

階層型信頼度伝搬法を用いた視差推定の性能評価

池谷 健佑[†] 久富 健介[†] 片山 美和[†] 岩舘 祐一[†]

[†]NHK 放送技術研究所 〒157-8510 東京都世田谷区砧 1-10-11

E-mail: [†]{ikeya.k-ec, hisatomi.k-ko, katayama.m-gm, iwadate.y-ja}@nhk.or.jp

あらまし 多視点カメラで撮影した被写体をインテグラル式立体テレビに表示するため、多視点画像間で視差推定を行い、被写体の3次元情報を取得する研究を行っている。近年その高い性能が注目されている信頼度伝搬法を用いた視差推定手法の適用を検討しているが、この方法を放送現場など実環境下で撮影した画像に適用した場合、照明反射などによる画像間の輝度差によって、他の視差推定アルゴリズムと同様に推定精度が低下するという課題がある。本報告では、この課題を改善するため、正規化相互相関を用いた階層型信頼度伝搬法による視差推定手法を提案する。実験により提案手法によって画像間の輝度差による推定精度の低下を抑制できることを確認した。

キーワード 信頼度伝搬法, 視差推定, 階層, 正規化相互相関

1. はじめに

多視点カメラで撮影した被写体をインテグラル式立体テレビに表示するため[1], 多視点画像間で視差推定を行い、被写体の3次元情報を取得する研究を行っている。視差推定には近年注目されている信頼度伝搬 (Belief Propagation 以下 BP) 法[2][3][4]を適用する。BP法は、視差情報に MRF (Markov Random Field) モデルを仮定し、データ項とスムーズ項で構成されるエネルギーを最小化することにより、滑らかな視差情報を生成する。ここでデータ項は画像間の対応箇所の相違度を、スムーズ項は視差値の不連続性を表す。BP法を適用することで SSD(Sum of Squared Difference)などの従来手法と比較し安定した視差を推定することができる。しかし、一方で BP法には画像間に輝度差が生じた場合、他の視差推定アルゴリズムと同様に視差推定精度が著しく低下するという課題がある。実環境下で撮影した画像には、照明反射、屋外撮影における日照変化、個々のカメラの特性差や明るさの調整誤差など様々な要因で画像間に輝度差が生じる可能性が高い。また、撮影後に画像間の輝度差を補正しても局所的には輝度差が残る可能性が高く、完全に輝度差をなくすことは難しい。そのため、実環境下で撮影した画像に BP法を適用した場合視差推定精度が低下する可能性が高い。

そこで、輝度差による視差推定精度の低下を抑制するため本論文では正規化相互相関(Zero-mean Normalized Cross-Correlation 以下 ZNCC)を用いた階層型 BP法による視差推定手法を提案する。提案手法では、データ項の算出に、画像間の輝度差の影響を受けにくい ZNCCを用いる。しかし、ZNCC から算出した値をデータ項に用いると BP法において一般的に用いられている SAD

(Sum of Absolute Difference) や RSSD(Root

Sum of Squared Difference)などと比較し処理結果に誤差が生じることがある。両者の違いは相違度のエネルギー分布の違いにあると仮定し、エネルギー分布を補正するため、ZNCCが内積演算であることに着目し、ZNCCを角度に変換した値をデータ項として用いる。これにより改善がみられるが、処理結果にはまだ誤差が残る個所があるため、誤差を軽減するため階層化手法を導入する。誤差の原因の一つに各画素に割り当てる視差値候補の範囲が広いことが考えられるため、BP法の処理を上位と下位の階層に分け、上位の処理結果を用いて下位の処理における視差値候補の範囲に拘束をかける。上位階層では隣接する複数の画素を1つのブロックとし、ブロック間でBP法を行うことで画像全体の大まかな視差値を推定する。そして、上位階層で求めた視差値を拘束条件とし、下位階層では画素間でBP法を行い、最終的に画素精度での視差値を算出する。実験により、提案手法の輝度差に対する性能、実環境下で撮影された画像への有効性、階層型 BP法の有効性、処理時間を示す。

