

# 距離情報と動き情報の統合による3次元シーンの復元<sup>1</sup>

5M-4

阿部 健一 長橋 宏<sup>2</sup>

東京工業大学 大学院総合理工学研究科<sup>3</sup>

## 1 はじめに

通常的环境下で撮影された画像から、3次元シーンの奥行き情報を精度良く抽出することは重要であり、古くから多くの研究がなされている。本稿では、ステレオカメラを動かして撮影された4枚の画像から視差とオプティカルフローを同時に推定して精度の良いマッチングを求める手法を検討する。

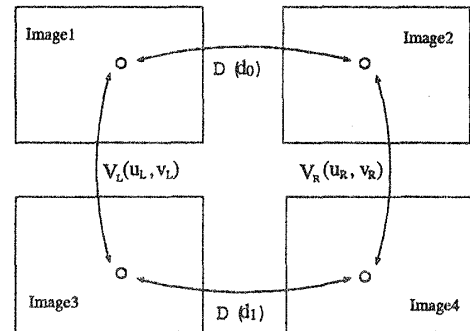


図1: 動きとステレオ対応

## 2 ステレオ情報と動き情報の統合

### 2.1 動きと視差の整合性

時系列のステレオ画像における、動き情報とステレオ情報の対応を図1に示す。

図1の4枚の画像の中で、Image1とImage2、Image3とImage4はステレオ画像のペアで、それぞれ視差  $d_0, d_1$  を持つ。Image1とImage3、Image2とImage4は動画像のペアで、それぞれ動きベクトル  $(u_L, v_L), (u_R, v_R)$  を持つ。

この4つの情報がそれぞれつじつまの合うマッチングを与える必要がある。本稿では、水平ベースラインステレオを仮定しているため、次の2つの条件が必要となる。

$$v_L = v_R \quad (1)$$

$$u_R - u_L = d_1 - d_0 \quad (2)$$

### 2.2 ステレオマッチングとオプティカルフロー

図1のステレオマッチング  $D(d_0), D(d_1)$  と動きベクトル  $V_L(u_L, v_L), V_R(u_R, v_R)$  は、いずれも時空間勾配法を用いて求める。それぞれの時空間勾配式  $D(d_0), D(d_1), V_L(u_L, v_L), V_R(u_R, v_R)$  と、滑らかさの拘束条件、式1、式2を全て含んだエネルギー関数を定義し、それを最小化する。

このようなエネルギー関数の式  $E$  は、以下のような滑らかさのエネルギー  $E_s$  と勾配法のエネルギー  $E_c$  を用いて、

$$E_s(V) = \int \int (u_x^2 + u_y^2 + v_x^2 + v_y^2) dx dy \quad (3)$$

$$E_c(V) = \int \int (I_x u + I_y v + I_t)^2 dx dy \quad (4)$$

$$E = E_c(D_0) + E_c(D_1) + E_c(V_L) + E_c(V_R) + \alpha \{E_s(D_0) + E_s(D_1)\} + \beta \{E_s(V_L) + E_s(V_R)\} + \gamma \{(v_l - v_r)^2 + (u_1 - u_2 - d_1 + d_2)^2\} \quad (5)$$

となる。

<sup>1</sup>3D scene reconstruction by integration of motion and stereo information

<sup>2</sup>Kenichi Abe Hiroshi Nagahashi

<sup>3</sup>Interdisciplinary Graduate School of Science and Engineering, Tokyo Institute of Technology

### 3 M 推定法を用いたエネルギー関数の最小化

本節では、複数のパラメータを同時に扱うことによる誤差の伝播に対応するため、最適化の段階でロバスト推定の一手法である M 推定法 [2] を導入する。

#### 3.1 M 推定法

$i$  番目のデータから推定値までの残差を  $\varepsilon_i$  とする。通常の最小自乗法は、 $\sum_i \varepsilon_i^2$  を最小化するもので、外れ値の影響を受けやすい。M 推定法は、外れ値の影響を減らすため、次の関数  $\rho(\varepsilon)$  の和の最小化を行う。

$$\min \sum_i \rho(\varepsilon) \quad (6)$$

これにより、誤差の大きなもの程小さな重みを付加することになり、外れ値が除去される。

#### 3.2 エネルギー最小化によるマッチング

まず、勾配法で大きな動きに対応するために、解像度の異なる画像を事前に用意し、各階層画像に対してエネルギー関数の最適化を行う [3]。上記の M 推定法を用いて、2 組のステレオマッチングとオプティカルフローを全て含んだエネルギー関数を最小化する。

最適化するエネルギー関数は、

$$\begin{aligned} E_s(V) &= \iint \rho(I_x u + I_y v + I_t) dx dy \\ E_c(V) &= \iint \{\rho(u_x, u_y) + \rho(v_x, v_y)\} dx dy \end{aligned} \quad (7)$$

として、

$$\begin{aligned} E &= E_c(D_0) + E_c(D_1) + E_c(V_L) + E_c(V_R) \\ &+ \alpha \{E_s(D_0) + E_s(D_1)\} \\ &+ \beta \{E_s(V_L) + E_s(V_R)\} \\ &+ \gamma \{(v_L - v_R)^2 + (u_L - u_R - d_1 + d_2)^2\} \end{aligned} \quad (8)$$

と表される。

ここで、 $\rho$  関数には、Geman-McClure の式

$$\rho(x) = \frac{x^2/2}{1+x^2} \quad (9)$$

を用いる。

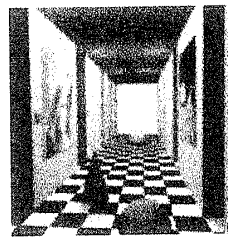


Image1

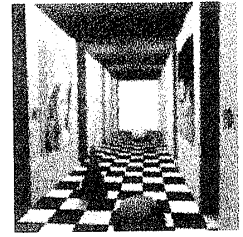


Image2

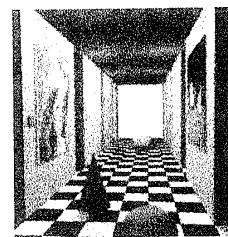


Image3

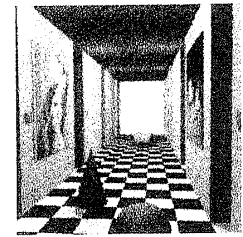


Image4

上に示す4つの画像は、それぞれ、サイズ  $192 \times 192$  の合成画像である。Image1, Image2 の水平ベースラインステレオ画像にアフィン変換を施して、奥行き方向の移動を表現したものが Image3, Image4 である。

これらの画像について、現在シミュレーションを実行中である。

## 5 おわりに

距離情報と動き情報を同時に用いて、3次元シーンの奥行き情報を復元する手法を示した。

## 参考文献

- [1] H.Baker and R.Bolles, "Realtime stereo and motion integration for navigation," Proc. of Image Understanding Workshop, pp.1295-1304, 1994.
- [2] 梅山 信二, "ロバスト統計の考え方をを用いた正則化における不連続の検出について", 信学技報, PRMU96-108, 1996.
- [3] 佐藤誠, 佐々木宏, "動画像における動きベクトルの階層的推定法", 信学論, '86/5 Vol.J69-D No.5.