位相限定相関法に基づく多視点画像からのデプスマップ生成の検討

酒井　修二† 伊藤　康一† 青木　孝文† 運天　弘樹††

† 東北大学 大学院情報科学研究科 〒 980–8579 宮城県仙台市青葉区荒巻字青葉 6–6–05
†† 凸版印刷株式会社 〒 112–8531 東京都文京区水道 1-3-3
E-mail: †{sakai,ito}@aoki.ecei.tohoku.ac.jp, aoki@ecei.tohoku.ac.jp, ††hiroki.unten@toppan.co.jp

あらまし 本稿では，位相限定相関法に基づく多視点画像からの 3 次元復元手法について述べる．デプスマップ統合に基づく多視点画像からの 3 次元復元手法では，高精度な 3 次元モデルを得るために，高精度なデプスマップ生成が重要となる．従来のデプスマップ生成手法では，奥行きを変化させながらウィンドウマッチングを行い，類似度が最大となる奥行きを真の奥行きとして決定していた．そのため，高精度なデプスマップを生成するためには，奥行き変化の刻み幅を非常に細かくする必要があり，マッチング回数が膨大になってしまう．これに対して，本稿では，位相限定相関法を用いたウィンドウマッチングと視差の正規化に基づく相関関数の統合を組み合わせることで，非常に少ないマッチング回数で奥行きを決定する手法を提案する．実験を通して，提案手法により，高速かつ高精度なデプスマップ生成が行えることを示す．

キーワード 多視点画像， 3 次元復元，デプスマップ生成，位相限定相関法，視差の正規化

1. はじめに

多視点画像からの 3 次元復元は，カメラを用いて撮影された複数枚の画像から物体の表面形状を復元する技術である [1]〜[5]．近年，3 次元復元技術の発展に伴い，カメラ画像のみを用いて非常に高精度な 3 次元形状モデルが生成できるようになった．それにより，多視点画像を用いた 3 次元復元は，コンピュータビジョンの研究コミュニティだけでなく，文化財のデジタルアーカイブやエンターテイメント産業など幅広い分野で注目されるようになった．特に，デジタルアーカイブでは，対象にレーザを当てることができなかったり，撮影場所が限られたりするため，カメラ画像のみを用いる 3 次元復元手法に対する需要が高い．

多視点画像を用いた 3 次元復元手法は，3 次元ボリュームに基づく手法，最適化による表面推定に基づく手法，特徴領域拡張に基づく手法，デプスマップ統合に基づく手法の 4 つに大きく分類される [1]．中でも，デプスマップ統合に基づく手法は，従来のステレオマッチングの技術を適用することが容易であること，および，初期値のメッシュモデルによって問題が収束しなかったり，復元対象によって復元される点群が極端に少なくなったりすることがないことから，その実用性が高い．本稿では，このデプスマップ統合に基づく手法に着目する．

デプスマップ統合に基づく手法は，(i) 多視点画像から複数のデプスマップを生成し，(ii) 複数のデプスマップからメッシュモデルを生成する 2 つのステップで構成される．ここで，(i) で生成されるデプスマップの精度は最終的に生成されるメッシュモデルの精度に大きく影響するので，高精度な 3 次元モデルを得るためにはデプスマップを高精度に求めることが重要である．

従来のデプスマップ統合に基づく手法の 1 つとして，Goesele らの手法 [2] が良く知られている．Goesele らの手法では，デプスマップ生成の際に，ある注目点に対して，その視線上で 3 次元点の奥行きを変化させながら正規化相互相関 (Normalized Cross-Correlation: NCC) を用いたウィンドウマッチングを行い，相関値が最大となる奥行きを真の奥行きとする．この手法では，複数の視点から計算される相関値を足し合わせることで，信頼性の高いデプスマップ生成が可能である．一方で，生成されるデプスマップの奥行きの分解能は奥行き変化の刻み幅に依存し，高精度な 3 次元復元を行うためには刻み幅を細かくする必要がある．しかし，奥行き変化の刻み幅を細かくすると，NCC を用いたマッチングの回数が増えるので，計算量が膨大になってしまう問題があった．

これに対して，本稿では，位相限定相関法 (Phase-Only Correlation: POC) [6], [7] に基づく多視点画像からのデプスマップ生成手法を提案する．提案手法では，ウィンドウ間のマッチングに理論的なピーク形状を有する POC を用い，各視点ペアでの視差を正規化し，相関関数を足し合わせることで，相関関数の相関ピークの高さだけでなく，相関ピークの位置を用いた奥行き探索を行う．また，相関関数のピーク位置推定と画像ピラミッドを用いた粗密探索を組み合わせることで，非常に少ないマッチング回数で奥行きの探索が行える．これにより，提案手法では，従来手法に比べ少ない計算量で高精度なデプスマップ生成が可能である．Herz-Jesu データセット [8] を用いた実験を通して，提案手法を用いることで，高速かつ高精度なデプスマップ生成が行えることを示す．

2. 位相限定相関法に基づく画像マッチング

本節では，1 次元位相限定相関法（1 次元 POC）に基づく画像マッチング手法について述べる．平行化されたステレオ画像ペアのように，ウィンドウ間の平行移動量が 1 次元方向に制限されている場合，1 次元 POC 関数を利用することで，高精度な平行移動量推定が可能である [7].

2.1 1 次元位相限定相関法

POC は画像信号の位相成分に着目した画像マッチング手法である [6], [7]．N 点の 2 つの 1 次元画像信号 $f(n)$ および $g(n)$ が与えられたとき，1 次元 POC を用いて 2 つの画像信号間の平行移動量を算出する方法について述べる．また，図 1 に 1 次元 POC の計算の流れを示す．ここで，1 次元画像信号の離散時間インデックス（整数）を，便宜上，$n = -M, \cdots, M$ とする．ただし，M は正の整数であり，$N = 2M + 1$ である．1 次元画像信号 $f(n)$ および $g(n)$ の 1 次元離散フーリエ変換 (Discrete Fourier Transform: DFT) を次式で定義する．

$$F(k) = \sum_{n=-M}^{M} f(n)W_N^{kn} = A_F(k)e^{j\theta_F(k)} \tag{1}$$

$$G(k) = \sum_{n=-M}^{M} g(n)W_N^{kn} = A_G(k)e^{j\theta_G(k)} \tag{2}$$

ここで，$k = -M, \cdots, M$ は離散周波数インデックス (整数) であり，$W_N = e^{-j\frac{2\pi}{N}}$ は回転因子である．また，$A_F(k)$ および $A_G(k)$ は振幅スペクトルであり，$\theta_F(k)$ および $\theta_G(k)$ は位相スペクトルを表す．このとき，正規化相互パワースペクトル $R(k)$ を次式で与える．

$$R(k) = \frac{F(k)\overline{G(k)}}{|F(k)\overline{G(k)}|} = e^{j(\theta_F(k)-\theta_G(k))} \tag{3}$$