本論文は、[5][6]の拡張であり、実環境下で撮影された新たな実験画像に対し提案手法を適用した結果、および提案手法の実装方法と処理時間のデータを加えた。

以下、2章でBP法による視差推定手法について説明し、3章でBP法を実環境に適用するための課題について述べる。4章で階層型BP法による視差推定手法を提案し、5章で実験結果と考察を、6章でまとめを述べる。

2. BP 法による視差推定手法

BP 法を用いた視差推定では、図 1 のように処理の対象となる画素 p の近傍画素 s から、メッセージ m と呼ばれる各画素にどの視差値が割り当てられるかという確率に関するエネルギーを受け取り、それらからメッセージを更新し周囲の画素 q に伝搬させる。この処理を画面内の全画素に対して繰り返し行うことで各画素の視差値を求める。処理手順を図 2 に示す。

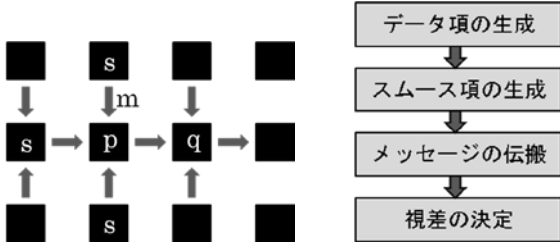


図 1 メッセージ伝搬 図 2 処理手順

まず、メッセージの生成に必要なデータ項 D およびスムース項 V を計算する。データ項の算出にはさまざまな方法があるが、データ項を RSSD で生成する場合は式(1)となる。また、スムース項は式(2)で計算される。

$$D_p(f_p) = \lambda \sum_{(i,j) \in R(p)} \sqrt{(I(i,j) - I'(i+f_p, j))^2} \quad (1)$$

$$V(f_p - f_q) = |f_p - f_q| \quad (2)$$

ここで p は処理対象の画素、 q は p のメッセージ伝搬先の隣接画素、 f_p は画素 p の視差値、 $R(p)$ は p を中心としたブロック内の画素集合、 I はブロック画像の画素値、 I' は処理対象画像の画素値、 λ はデータ項の重み係数を表わす。

次に、メッセージの生成式を式(3)に示す。

$$m_{p \rightarrow q}^t(f_p) = \min_{f_p} \left(V(f_p - f_q) + D_p(f_p) + \sum_{s \in N(p) \setminus q} m_{s \rightarrow p}^{t-1}(f_p) \right) \quad (3)$$

m は p から q へ伝搬されるメッセージであり、初期値は 0 である。 t は反復回数、 $N(p) \setminus q$ は p へメッセージを渡す q 以外の 4 近傍の画素集合を表し s はその集合の要素画素を示す。式(3)は漸化式になっており、受け取ったメッセージをもとにメッセージを更新する処理を繰り返し行う。この処理を全画素で行う。

設定した T 回まで反復が終了した段階で、画素 q の視差値 f に関するエネルギー b_q は以下の式(4)で表される。

$$b_q(f_q) = D_q(f_q) + \sum_{p \in N(q)} m_{p \rightarrow q}^T(f_q) \quad (4)$$

式(4)のエネルギー b_q が最小になる視差値 f を各画素の最終的な視差値とする。

3. BP 法を実環境に適用するための課題

BP 法によって画像間で安定した視差を推定することができる。しかし、一方で画像間に輝度差が生じた場合、視差推定精度は著しく低下する。われわれの目的は実環境下で撮影された被写体の 3 次元情報の取得であるが、実環境下の画像には照明反射、屋外撮影における日照変化、カメラの特性差や明るさの調整誤差などの要因で輝度差が生じる可能性が高い。そのため実環境下で撮影した多視点画像に BP 法をそのまま適用すると視差推定精度が著しく低下する可能性がある。BP 法を実環境に適用するためにこの輝度差による視差推定精度の低下を抑制することが課題と言える。実環境下の画像間の輝度差に関する事例を挙げる。

