

位相限定相関法に基づく高品質背景スプライト生成

沼 徳仁[†] 青木 孝文[†] 近藤 敏志^{††}

[†] 東北大学大学院情報科学研究科, 〒 980-8579 仙台市青葉区荒巻字青葉 6-6-05

^{††} 松下電器産業株式会社, 〒 571-8501 門真市大字門真 1006

あらまし 本稿では、位相限定相関法 (Phase-Only Correlation: POC) に基づく高品質な背景スプライト生成手法について述べる。背景スプライト生成は、複数枚の動画像フレームを張り合わせることで動画像の背景領域を生成する処理である。本稿では、POC に基づくサブピクセル精度の対応点探索手法によって得られた対応点組とその POC の相関値を用いた背景スプライト生成手法を提案する。生成した背景スプライトから切り出したフレームと原画像の背景領域とを PSNR を用いて評価した結果、提案手法により高品質な背景スプライトが得られることを示す。

キーワード スプライト生成, ビデオモザイク, 位相限定相関法, グローバルモーション

High-Quality Background Sprite Generation Based on Phase-Only Correlation

Norihito NUMA[†], Takafumi AOKI[†], and Satoshi KONDO^{††}

[†] Graduate School of Information Sciences, Tohoku Univ., 6-6-05, Aramaki Aza Aoba, Aoba-ku, Sendai-shi, 980-8579 Japan

^{††} Matsushita Electric Industrial Co., Ltd., 1006 Kadoma, Kadoma-shi, 571-8501 Japan

Abstract A background sprite is an image composed of pixels belonging to a video object visible throughout a video segment. This paper proposes a high-quality background sprite generation method using phase-based image matching. The key issues in creating a high-quality background sprite are accurate global motion estimation and foreground object removal. Our proposed method employs (i) a phase-based sub-pixel motion estimation technique using the Phase-Only Correlation (POC) function, and (ii) a foreground object masking technique using the peak value of the POC function as a reliability measure for motion estimation. Experimental observation shows that the proposed method can generate high-quality background sprites for many practical video sequences, where the image quality is evaluated by the PSNR between the reconstructed background image and the original one.

Key words sprite generation, video mosaicing, phase-only correlation, global motion

1. ま え が き

近年、動画像のフレームを張り合わせて背景の大きな静止画像 (背景スプライト) を得る背景スプライト生成やビデオモザイクに関する研究が盛んに行われている [1]~[6]。背景スプライト生成は、動画像処理における基本的な処理の 1 つであり、動画像符号化や動画像編集などに使われている。たとえば、生成された背景スプライトを利用することで、動画像の各フレームを前景オブジェクトと背景スプライトに分離し、効率的に動画像を圧縮することができる [3]。また、高臨場感な動画像の生成、オブジェクトの抽出 [4]、動画像理解、画質向上、カメラモーションの安定化などの動画像編集に利用することができる。

背景スプライトは、各フレームから前景オブジェクトを取り除いた背景画像を使ってカメラの動きによるフレーム間の画像変形を推定し、変形させたフレームを重ね合わせることで生成される。一般的には、フレーム間の動きベクトルを推定し、統計的にカメラの動きに一致しない動きベクトルを前景オブジェクトとして取り除く手法が用いられている [3], [5]。そのため、高品質な背景スプライトを得るためには、高精度な動きベクトル推定や高精度な前景オブジェクトの除去が重要となる。

筆者らの研究グループでは、位相限定相関法 (Phase-Only Correlation: POC) [7] を用いた高精度な対応点探索・移動量推定手法を開発してきた [8]~[10]。2枚の画像間の POC 関数を計算することで、画像間の移動量と類似度を相関関数のピークの座標と高さとして求めることができる。この手法を用いる

ことで、たとえば、画像サイズが 101×101 ピクセルのときに $1/100$ ピクセル精度で移動量を推定することができる [8]。

また、上記手法を局所的な画像ブロックのマッチングに適用し、coarse-to-fine 探索とサブピクセル窓あわせ法を組み合わせることにより、2枚の画像の対応点および動きベクトルをサブピクセル精度で検出できる [9], [11]。サイズが 11×11 の画像ブロックを用いて約 $1/20$ ピクセル精度の対応付けが可能である。