ただし，$\overline{G(k)}$ は $G(k)$ の複素共役を表す．また，$\theta_F(k) - \theta_G(k)$ は 2 つの 1 次元画像信号の位相差スペクトルである．ここで，1 次元 POC 関数 $r(n)$ は，正規化相互パワースペクトル $R(k)$ の 1 次元逆離散フーリエ変換 (Inverse Discrete Fourier Transform: IDFT) として以下の式で表される．

$$r(n) = \frac{1}{N}\sum_{k=-M}^{M} R(k)W_N^{-kn} \tag{4}$$

1 次元画像信号 $f(n)$ と $g(n)$ が互いに微小量 δ だけ平行移動した関係にあると仮定すると，$f(n)$ と $g(n)$ の POC 関数 $r(n)$ は次式で与えられる．

$$r(n) \simeq \frac{\alpha}{N}\frac{\sin(\pi(n+\delta))}{\sin\left(\frac{\pi}{N}(n+\delta)\right)} \tag{5}$$

上式は，1 次元画像信号が δ だけ微小に平行移動した場合の POC 関数の一般形を表している．ここで，α は相関ピークの高さを表現するために導入されたパラメータである．画像に対して，無相関なノイズが加わると α の値が減少するため，実際には $\alpha \leqq 1$ となる．このとき，POC 関数の相関ピークの高さ α は 2 つの信号の類似度の指標に，相関ピークの位置座標 δ は 2 つの信号の平行移動量に相当する．このパラメータ α と δ を推定することで，信号の類似度とサブピクセル精度の平行移動量を求めることができる．

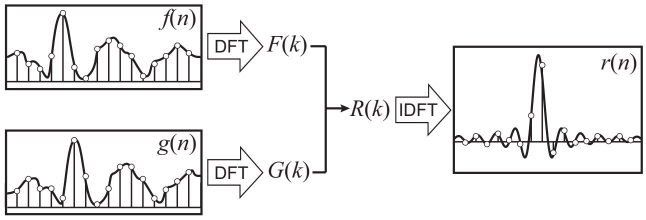

図 1　1 次元位相限定相関法 (1 次元 POC)

2.2 画像マッチングの高精度化

2.1 節では 2 つの画像信号が理想的に平行移動している場合の POC 関数を用いた平行移動量推定について説明した．しかし，実際の多視点ステレオ画像では，視点の基線長の変化に伴う歪みや，デジタル画像に生じる様々なノイズの影響により，平行移動量推定の際に誤差が生じる．そこで，以下では，1 次元 POC 関数を用いて平行移動量推定を行う際に重要となる各種の高精度化手法について述べる．

2.2.1 窓関数の適用

1 次元 DFT では，信号が周期的に循環することを仮定するため，端点での不連続性が問題となる．この不連続性の影響を軽減するため，1 次元画像信号 $f(n)$ および $g(n)$ に対して窓関数を適用する．本稿では，窓関数として次式で表される 1 次元ハニング窓を用いる．

$$w(n) = \frac{1}{2} + \frac{1}{2}\cos\left(\frac{\pi n}{M}\right) \tag{6}$$

2.2.2 スペクトルの重み付け

一般に，自然画像のエネルギーは低周波数成分に集中し，高周波数成分のエネルギーは相対的に小さいことが知られている．このため，エイリアシング，雑音，歪みなどの外乱が加わると，高周波数成分の SN 比が大幅に低下する．そこで，信頼性の低い高周波数成分の影響を抑制するために，正規化相互パワースペクトル $R(k)$ の計算の際に，低域通過型のスペクトル重み付け関数 $H(k)$ を適用する．本稿では，次式で表されるガウス関数を用いる．

$$H(k) = e^{-2\pi^2\sigma^2\frac{k^2}{N^2}} \tag{7}$$

ここで，σ はガウス関数の幅を表す定数であり，本稿では $\sigma = \sqrt{0.5}$ とする．このとき，相関ピークモデルも以下のようなガウス型になる．

$$r(n) \simeq \frac{\alpha}{\sqrt{2\pi}\sigma}e^{-\frac{(n+\delta)^2}{2\sigma^2}} \tag{8}$$

2.2.3 相関ピークモデルの関数フィッティングによるパラメータ推定

一般に，移動量 δ は実数値であり，POC 関数のピーク座標はサンプリング格子点の間に存在する．そこで，相関ピークのモデル式 (8) を，実際に計算された POC 関数の数値データに対してフィッティングすることで，画像のピクセル間に存在するピークの位置を推定する．このとき，α および δ がフィッティングパラメータとなる．本稿では，フィッティング手法として, 非線型最小 2 乗問題の解法の 1 つである Levenberg-Marquardt 法を用いる．

2.2.4 複数の 1 次元画像信号の利用

POC に基づく画像マッチング手法をステレオ画像間の対応付けに適用する場合，注目点を中心にマッチングウィンドウを切り出し，切り出されたウィンドウ間で POC 関数を計算する．2.1 節では，1 組の 1 次元画像信号を用いて，その 1 次元画像信号間の平行移動量を求める手法について説明した．しかし，実際に撮影した画像において，1 組の 1 次元画像信号のみを用いると，ノイズなどの影響によりピークの位置を高精度に推定することが困難である [7]．そこで，マッチングウィンドウを切り出す際に，注目点の近傍から複数の 1 次元画像信号の組を抽出し，それらの POC 関数の平均を平均 POC 関数とする．この平均 POC 関数を相関ピーク α および平行移動量 δ の推定に用いることでノイズの影響を抑えることが可能である．