2007 年に NHK で放送した「高専ロボコン」という番組で「多視点ハイビジョン映像生成システム」[7]を用いて多視点映像表現を行った。今回、その時に撮影した多視点画像の輝度を調査した。図 3 のような会場の床の「20」と描かれた領域に着目し、図 4 のように各カメラの撮影画像における着目領域の輝度をヒストグラムにした。



図 3 「高専ロボコン」会場の様子

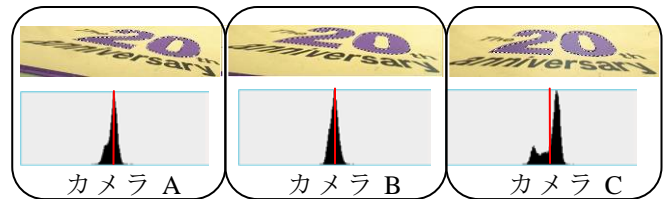


図 4 各カメラの画像と輝度ヒストグラム
(縦軸：ピクセル数 横軸：輝度値)

その結果、一見同じような輝度であっても照明の反射によってカメラ C のみ輝度が全体的に上がっていた。このような輝度差が生じた多視点画像に対し BP 法を適用すると視差推定精度が低下する。実環境下ではこのように多視点画像の画像間で輝度差が生じる可能性が高い。また、撮影後に輝度補正を行っても局所的には輝度差が残り完全に輝度差を補正することは難しい。

4. 提案手法

BP 法における画像間の輝度差による視差推定精度の低下を抑制するため、本論文では ZNCC を用いた階層型 BP 法による視差推定手法を提案する。提案手法では 2 章の処理手順におけるデータ項の生成に、ZNCC で算出した相関値を用いる。相関値をコサインの逆関数によって、ブロック画像と処理対象画像、双方のベクトルのなす角度に変換し、データ項を算出する。また、BP を上位と下位の階層に分け、階層的に行うことで推定誤差の軽減を図る。以下に、提案手法の詳細を述べる。

4.1 ZNCC を用いたデータ項の生成

画像の相関値は ZNCC を用いて式(5)で求める。

$$ZNCC(f_p) = \frac{\sum_{(i,j) \in R(p)} (I(i,j) - \bar{I})(I'(i+f_p, j) - \bar{I}')}{\sqrt{\sum_{(i,j) \in R(p)} (I(i,j) - \bar{I})^2 \times \sum_{(i,j) \in R(p)} (I'(i+f_p, j) - \bar{I}')^2}} \quad (5)$$

\bar{I} , \bar{I}' はブロック画像および処理対象画像のブロック内の画素値の平均値を表す。データ項 D は式(6)を用いて ZNCC から計算される。

$$D_p(f_p) = \lambda \cos^{-1} ZNCC(f_p) \quad (6)$$

4.2 階層型 BP

階層型 BP では図 5 のように上位と下位の階層に分けて処理を行う。上位階層では隣接する複数の画素で 1 つのブロックを生成し、ブロック間で BP を行う。

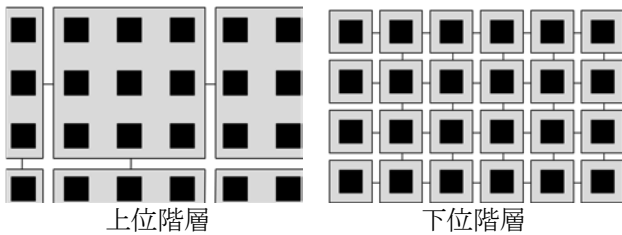


図 5 階層化

ブロックごとのデータ項 D^l およびスムース項 V^l は、それぞれ式(7), (8)で生成する。

$$D_{p'}^l(f_{p'}) = \sum_{(i,j) \in B_{p'}} D_{i,j}(f_{i,j}) \quad (7)$$

$$V^l(f_{p'} - f_{q'}) = \varepsilon V \left(\frac{f_{p'} - f_{q'}}{\varepsilon} \right) \quad (8)$$

p' は処理対象のブロック、 q' は p' の隣接ブロック、ブロックサイズは $\varepsilon \times \varepsilon$ である。 $B_{p'}$ はブロック p' 内の画素集合を表す。式(3)(4)において 1 つ画素を 1 つのブロックとみなし、データ項およびスムース項を式(7)(8)より生成したものに置き換えて BP を行い、ブロックごとの視差値を算出する。算出したブロックの視差値 f をブロック内の全ての画素に割り当てる。