本稿では、位相限定相関法に基づいた高品質な背景スプライト生成手法を提案する。提案する手法は、(i) POC に基づくサブピクセル精度の対応点探索手法 [9] と、(ii) 対応点組の相関値を動きベクトルの信頼性として利用した前景オブジェクトの除去手法を用いた手法である。動画像を使った実験により、生成した背景スプライトから切り出したフレームと原画像の背景領域とを PSNR (Peak Signal-to-Noise Ratio) を用いて評価し、提案手法を使うことで高品質な背景スプライトが得られることを示す。

2. 位相限定相関法に基づく動き推定

2.1 位相限定相関法

位相限定相関法 (Phase-Only Correlation: POC) の定義を示す。画像サイズ $N_1 \times N_2$ ピクセルの2つの画像を $f(n_1, n_2)$, $g(n_1, n_2)$ とする。定式化の便宜上、離散空間のインデックスを $n_1 = -M_1, \dots, M_1$, $n_2 = -M_2, \dots, M_2$ とし、画像サイズを $N_1 = 2M_1 + 1$ ピクセル, $N_2 = 2M_2 + 1$ ピクセルとする。これらの画像の2次元離散フーリエ変換 (2D DFT) をそれぞれ $F(k_1, k_2)$, $G(k_1, k_2)$ として次式で与える。

$$\begin{aligned} F(k_1, k_2) &= \sum_{n_1 n_2} f(n_1, n_2) W_{N_1}^{k_1 n_1} W_{N_2}^{k_2 n_2} \\ &= A_F(k_1, k_2) e^{j\theta_F(k_1, k_2)} \end{aligned} \quad (1)$$

$$\begin{aligned} G(k_1, k_2) &= \sum_{n_1 n_2} g(n_1, n_2) W_{N_1}^{k_1 n_1} W_{N_2}^{k_2 n_2} \\ &= A_G(k_1, k_2) e^{j\theta_G(k_1, k_2)} \end{aligned} \quad (2)$$

ここで、 $k_1 = -M_1, \dots, M_1$, $k_2 = -M_2, \dots, M_2$, $W_{N_1} = e^{-j\frac{2\pi}{N_1} k_1 n_1}$, $W_{N_2} = e^{-j\frac{2\pi}{N_2} k_2 n_2}$ であり、 $\sum_{n_1 n_2}$ は、 $\sum_{n_1=-M_1}^{M_1} \sum_{n_2=-M_2}^{M_2}$ である。 $A_F(k_1, k_2)$, $A_G(k_1, k_2)$ は振幅成分であり、 $e^{j\theta_F(k_1, k_2)}$, $e^{j\theta_G(k_1, k_2)}$ は位相成分である。

合成位相スペクトル $\hat{R}(k_1, k_2)$ は次のように定義される。

$$\begin{aligned} \hat{R}(k_1, k_2) &= \frac{F(k_1, k_2) \overline{G(k_1, k_2)}}{|F(k_1, k_2) \overline{G(k_1, k_2)}|} \\ &= e^{j\theta(k_1, k_2)} \end{aligned} \quad (3)$$

ここで $\overline{G(k_1, k_2)}$ は $G(k_1, k_2)$ の複素共役を示す。また、 $\theta(k_1, k_2) = \theta_F(k_1, k_2) - \theta_G(k_1, k_2)$ である。POC 関数 $\hat{r}(n_1, n_2)$ は $\hat{R}(k_1, k_2)$ の2次元離散フーリエ逆変換 (2D IDFT) であり、次式で定義される。

$$\begin{aligned} \hat{r}(n_1, n_2) &= \frac{1}{N_1 N_2} \sum_{k_1 k_2} \hat{R}(k_1, k_2) W_{N_1}^{-k_1 n_1} W_{N_2}^{-k_2 n_2} \end{aligned} \quad (4)$$