3. 位相限定相関法に基づく多視点画像からのデプスマップ生成手法

本節では，位相限定相関法に基づく多視点画像からのデプスマップ生成手法を提案する．提案手法では，複数のステレオペアにおける POC 関数を統合し，POC 関数のピーク位置を用いて奥行きを決定する．しかし，異なるステレオペアから計算された POC 関数では，そのステレオペアにおける視差の違いから，ピーク位置が異なるという問題がある．そこで，提案手法では，それぞれのステレオペアにおける視差を正規化することで，POC 関数を統合する．従来の相関関数の統合手法として Okutomi らの手法 [9] があるが，Okutomi らの手法では，すべてのカメラが同じ直線上にあることを仮定しており，一般的な多視点画像に適用できない．本節で述べる正規化視差に基づく相関関数の統合手法は，Okutomi らの手法をより一般化したものであり，カメラが同一の直線上にない場合においても相関関数を統合することが可能である．

提案手法では，カメラパラメータが既知の多視点画像 $\mathbf{V} = \{V_0, \cdots, V_{L-1}\}$ のうち，参照視点 $V_R \in \mathbf{V}$ とその近傍視点 $\mathbf{C} = \{C_0, \cdots, C_{K-1}\} \subset \mathbf{V} - \{V_R\}$ を入力として考える．ここで，L および K は，それぞれ多視点画像の数と近傍視点の数である．そして，V_R および $\mathbf{C}$ から K 組の平行ステレオペアを生成し [10]，それぞれのペアで計算された POC 関数を正規化視差に基づき統合することで，POC 関数のピーク位置から奥行きを決定する．この POC 関数による奥行き推定と画像ピラミッドを用いた粗密探索を組み合わせることで，従来手法と比べ，少ないマッチング回数で奥行き探索を実現することが可能である．

3.1 視差の正規化

平行ステレオペア V_R - C_i における 3 次元点 $\mathbf{M} = (X, Y, Z)$ の視差を考える [10]．ここで，$C_i \in \mathbf{C}$ である．また，世界座標と参照視点 V_R のカメラ座標が一致すると仮定する．さらに，参照視点 V_R と近傍視点 C_i でステレオ平行化を行う際に，参照視点におけるカメラ座標の回転行列を次式のように $\mathbf{R}_i$ とする．

$$\mathbf{R}_i = \begin{bmatrix} R_{i11} & R_{i12} & R_{i13} \\ R_{i21} & R_{i22} & R_{i23} \\ R_{i31} & R_{i32} & R_{i33} \end{bmatrix} \tag{9}$$

このとき，3 次元点 $\mathbf{M}$ と平行化後の座標 $\mathbf{M}_i = (X_i, Y_i, Z_i)$ の関係は次式で表される．

$$\mathbf{M} = \mathbf{R}_i \mathbf{M}_i \tag{10}$$

また，平行ステレオペア V_R - C_i における 3 次元点 $\mathbf{M}_i$ と視差 d_i の関係は次式で表される [10]．

$$\begin{bmatrix} X_i \\ Y_i \\ Z_i \end{bmatrix} = \begin{bmatrix} (u_i - u_{0i})B_i/d_i \\ (v_i - v_{0i})B_i/d_i \\ \beta_i B_i/d_i \end{bmatrix} \tag{11}$$

ここで，(u_i, v_i) は平行化後の参照視点における $\mathbf{M}_i$ の対応点のディジタル画像座標，(u_{0i}, v_{0i}) は平行化後の参照視点における画像中心，β_i は平行化後の焦点距離，B_i はステレオペア間の基線長を表す．一般的に，焦点距離は f で表されるが，本稿では，マッチングを行う 1 次元画像信号 f との混同を避けるため，β を用いる．式 (10) と式 (11) より，3 次元点 $\mathbf{M}$ と平行化後の視差 d_i との関係は次式で表される．

$$\mathbf{M} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \mathbf{R}_i \begin{bmatrix} (u_i - u_{0i})B_i/d_i \\ (v_i - v_{0i})B_i/d_i \\ \beta_i B_i/d_i \end{bmatrix} \tag{12}$$

同様に，参照視点 V_R と近傍視点 $C_j \in \mathbf{C} - \{C_i\}$ の間で次式が成り立つ．

$$\mathbf{M} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \mathbf{R}_j \begin{bmatrix} (u_j - u_{0j})B_j/d_j \\ (v_j - v_{0j})B_j/d_j \\ \beta_j B_j/d_j \end{bmatrix} \tag{13}$$

よって，式 (12) と式 (13) より，

$$X = \frac{B_i}{d_i}(R_{i11}(u_i - u_{0i}) + R_{i12}(v_i - v_{0i}) + R_{i13}\beta_i)$$

$$= \frac{B_j}{d_j}(R_{j11}(u_j - u_{0j}) + R_{j12}(v_j - v_{0j}) + R_{j13}\beta_j) \tag{14}$$

$$Y = \frac{B_i}{d_i}(R_{i21}(u_i - u_{0i}) + R_{i22}(v_i - v_{0i}) + R_{i23}\beta_i)$$
$$= \frac{B_j}{d_j}(R_{j21}(u_j - u_{0j}) + R_{j22}(v_j - v_{0j}) + R_{j23}\beta_j) \tag{15}$$

$$Z = \frac{B_i}{d_i}(R_{i31}(u_i - u_{0i}) + R_{i32}(v_i - v_{0i}) + R_{i33}\beta_i)$$
$$= \frac{B_j}{d_j}(R_{j31}(u_j - u_{0j}) + R_{j32}(v_j - v_{0j}) + R_{j33}\beta_j) \tag{16}$$

となり，平行ステレオペア V_R - C_i における視差 d_i と視点ペア V_R - C_j における視差 d_j の間には次の関係が成り立つ．

$$\begin{aligned} d_i &= \frac{R_{i11}(u_i - u_{0i}) + R_{i12}(v_i - v_{0i}) + R_{i13}\beta_i}{R_{j11}(u_j - u_{0j}) + R_{j12}(v_j - v_{0j}) + R_{j13}\beta_j}\frac{B_i}{B_j}d_j \\ &= \frac{R_{i21}(u_i - u_{0i}) + R_{i22}(v_i - v_{0i}) + R_{i23}\beta_i}{R_{j21}(u_j - u_{0j}) + R_{j22}(v_j - v_{0j}) + R_{j23}\beta_j}\frac{B_i}{B_j}d_j \\ &= \frac{R_{i31}(u_i - u_{0i}) + R_{i32}(v_i - v_{0i}) + R_{i33}\beta_i}{R_{j31}(u_j - u_{0j}) + R_{j32}(v_j - v_{0j}) + R_{j33}\beta_j}\frac{B_i}{B_j}d_j \end{aligned} \tag{17}$$