次に図 6 のように水平方向にブロックを移動させ、再び上記の手順により視差値を求める。

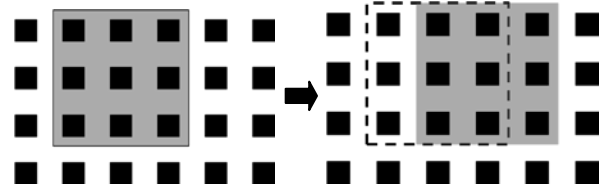


図 6 ブロックの移動

この処理を水平方向に ε 回繰り返す。次に、ブロックを初期位置から垂直方向に 1 つ移動させ、水平方向に同様に ε 回処理を行う。このような処理を図 7 のように垂直方向に ε 回繰り返す。各画素には処理を繰り返した回数分の視差値が割り当てられる。

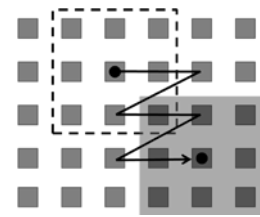


図 7 ブロックの移動











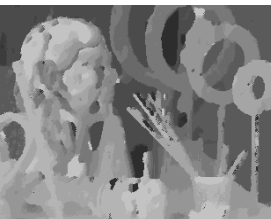





次に、下位階層において式(6)でデータ項、式(2)でスムース項を生成し、式(3)(4)で BP を行う。各画素の視差値は、上位階層で求めた複数の視差値の中から、下位階層における処理でエネルギー b_q が最小となるものを最終的に割り当てる。

5. 実験および考察

提案手法の有効性を確認するため提案手法を実装し実験を行った。実験では提案手法の輝度差に対する性能、実環境下で撮影された画像への有効性、階層型 BP 法の有効性、処理時間について調べた。以下に実験の詳細を述べる。

5.1 性能評価実験

ステレオ画像(Middlebury stereo data[8])において片方の輝度を変動させ、意図的に輝度差を生じさせたものを入力画像とし、BP_{RSSD}法と提案手法で視差推定精度の比較を行った。輝度を変動させた画像には輝度振幅を 1%, 3% および 10%低下させたものを使用した。処理結果を図 8 に示す。

	入力画像	輝度低下 1%	輝度低下 3%	輝度低下 10%
L	 I_1^0	 $I_1 = I_1^0$	 $I_1 = I_1^0$	 $I_1 = I_1^0$
R	 I_2^0	 $I_2 = 0.99I_2^0$	 $I_2 = 0.97I_2^0$	 $I_2 = 0.9I_2^0$
BP _{RSSD}				
提案手法				

(BP_{RSSD} : R:3×3, λ=0.07, T=10 提案手法 : 上位階層 R:3×3, λ=5, ε=3, T=10 下位階層 R:3×3, λ=15, T=10)

図 8 視差画像 (Middlebury stereo data)

また、図 9 の 3 台のカメラを用いたステレオ画像 (Tsukuba stereo pair) において右カメラ画像のみの輝度振幅を徐々に低下させ、 BP_{RSSD} 法と提案手法を用いて左カメラ画像と右カメラ画像で視差推定を行った。そして、視差推定結果から中央カメラ画像を撮影したカメラ位置からの仮想視点カメラ画像を生成した。生成した仮想視点カメラ画像には視差推定精度が低いほど多くの誤差が生じる。例として右カメラ画像の輝度振幅を 10% 低下させて視差推定を行い、その推定結果を用いて生成した仮想視点カメラ画像を図 10 に示す。提案手法より BP_{RSSD} 法の視差推定精度が低かったため、 BP_{RSSD} 法を用いた生成画像にはより多くの誤差が生じている。このような仮想視点カメラ画像と中央カメラ画像とを比較し誤差を測定した。測定には Mean Square Error (以下、MSE) を用いた。結果を図 11 に示す。