ここで、 $\sum_{k_1 k_2}$ は、 $\sum_{k_1=-M_1}^{M_1} \sum_{k_2=-M_2}^{M_2}$ である。

連続空間で定義された2次元画像 $f_c(x_1, x_2)$ を考える。ここで、 x_1, x_2 は実数である。 δ_1, δ_2 をそれぞれ x_1, x_2 方向についての $f_c(x_1, x_2)$ の微小な移動量をあらわす実数とすると、微小にずれた画像は $f_c(x_1 - \delta_1, x_2 - \delta_2)$ と表現できる。 $f_c(x_1, x_2)$, $f_c(x_1 - \delta_1, x_2 - \delta_2)$ を標準化間隔 T_1, T_2 で標準化した2次元画像をそれぞれ $f(n_1, n_2)$, $g(n_1, n_2)$ とおき、次式で定義する。

$$\begin{aligned} f(n_1, n_2) &= f_c(x_1, x_2)|_{x_1=n_1 T_1, x_2=n_2 T_2} \\ g(n_1, n_2) &= f_c(x_1 - \delta_1, x_2 - \delta_2)|_{x_1=n_1 T_1, x_2=n_2 T_2} \end{aligned}$$

ここで、 $n_1 = -M_1, \dots, M_1$, $n_2 = -M_2, \dots, M_2$ である。このとき、 $f(n_1, n_2)$, $g(n_1, n_2)$ に関する POC 関数 $\hat{r}(n_1, n_2)$ は次式で与えられる。

$$\begin{aligned} \hat{r}(n_1, n_2) &\approx \frac{\alpha}{N_1 N_2} \frac{\sin\{\frac{\pi}{N_1}(n_1 + \delta_1)\}}{\sin\{\frac{\pi}{N_1} n_1\}} \frac{\sin\{\frac{\pi}{N_2}(n_2 + \delta_2)\}}{\sin\{\frac{\pi}{N_2} n_2\}} \end{aligned} \quad (5)$$

ここで、 $\alpha \leq 1$ である。上式は、画像間に微小移動量 δ_1 と δ_2 がある場合の POC 関数の一般形を表している。相関ピークの座標は画像間の移動量を表し、相関ピークの高さ α は画像間の類似度の指標となる。画像の微小変化によって α の値が変化する。このように微小に位置ずれした2枚の同一の画像の POC 関数は1ピクセルの幅の急峻なピークを持つため、画像マッチングにおけるロバスト性と移動量推定の精度がきわめて高い。

POC 関数を用いた画像間のサブピクセル精度の移動量推定には、以下の高精度化手法が重要である [8]。2枚の画像間の POC 関数 (式 (4)) は、離散点のみで構成されるデータである。このとき、計算されたデータに相関ピークモデル (式 (5)) をフィッティングすることで、サブピクセル精度の移動量 δ_1, δ_2 と相関ピーク値 α を推定することができる。図1は、真の相関ピーク座標と相関ピーク値を推定するために、式 (5) で与えられる相関ピークモデルをフィッティングした例である。その際に、移動量と相関ピーク値の推定精度を向上させるために、(i) 窓関数による画像端の不連続性の除去、(ii) 低域通過型のスペクトル重み付け関数の適用による信頼性の低い高周波成分の除去などが重要である。実験においては、窓関数としてハンギング窓を適用し、ガウス関数型のスペクトル重み付けを用いた。

2.2 POC に基づく対応点探索

POC に基づく対応点探索は、画像ピラミッドを使った coarse-to-fine 探索である [9]。最も疎な画像から密な画像へ徐々に探索を進める。各階層で検出した移動量は、次のより密な画像階層での探索に伝播する。そして、最終的に得られた移動量から動きベクトルを求める。この手法の概略を図2に示し、処理手順を以下に示す。

処理手順

入力:

基準画像 $I(n_1, n_2)$ ($= I_0(n_1, n_2)$ とする)

参照画像 $J(n_1, n_2)$ ($= J_0(n_1, n_2)$ とする)

基準画像 $I(n_1, n_2)$ 内の探索点座標 \mathbf{p} ($= \mathbf{p}_0$ とする)

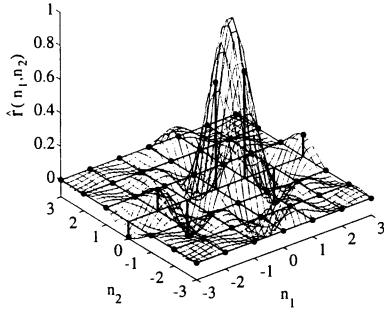


図1 関数フィッティング

出力:

探索点 p の動きベクトル v_p

対応点組 p, q の相関値 α_p

ステップ 1: $l = 1, 2, \dots, l_{max}$ の階層において, $I_0(n_1, n_2)$, $J_0(n_1, n_2)$ の縮小画像を以下のように作る.

$$I_l(n_1, n_2) = \frac{1}{4} \sum_{i_1=0}^1 \sum_{i_2=0}^1 I_{l-1}(2n_1 + i_1, 2n_2 + i_2)$$

$$J_l(n_1, n_2) = \frac{1}{4} \sum_{i_1=0}^1 \sum_{i_2=0}^1 J_{l-1}(2n_1 + i_1, 2n_2 + i_2)$$

実験においては, $l_{max} = 3$ とした. 一般的には, この値は, 動きベクトルの大きさに依存して決定する必要がある.