よって，式 (17) で示したように，異なる平行ステレオペアにおける視差 d_i と d_j はある倍率によって関係付けられ，その倍率はカメラパラメータと参照視点における 3 次元点の対応点のディジタルカメラ座標に依存する．以上より，各ステレオペアにおける視差の倍率を考慮することで，正規化視差 d を定義する．平行ステレオペア V_R - C_i $(i = 0, \cdots, K-1)$ が与えられたとき，各平行ステレオペアにおける視差 d_i と正規化視差 d の間には以下の関係が成り立つ．

$$d_i = s_i d \tag{18}$$

ここで，s_i は視差の倍率である．平行化前の参照視点における注目点 (u, v) について，平行化後の各平行ステレオペアにおける視差倍率 s_i を式 (17) から次式のように定義する．

$$s_i = \frac{1}{|s|}(R_{i31}(u_i - u_{0i}) + R_{i32}(v_i - v_{0i}) + R_{i33}\beta_i)B_i \tag{19}$$

ここで，$|s|$ は

$$|s| = \frac{1}{K}\sum_{j=0}^{K-1} (R_{j31}(u_j - u_{0j}) + R_{j32}(v_j - v_{0j}) + R_{j33}\beta_j)B_j \tag{20}$$

である．このとき，平行ステレオペア V_R - C_i における 3 次元点 $\mathbf{M}_i$ は次式で与えられる．

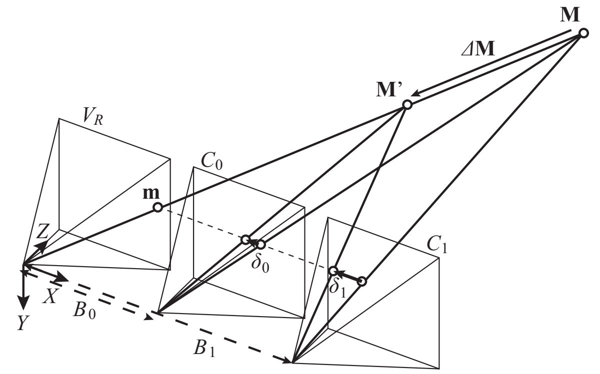

図 2　3 次元点の変移と視差の変移の関係

$$\mathbf{M}_i = \begin{bmatrix} X_i \\ Y_i \\ Z_i \end{bmatrix} = \begin{bmatrix} (u_i - u_{0i})B_i/(s_i d) \\ (v_i - v_{0i})B_i/(s_i d) \\ \beta_i B_i/(s_i d) \end{bmatrix} \tag{21}$$

3.2　正規化視差に基づく POC 関数の統合

図 2 のように，参照視点 V_R から 3 次元点 $\mathbf{M}$ を通る視線上を微小量 $\Delta\mathbf{M} = (\Delta X, \Delta Y, \Delta Z)$ だけずれた点 $\mathbf{M}' = \mathbf{M} + \Delta\mathbf{M}$ について，平行ステレオペア V_R - C_i $(C_i \in \mathbf{C})$ における視差を考える．このとき，式 (12) より，3 次元点 $\mathbf{M}'$ は，平行ステレオペア V_R - C_i における視差 d'_i を用いて

$$\mathbf{M}' = \mathbf{R}_i \begin{bmatrix} (u_i - u_{0i})B_i/d'_i \\ (v_i - v_{0i})B_i/d'_i \\ \beta_i B_i/d'_i \end{bmatrix} \tag{22}$$

と表される．ここで，$\mathbf{M}$ を真の 3 次元点とすると，平行ステレオペア V_R - C_i における視差は誤差 δ_i が生じるので，$\mathbf{M}$ に対する視差 d_i と $\mathbf{M}'$ に対する視差 d'_i の関係は

$$d'_i = d_i + \delta_i \tag{23}$$

となる．局所的な画像変形を平行移動のみと仮定すると，対応点を中心に V_R から切り出したウィンドウ f_i と C_i から切り出したウィンドウ g_i の間には δ_i の平行移動が生じていると考えられる．この平行移動量 δ_i は，2. 節で述べたように，f_i と g_i の 1 次元 POC 関数を計算し，そのピーク位置を推定することで求められる．しかしながら，異なる平行ステレオペア V_R - C_i および V_R - C_j $(C_j \in \mathbf{C} - \{C_i\})$ において，平行移動量 δ_i と δ_j は必ずしも一致せず，それぞれの平行ステレオペアから切り出されたウィンドウで計算される 1 次元 POC 関数のピーク位置は異なる．例えば，図 2 では，平行ステレオペア V_R - C_0 の基線長 B_0 よりも V_R - C_1 の基線長 B_1 の方が長く，同じ奥行きの変化 ΔZ に対して V_R - C_1 上の対応点の変動 δ_1 の方が δ_0 よりも大きくなる．

そこで，提案手法では，正規化視差に基づき各平行ステレオペアで切り出すウィンドウを拡大縮小することで，それぞれの平行ステレオペアで計算される 1 次元 POC

関数のピーク位置を統一する．$\mathbf{M}$ における正規化視差を d，$\mathbf{M}'$ における正規化視差を $d' = d + \delta$ とすると式 (21) より $\mathbf{M}'$ は

$$\mathbf{M}' = \mathbf{R}_i \begin{bmatrix} (u_i - u_{0i})B_i/(s_i(d+\delta)) \\ (v_i - v_{0i})B_i/(s_i(d+\delta)) \\ \beta_i B_i/(s_i(d+\delta)) \end{bmatrix} \tag{24}$$

と表される．よって，平行ステレオペア V_R - C_i において対応点を中心に切り出したウィンドウ f_i と g_i 間の平行移動量は $s_i\delta$ となる．ここで，基準となるウィンドウサイズ w に対して，f_i，g_i を切り出す際のウィンドウサイズを $s_i w$ とする．この $s_i w$ の大きさの 1 次元画像信号を w の大きさにそろえるため，$1/s_i$ 倍に拡大縮小した信号を f_i'，g_i' とする．このとき，f_i'，g_i' 間の平行移動量は δ となり，f_i, g_i' から計算される 1 次元 POC 関数のピーク位置は δ となる．

