直交する方向を用いる。

$$\theta = \tan^{-1} \left(\frac{u_{11}}{u_{12}} \right) \quad (8)$$

3. 実験結果と考察

本手法の性能を検証するため実験を行った。推定値の比較として、筆者らが以前に報告した可変ブロックによる固有値勾配法[9] (以下 VEV と略記) と、文献[14]の WLS (Weighted Least Square)法を用いた。

VEV は信頼性を判定基準として、Type 1 および Type 3 に対してブロックサイズを大きく、Type2 に対してブロックサイズを小さくして動き推定を行う手法である。

WLS は初期推定として LMedS 推定を行い、その後、(10)式の重み係数により外れ値を除去し、さらに厳しい(11)式の重み係数により外れ値を除去して最小自乗法により動きを推定するロバスト動き推定手法である[12,14]。

$$w_i^0 = \begin{cases} 1: & \text{if } |r_i / s^0| \leq 2.5 \\ 0: & \text{otherwise} \end{cases} \quad (10)$$

$$s^0 = 1.4826(1 + 5/(N-2)) \sqrt{\text{med}_i r_i^2}$$

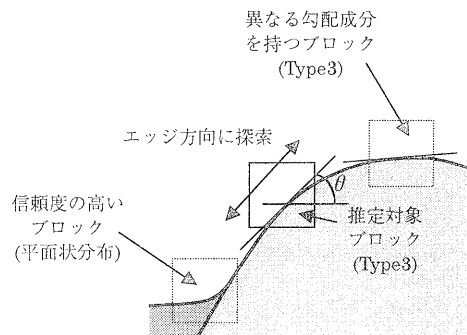


図5 ブロックの探索

$$w_i^* = \begin{cases} 1: & \text{if } |r_i/\sigma^*| \leq 2.5 \\ 0: & \text{otherwise} \end{cases} \quad (11)$$

$$\sigma^* = \sqrt{\frac{\sum_{i=1}^N w_i^0 r_i^2}{\sum_{i=1}^N w_i^0} - 2}$$

本実験では、時空間輝度勾配の計算には、全ての手法共通に Zucker の 3x3x3 のエッジ検出オペレータ [18] を用いた。動き推定は対象画素位置を中心とした 15x15 画素のブロックを用い、画像フレーム内の全画素に対して動き推定を行った。VEV の最小ブロックサイズは 3x3 画素、最大ブロックサイズは 47x47 画素とした。

提案手法および VEV でブロック内の画像の分類に用いる閾値は、文献 [9] の $Th_l = 0.01$, $Th_s = 0.00137$, $Th_n = 1000$ を用いた。提案手法のロバスト推定には文献 [15] より、繰り返し回数 $l=4$, 定数 $c=4$ を用いた。固有ベクトル補完で用いる閾値 Th_θ は予備実験から 45° とした。提案手法と WLS で用いる LMedS 推定の計算法は文献 [12] のランダムサブサンプリング法を用いた。繰り返し回数は 30 回とした。

図 6 に実験結果を示す。原画像 (a.1) [14] は中心の矩形領域が止まり、周囲が右上に動いている画像であり、(b.1) [14] はテーブルが回転している画像、(c.1) [19] は移動している人物をカメラがフォローしている画像である。図 (a.2), (b.2), (c.2) は分類結果である。

図 6 (a) の結果では、VEV は動きの境界付近で誤推定が生じている。これは境界部分では推定に用いるブロックサイズが小さくなりすぎるためである。一方、ロバスト推定を用いる WLS および提案手法では動きの境界部分においても正しく推定が行われている。図 6 (b) の結果では、回転するテーブルの部分がブロック内に 1 種類の勾配のみ含まれる Type3 となるため、WLS では推定が行えない部分があるのに対し、提案手法では多くの部分で正しい推定結果が得られた。また、図 6 (c) の結果からも Type3 の領域に対して提案手法が有効であることが分かる。

次に、各手法の演算時間を表 1 に示す。提案手法が他の手法より演算量が少ない結果となった。これは、WLS が全ての画素位置に対して演算コストの高いロバスト推定を行うのに対し、提案手法はロバスト推定が有効に作用する信頼性の低い Type3 の部分に対してのみ演算を行