ステップ 2: $l = 1, 2, \dots, l_{max}$ の各階層において, 探索点座標 p_0 に対応する座標 $p_l = (p_{l1}, p_{l2})$ を以下のように求める.

$$p_l = \lfloor \frac{1}{2} p_{l-1} \rfloor = (\lfloor \frac{1}{2} p_{l-1,1} \rfloor, \lfloor \frac{1}{2} p_{l-1,2} \rfloor) \quad (6)$$

ステップ 3: $q_{l_{max}} = p_{l_{max}}$, $l = l_{max} - 1$ とする.

ステップ 4: l 層目の画像 $I_l(n_1, n_2)$, $J_l(n_1, n_2)$ から, それぞれ中心座標 p_l , $2q_{l+1}$, 画像サイズ $W \times W$ ピクセルの画像ブロック $f_l(n_1, n_2)$, $g_l(n_1, n_2)$ を取り出す.

ステップ 5: $f_l(n_1, n_2)$, $g_l(n_1, n_2)$ の間の移動量を POC により推定する (ピクセル精度で推定). 推定した移動量を δ_l とおくと, l 層目の対応点座標 q_l は以下のように決まる.

$$q_l = 2q_{l+1} + \delta_l \quad (7)$$

ステップ 6: $l = l - 1$ として $l \geq 0$ となるまで, ステップ 4 からステップ 6 を繰り返す.

ステップ 7: q_0 を考慮して切り出し位置を正規化したうえで, 画像ブロック $g(n_1, n_2)$ を取り出し, $f(n_1, n_2)$ ($= f_0(n_1, n_2)$) との POC 関数を計算する. 計算した POC 関数に対して関数フィッティングを行い, 移動量をサブピクセル精度で推定する. 推定した移動量を δ とおくと, 対応点座標 q は以下のように決まる.

$$q = q_0 + \delta \quad (8)$$

また, このときの相関ピーク値を相関値 α_p とする.

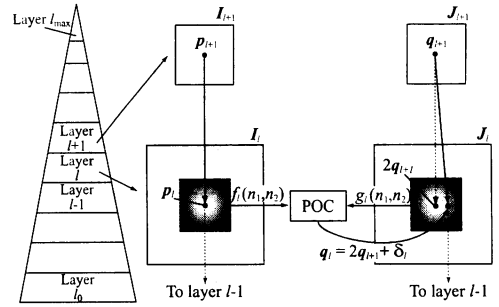


図2 POCに基づく階層探索の概要

ステップ 8: 動きベクトル $v_p = q - p_0$ を求める.

高精度なマッチングには, ある程度の画像ブロックサイズが必要となる. 実験では 32×32 ピクセルの画像ブロックを用いた.

3. POCに基づく背景スプライト生成

本稿で提案する POC に基づく背景スプライト生成について述べる.

背景スプライトは, 各フレームから前景オブジェクトを取り除いた背景画像を使ってカメラの動きによるフレーム間の画像変形を推定し, 変形させたフレームを重ね合わせることで生成される. 提案手法では, カメラの動きによるフレーム間の画像変形を 8 パラメータの平面射影変換で表されると仮定する. このとき, フレーム t からフレーム $t+1$ への射影変換行列 $H_{t \rightarrow t+1}$ は次式で表される.

$$H_{t \rightarrow t+1} = \begin{pmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & 1 \end{pmatrix} \quad (9)$$

また, フレーム間の対応点組 (p_x, p_y) , (q_x, q_y) の関係は次式で表される.

$$q_x = \frac{h_1 p_x + h_2 p_y + h_3}{h_7 p_x + h_8 p_y + 1}, \quad q_y = \frac{h_4 p_x + h_5 p_y + h_6}{h_7 p_x + h_8 p_y + 1} \quad (10)$$

本稿では, 高品質な背景スプライトを生成するために, POC に基づく動き推定を使った変換行列の推定手法と画像を重ね合わせる際の重み係数の決定法を提案する.