(a)左カメラ画像 (b)中央カメラ画像 (c)右カメラ画像
図 9 ステレオ画像 (Tsukuba stereo pair)



(a) BP_{RSSD} 法 (b) 提案手法
図 10 仮想視点カメラ画像

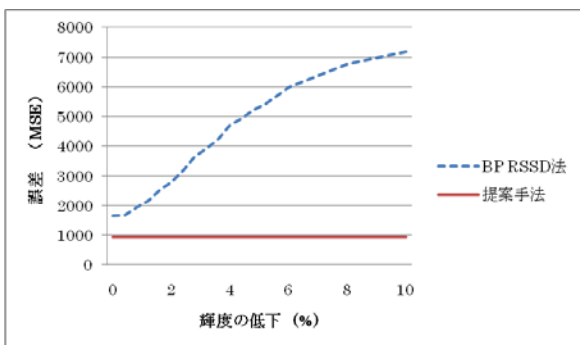


図 11 仮想視点カメラ画像の誤差

実験結果からステレオ画像に輝度差が生じると BP_{RSSD} では視差推定の精度が低下するが、提案手法では精度の低下を抑制できることが確認できた。図 8 から BP_{RSSD} 法では輝度の振幅が 1% 低下しただけでも視差推定精度が低下し、輝度の振幅を低下するにつれて推定精度も著しく低下することが分かる。一方、提案手法は輝度の振幅を 10% まで低下させても安定した結果を得ることができた。同様に図 11 の結果からも輝度の振幅が低下するにつれて BP_{RSSD} 法では誤差が増加するのに対し、提案手法では誤差の増加を抑制できることを数値的に確認できた。実環境下の多視点画像では照明の反射具合によっては 10% 程度の輝度差が生じる可能性もある。そういった輝度差が大きい画像に対しても提案手法を用いることで安定した処理結果が得られることが分かり、提案手法の有効性を確認できた。

5.2 実環境画像を用いた実験

実環境下で撮影された画像に対する提案手法の有効性を検証するため、屋外で被写体を撮影し撮影画像に提案手法を用いて視差推定実験を行った。カメラと被写体の位置関係を図 12 に示す。撮影には同期させた 2 台のカメラを使用し、撮影画像サイズは 1920×1080 画素、カメラから被写体までの距離は約 7m、2 台のカメラ間隔は約 1m だった。撮影時の天候は曇りで時間の経過につれて日照の状況が刻々と変わるため、撮影前にはカメラ間で輝度差が生じないように十分に輝度調整を行った。撮影後 2 台のカメラの撮影画像を画像処理で平行化した後、画像サイズを 480×270 画素にダウンコンバートし視差推定実験を行った。2 種類のシーンを撮影し、撮影画像に平行化を行ったものをそれぞれ図 13, 14 に示す。撮影前に輝度調整を十分に行ったが、撮影後図 13 の画像間の輝度を比較したところ輝度の振幅に約 3% の差が生じていた。この撮影画像に BP_{RSSD} 法を用いて視差推定を行った結果を図 15 (a) に示す。次に画像間の輝度差を約 0.08% まで補正し、補正した画像に BP_{RSSD} 法を適用した結果を図 15 (b) に示す。補正を行わず輝度差が約 3% 生じた撮影画像に提案手法を用いて視差推定を行った結果を図 15 (c) に示す。同様に図 14 の撮影画像に輝度補正を行わず提案手法を適用した結果を図 16 に示す。