っているためである。また、VEV はブロックサイズを変えながら固有値勾配法の演算を繰り返し行うため、信頼性の低い領域が多い画像では演算量が増加するのに対し、提案手法では Type3 の領域に対して行う固有ベクトル補完処理は単純な探索を行うだけなので、演算量が少なくてすむためである。

表 1 演算時間の比較 (sec)

	VEV	WLS	提案
New-Sinusiod1	1	35	13
Rubik	51	348	22
Japanese Room	517	1293	187

CPU: Alpha 21164A (466MHz)

4. まとめ

画像の局所的な特徴を考慮した固有値勾配法の推定精度を向上させる手法を提案した。提案手法は、始めに固有値勾配法の動き推定の導出過程で求められる固有値を使った指標により、信頼性の低いブロックを画像の特徴で分類する。ブロック内に複数の動きが含まれるブロックに対してはロバスト推定を適用することで、ブロック内の主要な動きに関連する輝度勾配のみを用いて動き推定を行う。一方、特徴的な勾配が一種類しか含まないブロックに対しては、同一エッジを有する画像の異なるブロック位置の固有ベクトルで推定対象ブロックの第 2 固有ベクトルを補完することで推定を行う。

実験の結果、提案手法は複数の動きを含むブロックおよび、単一の勾配のみを含むブロックという信頼性の低いブロックに対して推定精度を改善できることを確認した。また、提案手法は、推定精度の低い部分にのみ計算コストの高いロバスト推定を用いるため、少ない演算量で効果的に精度を改善できることが分かった。

謝辞

本研究を行うにあたり御助言を頂いた神奈川大学工学部齋藤隆弘教授に感謝いたします。

参考文献

- [1] C. Cafforio, F. Rocca, "Methods for Measuring Small Displacement of Television Images," IEEE Trans. Information Theory, IT-22, 5, pp.573-579, 1976
- [2] 吹抜, "画像信号による動対象の移動量, 速度の測定," 信学技報, IE78-67, 1978
- [3] B. K. P. Horn, B. G. Schunck, "Determining Optical Flow," Artificial Intell., vol.17, pp.185-203, 1981
- [4] J. L. Barron, D. J. Fleet, S. S. Beauchemin, "Systems and Experiment Performance of Optical Flow Techniques," Intern. J. Comput. Vis. 12:1, pp.43-77,