3.1 背景スプライト生成

提案する背景スプライト生成の処理手順を示す (図 3).

処理手順

入力:

入力画像シーケンス F_t ($t = 1, \dots, n$)

出力:

背景スプライト S ($= S_n$)

各フレームの変換行列 $H_{1 \rightarrow t}$

ステップ 1: 開始フレームを初期スプライトとする ($S_1 = F_1$). また, 入力フレーム番号 t と参照フレームの番号 t_{ref} を $t = t_{ref} = 1$, 射影変換行列を $H_{1 \rightarrow 1} = I$ とする. ただし, I は 3×3 の単位行列である.

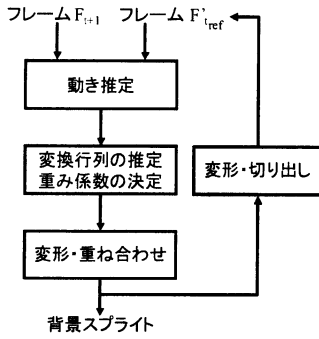


図3 背景スプライト生成の処理フロー

ステップ 2: 背景スプライト S_t からフレーム $F_{t_{ref}}$ に対応する領域を変換行列 $H_{1 \rightarrow t_{ref}}$ を用いて変形し、参照フレーム $F'_{t_{ref}}$ を切り出す。

ステップ 3: 切り出した参照フレーム $F'_{t_{ref}}$ から入力フレーム F_{t+1} への変換行列 $H_{t_{ref} \rightarrow t+1}$ を求める。また、背景スプライトに重ね合わせる際の各画素に対する重み w を求める。

ステップ 4: 変換行列 $H_{1 \rightarrow t+1} = H_{t_{ref} \rightarrow t+1} H_{1 \rightarrow t_{ref}}$ を用いて入力フレーム F_{t+1} を変形し、 w で重み付けをして背景スプライト S_t に重ね合わせる。

ステップ 5: 参照フレーム $F'_{t_{ref}}$ と入力フレーム F_{t+1} の間の画像全体の平行移動量が 1 ピクセル以上の場合には、 $t_{ref} = t$ とする。それ以外の場合 (カメラの動きが小さい場合) は t_{ref} を更新しない。

ステップ 6: $t = t + 1$ として、 $t > n$ となるまでステップ 2 からステップ 6 を繰り返す。

3.2 変換行列の推定

POC に基づいた動き推定を使った変換行列推定手法について述べる。フレーム間の背景領域の変形は射影変形を仮定しているが、局所領域では平行移動モデルのレジストレーション手法で対応できる。そこで、POC に基づく動き推定によりフレーム間の対応点組を求め、射影変換行列を推定する。また、高精度な射影変換行列を得るために、対応点組の相関値を用いる。一般的に、オブジェクトの境界領域では対応点組の相関値が小さくなる傾向にある。このような領域の対応点組は、フレーム間の背景領域の変形を表す変換行列の推定において、必要のない局所領域の情報である。そこで、射影変換行列を推定する際には相関値の低い対応点組を除去する。

変換行列 $H_{t_{ref} \rightarrow t+1}$ の推定手順を示す。

ステップ 1: フレーム F_{t+1} を基準画像、フレーム $F'_{t_{ref}}$ を参照画像として、POC に基づく動き推定によりフレーム $F'_{t_{ref}}$ におけるフレーム F_{t+1} の対応点を求める。このとき、フレーム F_{t+1} 上にある点群を $P = \{p_1, p_2, \dots, p_N\}$ 、 $F'_{t_{ref}}$ における対応点群を $Q = \{q_1, q_2, \dots, q_N\}$ 、各対応点組の相関値を $\alpha = \{\alpha_1, \alpha_2, \dots, \alpha_N\}$ とする。

ステップ 2: $\alpha \leq 0.3$ となる P と Q の対応点組を除去する。

ステップ 3: $\alpha \geq \mu(\alpha) - \sigma(\alpha)$ となる P と Q の対応点組から最小 2 乗法を使って変換行列 $H'_{t_{ref} \rightarrow t+1}$ を推定する。ここで、