ここで，平行ステレオペア V_R - C_j $(C_j \in \mathbf{C} - \{C_i\})$ について同様に考えると，切り出すウィンドウのサイズ $s_j w$ は $s_i w$ とは異なるが，1 次元 POC 関数のピーク位置は δ となり，平行ステレオペア V_R - C_i で計算される 1 次元 POC 関数と一致する．実際に計算される POC 関数の例を図 3 に示す．図 3 (a) は，視差の倍率を考慮せず，すべての平行ステレオペアにおいて同じウィンドウサイズ w でウィンドウを切り出した場合の POC 関数である．この場合，ウィンドウ間の平行移動量 δ_i はそれぞれ異なり，計算される POC 関数のピーク位置は一致しない．これに対して，図 3 (b) は，各ステレオペアにおいて，正規化視差に基づき切り出すウィンドウサイズを $s_i w$ とした場合の POC 関数である．図 3 (b) より，視差の倍率を考慮することで，それぞれの POC 関数のピーク位置 δ が一致することがわかる．これにより，異なる平行ステレオペアで計算される 1 次元 POC 関数を統合することが可能となる．本稿では，複数の 1 次元 POC 関数の統合として，各平行ステレオペアで計算された 1 次元 POC 関数を平均化する．統合された 1 次元 POC 関数のピーク位置 δ を推定することで真の 3 次元点 $\mathbf{M}$ を以下の式より求めることができる．

$$\mathbf{M} = \mathbf{R}_i \begin{bmatrix} (u_i - u_{0i})B_i/(s_i(d'-\delta)) \\ (v_i - v_{0i})B_i/(s_i(d'-\delta)) \\ f_i B_i/(s_i(d'-\delta)) \end{bmatrix} \tag{25}$$

ここで，オクルージョンにより 3 次元点 $\mathbf{M}$ が近傍視点 $C_i \in \mathbf{C}$ の画像に写っていな場合や，物体境界において複数の視差を持つ領域でウィンドウを抽出した場合は，そのウィンドウから計算される POC 関数のピーク位置に非常に大きい誤差が生じることが予想される．一方で，マッチングウィンドウ間の画像変形が平行移動のみに近似できない場合は，そのウィンドウから計算される POC 関数のピーク値 α が低くなることが実験的に知られている．そこで，複数の 1 次元 POC 関数の平均化の際に，POC 関数のピーク値 α が閾値 th_{corr} を超えるもののみを用いて計算することで，オクルージョンや物体境界の影響などを抑えることが可能である．

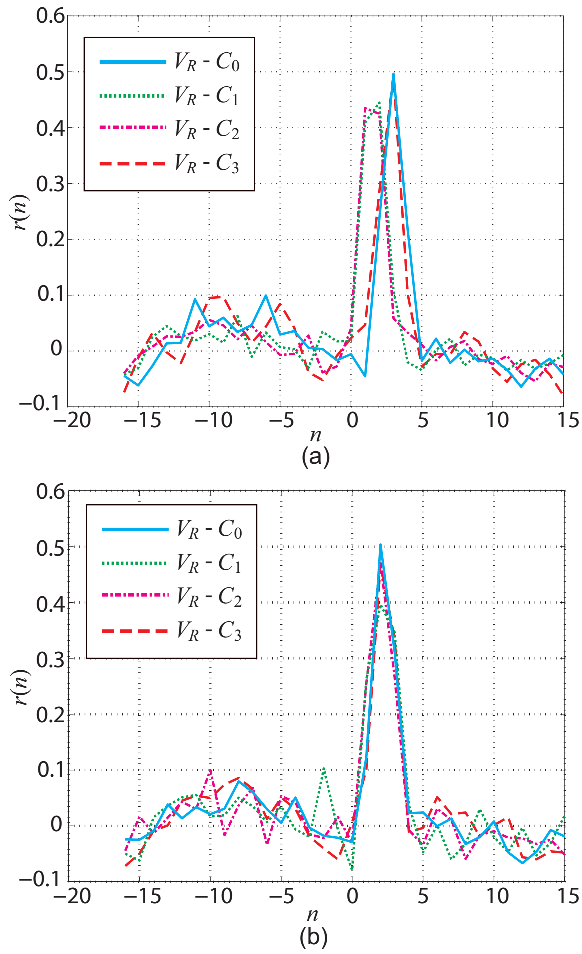

図 3　正規化視差に基づく相関関数の統合: (a) ウィンドウサイズ w で計算された POC 関数, (b) ウィンドウサイズ $s_i w$ で計算された POC 関数

3.3　位相限定相関法に基づく多視点画像からのデプスマップ生成手法

POC に基づく多視点画像からのデプスマップ生成手法を提案する．提案手法では，3.2 節で述べた POC 関数の統合による奥行き推定手法と，画像ピラミッドを用いた粗密探索とを組み合わせることで，高速かつ高精度なデプスマップ生成を行う．以下に，処理手順を説明する．ここで，入力は参照視点 V_R とその近傍視点 $\mathbf{C} = \{C_0, \cdots, C_{K-1}\}$ とし，出力はデプスマップ dep および相関値マップ $corr$，信頼値マップ $conf$ とする．また，各パラメータとして，相関ピーク値の閾値を th_{corr}，画像ピラミッドの階層数を H，基準となるウィンドウの大きさを w，平均 POC の計算に用いるライン数を l とする．

まず，参照視点 V_R と近傍視点 $\mathbf{C}$ から平行ステレオペアおよびその階層画像を生成する．$i = 0, \cdots, K-1$ について，ペア V_R - C_i に対してステレオ平行化 [10] を