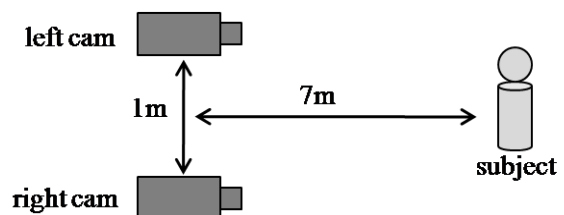


図 12 カメラと被写体の位置関係



(a) 右カメラの画像



(b) 左カメラの画像
図 13 撮影画像 A



図 15 (a) 撮影画像 A に BP_{RSSD} 法を適用

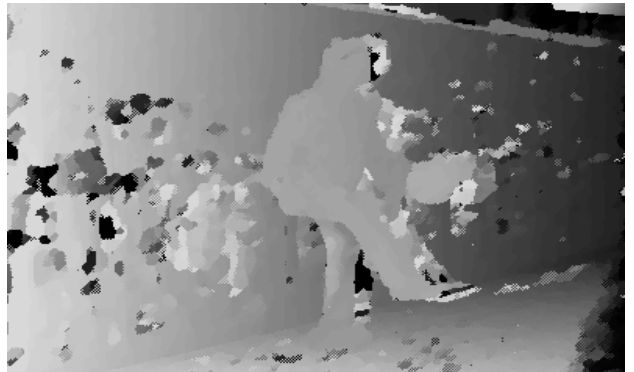


図 15 (b) 輝度差を補正した撮影画像 A に BP_{RSSD} 法を適用



(a) 右カメラの画像



(b) 左カメラの画像
図 14 撮影画像 B

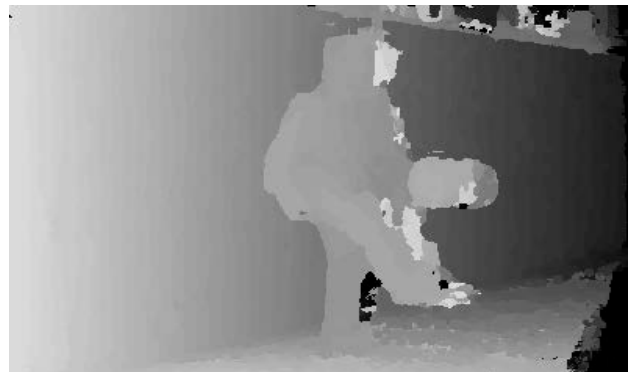


図 15 (c) 撮影画像 A に提案手法を適用

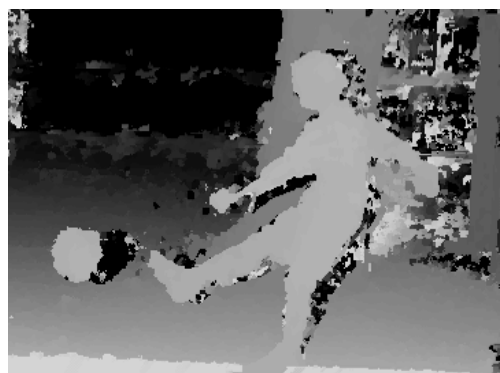


図 16 撮影画像 B に提案手法を適用

実験結果から BP_{RSSD} 法ではステレオ画像間に約 3%程度の輝度差が生じただけでも図 15 (a)のように視差推定精度が低下することが分かった。また、輝度補正を行った画像に BP_{RSSD} 法を適用しても、図 15 (b)のように図 15 (a)と比較し多少の視差推定精度の向上は見られたものの精度は低下したままだった。これは、撮影画像に輝度補正を行っても、完全に補正することは難しく、局所的には輝度差が残ることが原因だと考えられる。一方、提案手法では輝度補正を行っていない画像に対しても図 15 (c)や図 16 のように精度の低下を抑制して視差推定が可能であることが分かった。図 15(c), 16 の被写体とボールの右側面の誤差および図 16 の木の右側面の誤差はオクルージョンによるものと思われる、これは撮影カメラ台数を増やし、提案手法を適用することで改善が可能である。一方、図 16 における建物の鏡面反射が起こった窓や無限遠に近い被写体には誤差が生じる可能性が高く、今後改善の余地がある。実環境下では今回のように各カメラの信号レベルを事前に十分に調整しても撮影した多視点画像に輝度差が生じる可能性が高い。また、撮影後に輝度補正を行っても局所的には輝度差が残り完全に輝度差を補正することは難しいため、提案手法は有効だと言える。