$\mu(\alpha), \sigma(\alpha)$ は、それぞれ各対応点組の相関値の平均値と標準偏差である。

ステップ 4: 推定した変換行列 $H_{t_{ref} \rightarrow t+1}$ を用いて、フレーム F_{t+1} 上の点群 P をフレーム $F'_{t_{ref}}$ に射影し、対応点の計算値 Q^{calc} を求める。

ステップ 5: $|q_i - q_i^{calc}| > |q_i - p_i|/2$ となる P と Q 対応点組を一時的に除去する。

ステップ 6: 一時的に除去する対応点組が変化しなくなるまで、ステップ 3 からステップ 5 を繰り返す。

3.3 背景スプライトへの重ね合わせ

オブジェクトマスクを必要としない重ね合わせ手法を提案する。高品質な背景スプライトを得るためには、フレーム F_{t+1} を背景スプライト S_t に重ね合わせる時に、オブジェクト領域を重ね合わせないことが望ましい。そこで、オブジェクトの境界領域で対応点組の相関値が小さくなることを利用し、相関値を基準とした重みをつけて加重平均により背景スプライトへ重ね合わせる。変換行列を推定する際に求めた計算値 Q^{calc} 、フレーム $F'_{t_{ref}}$ における対応点群 Q 、相関値 α を使用すると、加重平均の重み係数は以下の処理によって決定される。

ステップ 1: 各点 $p_i \in P$ における重み係数 w_i を以下の操作によって求める。

if $|q_i - q_i^{calc}| \leq 1/2$ then

$$w_{p_i} = \alpha_{p_i}$$

else

フレーム F_{t+1} の p_i 近傍 (16×16 ピクセル) での輝度値の標準偏差 σ_{p_i} を計算する。

フレーム $F'_{t_{ref}}$ の q_i 近傍 (16×16 ピクセル) での輝度値の標準偏差 σ_{q_i} を計算する。

if $\sigma_{p_i} \leq Th$ かつ $\sigma_{q_i} \leq Th$ かつ $\alpha_i \leq 0.5$ then

$$w_i = 1/2$$

else

$$w_i = 1/16$$

end

end

ここで、 $q_i, q_i^{calc}, \alpha_i$ は、それぞれフレーム F_{t+1} における p_i の対応点、対応点の計算値、および対応点組の相関値である。また、画像の輝度値を 0 から 1 に正規化した場合、 $Th = 1/16$ とする。

ステップ 2: フレーム F_{t+1} における各画素の重みを求めるために、 w_i を基準として各画素の重み係数を補間処理によって求める。

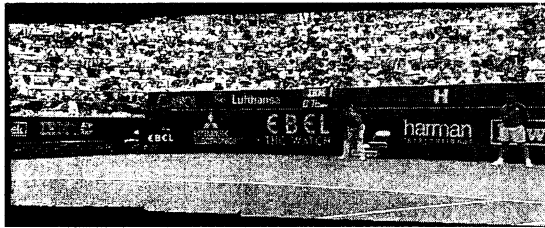
ステップ 3: 推定した変換行列 $H_{t_{ref} \rightarrow t+1}$ を使って射影変形したフレーム F_{t+1} と背景スプライト S_t を求めた重み係数に基づいて加重平均により重ね合わせる。

4. 実験

いくつかの動画像を用いて背景スプライトを生成した実験と生成した背景スプライトの品質評価について述べる。品質評価には、生成した背景スプライトから切り出した画像と対応する原画像の背景領域との PSNR を用いた。原画像の背景領域は



(a) 単純平均で重ね合わせた背景スプライト



(b) 提案手法で重ね合わせた背景スプライト

図 4 stefan から生成した背景スプライト

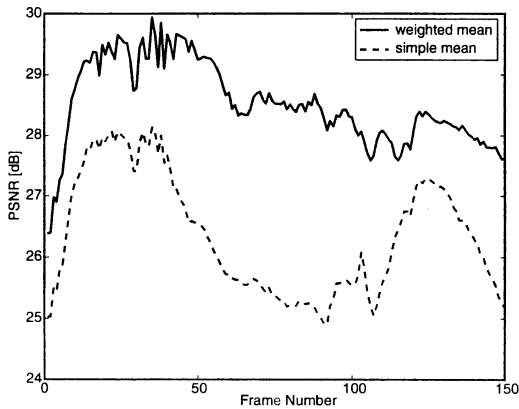


図 5 stefan の PSNR (実線：提案手法、点線：単純平均)