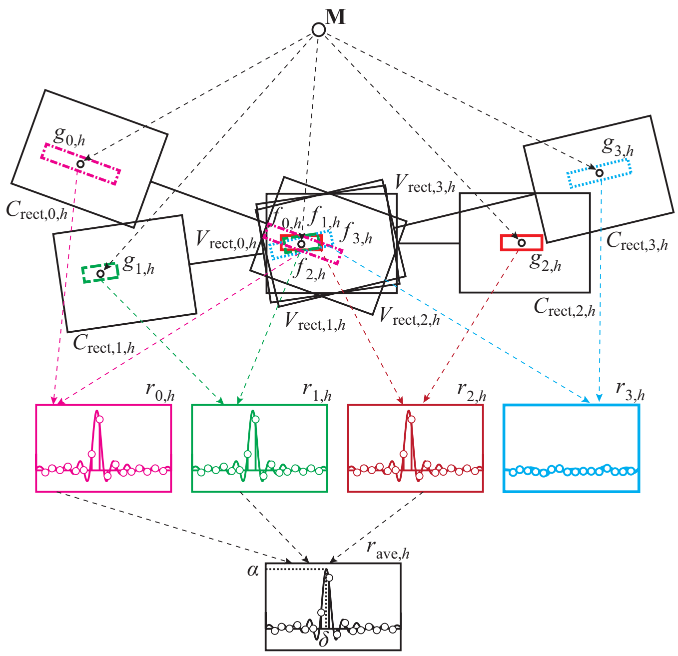

図 4　複数の POC 関数の統合による平行移動量推定

行い，平行化後の画像をそれぞれ $V_{\mathrm{rect},i,0}$, $C_{\mathrm{rect},i,0}$ とする．そして，$h = 1, \cdots, H-1$ について，$V_{\mathrm{rect},i,0}$ および $C_{\mathrm{rect},i,0}$ を 2^{-h} 倍した画像を $V_{\mathrm{rect},i,h}$ および $C_{\mathrm{rect},i,h}$ とする．次に，画像ピラミッドの最上層 $H-1$ において，画像全体のマッチングを行い，奥行きの探索開始座標 Z_{init} を決定する．この Z_{init} は，$i = 0, \cdots, K-1$ について，$V_{\mathrm{rect},i,H-1}$ と $C_{\mathrm{rect},i,H-1}$ の POC 関数を計算し，そのピーク位置からおよその視差を求めることで決定する．そして，参照視点 V_R 上の任意の点 $\mathbf{m} = (u, v)$ について，以下の **Step1** から **Step7** の処理により奥行きおよび相関値，信頼値を決定し，$dep(\mathbf{m})$, $corr(\mathbf{m})$, $conf(\mathbf{m})$ に書き込む．この $\mathbf{m}$ の座標を V_R 上で変化させながら処理を繰り返すことで，dep, $corr$, $conf$ を決定する．画像ピラミッドの各階層における処理の流れを図 4 に示す．

Step1:　$i = 0, \cdots, K-1$ について，$\mathbf{m}$ における視差倍率 s_i を式 (19) から決定する．このとき，平行化後の参照視点 $V_{\mathrm{rect},i,0}$ における $\mathbf{m}$ の対応点 $\mathbf{m}_i = (u_i, v_i)$ はステレオ平行化の変形に用いる射影行列から計算される．

Step2:　$h \leftarrow H-1$ および 3 次元点 $\mathbf{M} \leftarrow \mathbf{M}_{\mathrm{init}}$ として探索を開始する．この $\mathbf{M}_{\mathrm{init}}$ は，参照視点 V_R において座標 $\mathbf{m}$ を通る視線上で奥行きが Z_{init} となる 3 次元点である．

Step3:　$i = 0, \cdots, K-1$ について，$\mathbf{M}$ をステレオペア $V_{\mathrm{rect},i,h}$ - $C_{\mathrm{rect},i,h}$ に投影した点をそれぞれ $\mathbf{m}_{i,h} = (u_{i,h}, v_{i,h})$, $\mathbf{m}'_{i,h} = (u'_{i,h}, v'_{i,h})$, とする．さらに，$V_{\mathrm{rect},i,h}$ から $\mathbf{m}_{i,h}$ を中心に $s_i w \times l$ の大きさで切り出されたウィンドウ $f_{i,h}$ と $C_{\mathrm{rect},i,h}$ から $\mathbf{m}'_{i,h}$ を中心に $s_i w \times l$ の大きさで切り出されたウィンドウ $g_{i,h}$ から 1 次元 POC 関数 $r_{i,h}$ を計算する．

Step4:　$r_{i,h}$ のうち POC 関数のピーク値 $\alpha_i > th_{corr}$ であるものを平均化し，$r_{\mathrm{ave},h}$ とする．ここで，$\alpha_i > th_{corr}$ となる POC 関数の個数を K' とする．

Step5:　1 次元 POC 関数 $r_{\mathrm{ave},h}$ に対して関数フィッティングを行うことで，POC 関数のピーク値 α とピーク位置 δ を得る．この δ より式 (25) を用いて 3 次元点 $\mathbf{M}$ を更新する．

Step6:　$h = 0$ となるまで $h \leftarrow h-1$ とし，**Step3** から **Step6** を繰り返す．

Step7:　$h = 0$ のときの 1 次元 POC 関数 $r_{\mathrm{ave},0}$ と 3 次元点 $\mathbf{M} = (X, Y, Z)$ より，$dep(\mathbf{m})$, $corr(\mathbf{m})$, $conf(\mathbf{m})$ を以下のように定義する．

$$dep(\mathbf{m}) = Z \tag{26}$$

$$corr(\mathbf{m}) = \alpha \tag{27}$$

$$conf(\mathbf{m}) = \frac{K'}{K} \frac{(\alpha - th_{corr})}{(1 - th_{corr})} \tag{28}$$

以上の処理により決定されたデプスマップ dep について，相関値マップ $corr$，信頼値マップ $conf$ が閾値以上の点のみを復元することにより，信頼性の高い 3 次元点群を得ることが可能である．

4. 評価実験

データセット Herz-Jesu [8] を用いて従来手法と提案手法を比較する．データセット Herz-Jesu には，多視点画像とそのカメラパラメータ，および復元対象の真値のメッシュメッシュモデル，バウンディングボックスが含まれる．Herz-Jesu の画像のうち 1 枚を参照視点 V_R（図 5 (a) ），その近傍 4 枚の画像を近傍視点 $\mathbf{C} = \{C_0, C_1, C_2, C_3\}$（図 5 (b)-(e) ）とし，従来手法と提案手法を用いて V_R におけるデプスマップを生成する．ここで，Herz-Jesu の画像は $3,072 \times 2,048$ ピクセルと高解像度の画像であるが，処理時間短縮のため，$1,536 \times 1,024$ ピクセルに縮小した画像を用いた．従来手法として Goesele らの手法 [2] を用いる．Goesele らの手法では，奥行き変化の刻み幅 ΔZ が精度と計算量に大きく影響する．そこで，本稿では，$\Delta Z = 200, 100, 50, 20, 10$ mm についてそれぞれ実験を行う．このとき，奥行きの探索範囲はデータセットに含まれるバウンディングボックス内とする．また，NCC のマッチングウィンドウの大きさは 5×5 ピクセルとする．