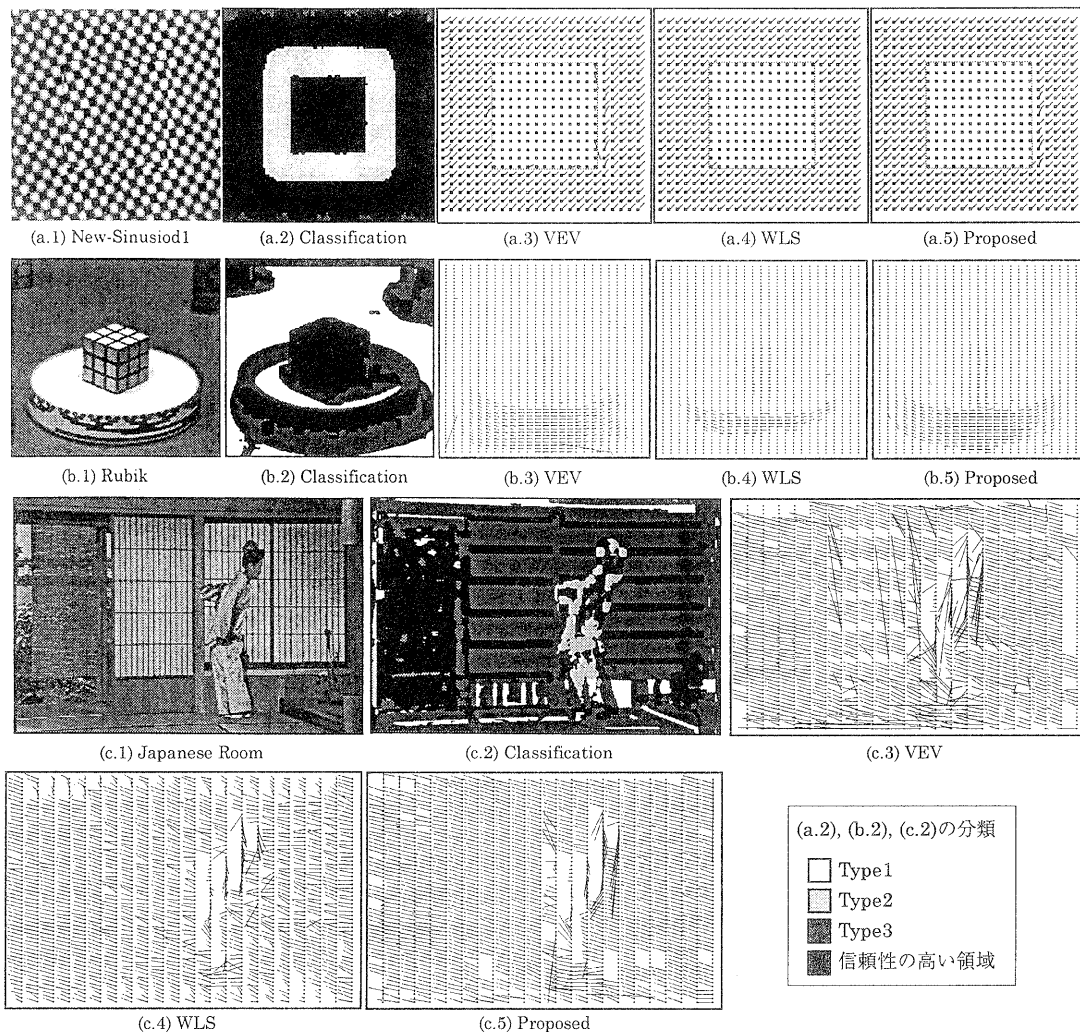


図6 実験結果

- 1994
- [5] H. Gharavi, M. Mills, "Block Matching Motion Estimation Algorithms - New Results," IEEE Trans. Circuits and Systems, vol.37, pp.649-651, 1990
- [6] N. Ohta, "Image Movement Detection with Reliability Indices," IEICE Trans., vol.E74, No.10, pp.3379-3388, 1991
- [7] Y. Kaneko, Y. Shishikui, Y. Tanaka, "Improvement of Motion Estimation Accuracy using Gradient Vector Distribution in Advanced Picture Coding," IEEE Trans. on Broadcasting, vol.42, No.3, pp.237-244, 1996
- [8] 吉田, 宮本, 酒井, "動画像の動きベクトルに対する信頼度関数とその応用," 信学論, Vol.J80-D-II, No.5, pp.1192-1201, 1997
- [9] Y. Kaneko, Y. Shishikui, Y. Nojiri, Y. Tanaka, "Advanced Motion Estimation using Eigenvalue Algorithm with Optimized Block Size," proc. Int. Workshop on Very Low Bitrate Video, VLVB98, 5-7, Oct., 1998
- [10] 金谷, "画像理解-3次元認識の数理-, " p.1-27, 1990
- [11] 山本, "動画像と距離情報の併用による3次元運動パラメータの直接的推定法," 信学論, Vol.J68-D, No.4, pp.562-569, 1985
- [12] P. J. Rousseeuw, A. M. Leroy, "Robust Regression and Outlier Detection," John Wiley & Sons, 1987
- [13] M. J. Black, P. Anandan, "A Framework for the Robust Estimation of Optical Flow," proc. Int. Conf. on Computer Vision, ICCV-93, pp.231-236, 1993
- [14] A. B. Hadiashar, D. Suter, "Robust Optical Flow Computation," Intern. J. Comput. Vis. 29(1), pp.59-77, 1998
- [15] 金子, 鹿喰, 金次, "ロバスト推定を適用した固有値勾配法による動き推定の検討," 信学ソ大, D-11-1, 1998
- [16] E.P. ONG, M. Spann, "Robust Optical Flow Computation Based on Least-Median-of-Squares Regression," Intern. J. Comput. Vis. 31(1), pp.51-82, 1999
- [17] 伊藤, 小松, 齋藤, "ロバスト動き推定法とその評価," 映像年次大, 22-5, 1999
- [18] S. W. Zucker, R. A. Hummel, "A Tree-Dimensional Edge Operator," IEEE Trans. PAMI, Vol.PAMI-3, No.3, pp.324-331, 1981
- [19] 映像情報メディア学会, "標準テレビジョン方式・システム評価用標準動画像"