前景オブジェクト領域を手動で抽出して作成した。実験では、提案する重み付け手法の効果を評価するために、単純平均で重ね合わせた場合と比較した。

カメラが大きく動いている動画として、テニスプレイヤーが左右に動いているテスト動画 stefan の 1 フレーム目から 150 フレーム目までを使った。図 4(a) は単純平均で重ね合わせた場合の背景スプライトであり、図 4(b) は提案手法で重ね合わせた場合の背景スプライトである。単純平均で重ね合わせた場合は、テニスプレイヤーや画像端が現れている。一方、提案手法で重ね合わせた場合は、テニスプレイヤーが消え、高品質な背景スプライトが得られている。

生成した背景スプライトから切り出した画像と対応する原画像の背景領域との PSNR を図 5 に示す。単純平均を使った場合の平均 PSNR は 26.3dB であり、提案手法を使った場合の



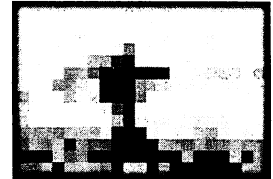
(a) フレーム 2



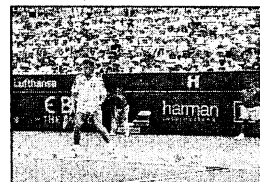
(b) (a) に対する重み係数



(c) フレーム 25



(d) (c) に対する重み係数



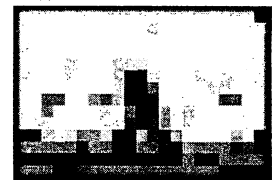
(e) フレーム 50



(f) (e) に対する重み係数



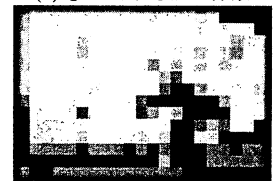
(g) フレーム 75



(h) (g) に対する重み係数



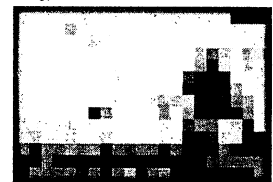
(i) フレーム 100



(j) (i) に対する重み係数



(k) フレーム 125



(l) (k) に対する重み係数



(m) フレーム 150



(n) (m) に対する重み係数

図 6 加重平均に用いる重み係数 (黒：重み小、白：重み大)

平均 PSNR は 28.5dB である。提案手法を用いることにより、PSNR が 2.2dB 向上した。

提案手法で求めた重み係数の例を図 6 に示す。左側がフレーム画像で、右側がそれぞれのフレームに対する重み係数である。重み係数が小さい部分を黒で表示し、重み係数が大きくなるにつれて白で表示している。求めた重み係数より、前景オブジェクトであるプレイヤーがいる領域の重みは小さくなっている。提案手法を使うことで前景オブジェクト領域の重みが小さくなり、背景スプライトに前景オブジェクトの情報があまり入らないため、PSNR が向上している。

カメラがほぼ静止している動画像として、オフィスを人が行き交うテスト動画像 hall monitor の 1 フレーム目から 300 フレーム目までを使った。図 7(a) は単純平均で重ね合わせた場合の背景スプライトであり、図 7(b) は提案手法を使った場合の背景スプライトである。得られた背景スプライト画像を比較すると、単純平均で重ねた場合よりも提案手法を使った場合の方がオブジェクトが消えている。提案手法を用いることで、カメラが固定されている場合でも、前景オブジェクトが動いていれば高品質な背景領域を抽出できることがわかる。