提案手法は 3.3 節の手法とし，各パラメータは次のように設定する．まず，画像ピラミッドの階層数 H は 4 とする．次に，画像ピラミッドの上位層 $h > 0$ では，相関値の閾値 th_{corr} を 0.3，基準となるウィンドウの大きさ w を 32 ピクセルとする．画像ピラミッドの最下層 $h = 0$ では，相関値の閾値 th_{corr} を 0.7，基準となるウィンドウの大きさ w を 8 ピクセルとする．さらに，平均 POC の計算に用いるライン数 l は，どの階層においても $l = w/2 + 1$ とする．ここで，POC 関数の計算

(a) (b) (c) (d) (e)

図 5　実験に用いた画像: (a) 参照視点 V_R , (b)-(e) 近傍視点 $\mathbf{C} = \{C_0, C_1, C_2, C_3\}$

表 1　実験に用いた計算機の仕様

CPU	Intel(R) Core(TM)2 Extreme CPU X9650 3.00GHz
Memory	3GB
OS	Vine Linux 4.2
C complier	gcc version 3.3.6

では，2.2.1 節で述べたようにハニング窓をかけるため，最下位層でのマッチングウィンドウの大きさ 8×5 ピクセルは，従来手法で用いる 5×5 ピクセルのウィンドウの大きさに相当することに注意されたい．

従来手法と提案手法を C 言語で実装し，1 つの参照視点からのデプスマップ生成について比較する．従来手法と提案手法の比較として，生成されたデプスマップから信頼値が 0.6 以上となる点のみを 3 次元復元し，その復元点数，誤対応点数，誤対応率，真値のメッシュモデルとの誤差の RMS (Root Mean Square) ，処理時間で評価する．誤対応点は真値のメッシュモデルとの誤差が 200 mm 以上の点とし，誤対応率は復元点数のうち誤対応点が含まれる割合を表す．ここで，真値のメッシュモデルとの誤差が 200 mm 以上の点は，実験に用いた参照視点-近傍視点ペアのうち，最も基線長が短いペアにおいて ± 1 ピクセルの視差の誤差に相当する．また，処理時間の測定に用いた計算機の仕様を表 1 に記す．

実験結果を表 2 に示す．また，従来手法と提案手法を用いて生成されたデプスマップより復元された 3 次元点群を図 6 に示す．まず，復元点数と誤対応点数について比較すると，提案手法は Goesele らの手法に比べ，復元点数が多く，誤対応点が少ないことがわかる．もちろん，Goesele らの手法において，復元する点の信頼値の閾値を上げることで誤対応点数を減らすことは可能であるが，その場合，復元点数も減少することに注意されたい．この結果は図 6 からも確認できる．提案手法では，画像ピラミッドの上位層においてウィンドウサイズを大きくし，多くの情報を用いることで，大きな誤対応の発生を抑制しているためであると考えられる．また，POC 関数に基づく画像マッチングでは，周波数領域で重み付けを行い，ノイズの影響を抑えるため，テクスチャの少ない領域でもノイズの影響が少なく，高精度なマッチングを行える．

次に，復元された 3 次元点群の精度と処理時間について比較する．提案手法は，Goesele らの手法の $\Delta Z = 10$ mm の場合よりも高精度に点群を求められており，そのときの処理は 20 倍以上高速である．また，Goesele らの手法の $\Delta Z = 200$ mm の場合について，処理時間はほぼ同じであるが，復元精度は提案手法の方が 2 倍以上高精度であるといえる．Goesele らの手法の $\Delta Z = 200$ mm の場合に奥行きの分解能が不十分であることが，図 6 (a) からも確認できる．また，提案手法の方が高速に奥行き探索を行えていることについては，提案手法の方が従来手法に比べマッチング回数が非常に少ないためであると考えられる．実験に用いたデータの場合，Goesele らの手法では，1 点の注目点の奥行き探索ついて，$\Delta Z = 200$ mm の時に 85 回，$\Delta Z = 10$ mm の時に 1,701 回のマッチングを行う．これに対して，提案手法では，画像ピラミッドの各階層で 1 回ずつマッチングを行うため，1 点の注目点について 4 回という非常に少ないマッチング回数で奥行きの探索を行える．以上の結果より，提案手法を用いることで，従来手法に比べ高速かつ高精度なデプスマップ生成が行える．

5. まとめ

本稿では，位相限定相関法に基づく多視点画像からのデプスマップ生成手法を提案した．また，実験を通して，提案手法を用いることで，従来手法に比べ高速かつ高精度なデプスマップ生成が可能であることを示した．

今後の課題として，提案手法により生成されたデプスマップからのメッシュモデル生成手法を検討している．提案手法を用いることで，非常に高精度なデプスマップを生成することが可能であり，従来のデプスマップ統合によるメッシュモデル生成においても，高精度な 3 次元モデルが生成できると考えられる．また，基線長の長いステレオ画像に対してもロバストなマッチングを行える Scaled Window-POC (SW-POC) [11] を導入することで，よりロバスト性の高い手法への改良を検討している．

文　献

[1] S. M. Seitz, B. Curless, J. Diebel, D. Scharstein and R. Szeliski: "A comparison and evaluation of multi-view stereo reconstruction algorithms", Proc. Int'l Conf. Computer Vision and Pattern Recognition, pp. 519–528 (2006).

[2] M. Goesele, N. Snavely, B. Curless, H. Hoppe and S. M. Seitz: "Multi-view stereo for community photo collections", Proc. IEEE Int'l Conf. Computer Vision,

表 2 実験結果

	ΔZ [mm]	復元点数	誤対応点数	誤対応率 [%]	RMS 誤差 [mm]	処理時間 [sec]
Goesle	200	579,038	13,089	2.2605	48.9632	2,230
	100	699,480	10,222	1.4614	33.2230	4,433
	50	737,107	9,488	1.2872	25.9389	8,750
	20	734,730	9,160	1.2467	23.4032	22,060
	10	728,797	9,007	1.2359	22.9967	44,117
提案手法		937,177	4,451	0.4749	22.2474	2,128

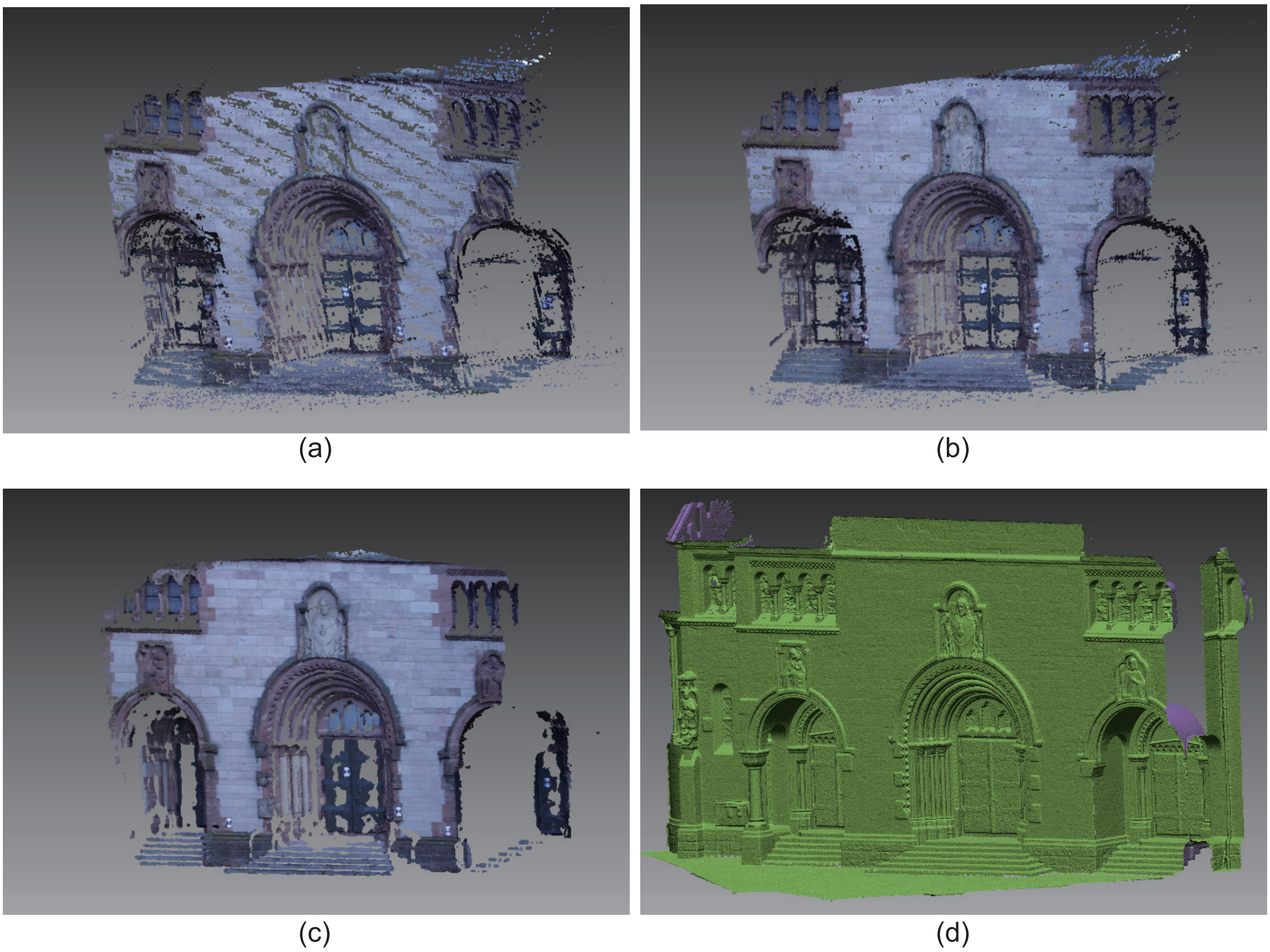

図 6 復元された 3 次元点群: (a) Goesele らの手法 ($\Delta Z = 200$ mm) , (b) Goesele らの手法 ($\Delta Z = 10$ mm) , (c) 提案手法 , (d) 真値のメッシュモデル

pp. 1–8 (2007).

[3] M. Goesele, J. Ackermann, S. Fuhrmann, R. Klowsky, F. Langguth, P. Mucke and M. Ritz: "Scene reconstruction from community photo collections", IEEE Computer, **Vol. 43**, pp. 48–53 (2010).

[4] Y. Furukawa and J. Ponce: "Accurate, dense, and robust multiview stereopsis", IEEE Trans. Pattern Analysis and Machine Intelligence, **Vol. 32**, No. 5, pp. 1362–1376 (2010).

[5] S. Agarwal, Y. Furukawa, N. Snavely, B. Curless, S. M. Seitz and R. Szeliski: "Reconstructing Rome", IEEE Computer, **Vol. 43**, pp. 40–47 (2010).

[6] 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 (2004).

[7] 柴原, 沼, 長嶋, 青木, 中島, 小林："一次元位相限定相関法に基づくステレオ画像の高精度サブピクセル対応付け手法", 電気情報通信学会論文誌 D, **Vol. J91-D**, No. 9, pp. 2343–2356 (2008).

[8] C. Strecha, W. von Hansen, L. V. Gool, P. Fua and U. Thoennessen: "On benchmarking camera calibration and multi view stereo for high resolution imagery", Proc. Int'l Conf. Computer Vision and Pattern Recognition, pp. 1–8 (2008).

[9] M. Okutomi and T. Kanade: "A multiple-baseline stereo", IEEE Trans. Pattern Analysis and Machine Intelligence, **Vol. 15**, No. 4, pp. 353–363 (1993).

[10] 徐, 辻："3 次元ビジョン", 共立出版 (1998).

[11] 酒井, 高橋, 伊藤, 青木："位相限定相関法を用いた 3 次元計測の高精度化と性能評価", 映像情報メディア学会技術報告, **Vol. 34**, No. 34, pp. 43–46 (2010).