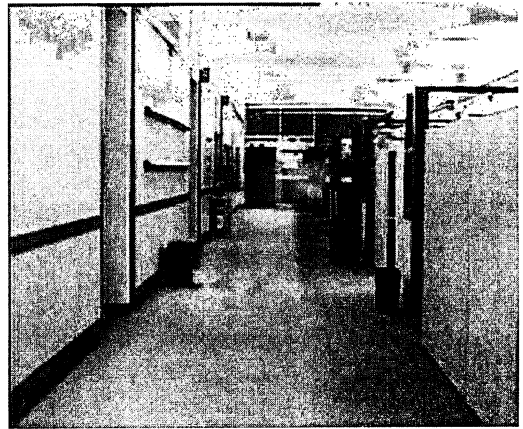
5. ま と め

本稿では、位相限定相関法 (Phase-Only Correlation: POC) に基づく高品質な背景スプライト生成を提案した。提案手法は、(i) POC に基づくサブピクセル精度の対応点探索手法を利用することで高精度な変換行列を推定することができ、(ii) 対応点組の相関値を利用することで適切に背景領域を重ね合わせることができる。また、生成した背景スプライトから切り出した画像と対応する原画像の背景領域との PSNR より、提案手法は、高品質な背景スプライトを生成できることを示した。

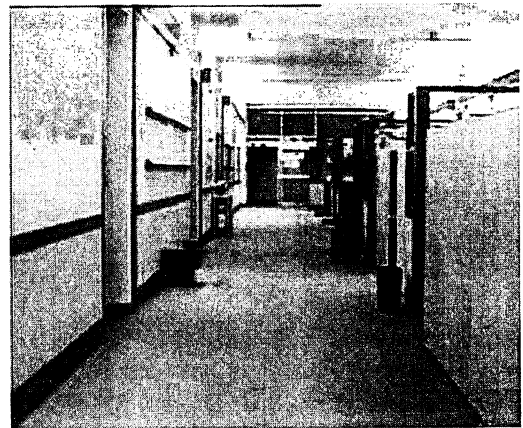
今後は、品質を向上するために提案手法を改善するとともに、背景差分によるオブジェクト抽出などのビデオモザイクングを使った応用を検討する予定である。

文 献

- [1] R. Szeliski, "Video mosaics for virtual environments," *IEEE Computer Graphics and Applications*, Vol. 16, No. 2, pp. 22-30, 1996.
- [2] M. Irani, P. Anandan, J. Bergen, R. Kumar, and S. Hsu, "Efficient representations of video sequences and their applications," *Signal Processing: Image Communication*, Vol. 8, pp. 327-351, 1996.
- [3] A. Smolic, T. Sikora, and J.-R. Ohm, "Long-term global motion estimation and its application for sprite coding, content description and segmentation," *IEEE Trans. Circuits Syst. Video Techn.*, Vol. 9, No. 8, pp. 1227-1242, 1999.
- [4] D. Farin, P. H. N. de With, and W. Effelsberg, "Video-object segmentation using multi-sprite background subtraction," *In ICME*, pp. 343-346, 2004.
- [5] Y. Lu, W. Gao, and F. Wu, "Efficient background video coding with static sprite generation and arbitrary-shape spatial prediction techniques," *IEEE Trans. Circuits Syst. Video Techn.*, Vol. 13, No. 5, pp. 394-405, 2003.
- [6] D. Capel, *Image Mosaicing and Super-resolution*, Springer, 2004.
- [7] C. D. Kuglin and D. C. Hines, "The phase correlation image alignment method," *Proc. Int. Conf. on Cybernetics and Society*, pp. 163-165, 1975.
- [8] K. Takita, T. Aoki, Y. Sasaki, T. Higuchi, and K. Kobayashi, "High-accuracy subpixel image registration based on phase-only correlation," *IEICE Trans. Fundamentals*, Vol. E86-A, No. 8, pp. 1925-1934, August 2003.
- [9] K. Takita, M. A. Muquit, T. Aoki, and T. Higuchi, "A sub-pixel correspondence search technique for computer vision applications," *IEICE Trans. Fundamentals*, Vol. E87-A, No. 8, pp. 1913-1923, August 2004.
- [10] 電子情報通信学会技術研究報告, vol. 105, no. 145-146, Jun 2005 (位相限定相関法関係 5 件発表: SIP2005-35, SIP2005-36, SIP2005-37, SIP2005-40, SIP2005-42).
- [11] H. C. Loy and T. Aoki, "Robust motion estimation for video sequences based on phase-only correlation," *Proceedings of the 6th IASTED International Conference on Signal and Image Processing*, pp. 441-446, August 2004.



(a) 単純平均で重ね合わせた背景スプライト画像



(b) 提案手法で重ね合わせた背景スプライト画像

図 7 hall monitor から生成した背景スプライト