5.3 階層型 BP 法の有効性

提案手法では階層型処理を用いている。階層型処理の有効性を検証するために、階層型処理を導入していない BP_{ZNCC} 法と導入した階層型 BP_{ZNCC} 法の処理結果を比較した。 BP_{ZNCC} の処理結果を図 17 に示す。階層型 BP_{ZNCC} の処理結果は図 15(c)である。

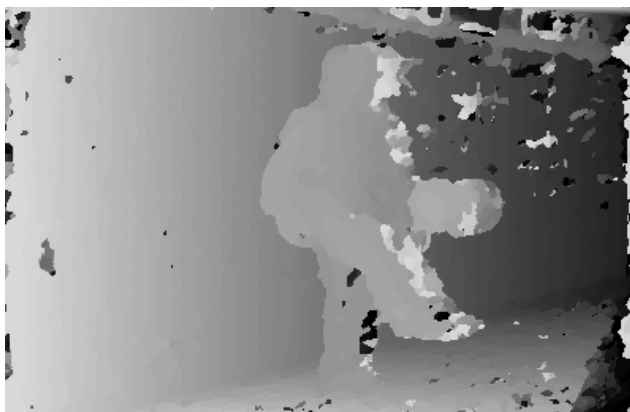


図 17 BP_{ZNCC} の処理結果

実験の結果、 BP_{ZNCC} では図 17 のように画像全体に誤差が生じるが、階層型 BP_{ZNCC} では誤差を抑制し安定した視差を推定できることが分かった。これは上位階層で各画素に割り当てる視差候補の範囲を狭め、下位階層では狭めた視差候補の範囲から視差値を決定したためである。また、上位階層でブロックを移動させて処理することで被写体の輪郭部には前景と背景の視差値を含む視差候補の範囲を、一方、被写体の表面には近隣画素の視差値が似ているため、より狭い視差候補の範囲を割り当てることができ視差推定精度を向上させることができた。実験を通して提案手法の有効性を確認することができた。

5.4 実装手法と処理時間

提案手法を実装するにあたり表 1 の実験環境下で処理を高速化するため、マルチコア CPU を用いた並列化処理と最適化処理を行った。

- ・階層処理の並列化

階層型 BP 法における階層処理を並列化した。上位階層と下位階層の処理は独立に実行可能であるため CPU の別々のコアで並列実行した。

- ・ZNCC によるデータ項生成処理の並列化

ZNCC を用いたデータ項の生成に関しても ZNCC の計算を並列化した。ZNCC の計算はデータ処理の順番に依存性がないため、利用可能な CPU の最大コア数を使用して並列処理を行った。

- ・BP 法のメッセージ伝搬処理の最適化

式(3)において画素 p から 4 近傍の画素 q へと上下左右方向にメッセージを伝搬する際、上下左右方向への伝搬処理の順番を可能な限りメモリに連続アクセスするよう最適化し、パイプライン処理を行った。

表 1 実験環境

開発 PC	HP Z800 Workstation
CPU	Intel Xeon W5590 3.33GHz (4core)
メモリ	4GB
OS	Windows XP Professional 32bit SP3

上記のように提案手法を実装し、提案手法の処理時間について調査した。画像サイズは480×270画素で視差の探索範囲は25から80である。処理時間を図18に示す。処理時間の内訳はZNCCによるデータ項の生成に約600ms、画像入出力やデータの初期化などに約280msとなっており、その他がBP法のメッセージ伝搬処理となる。図18からメッセージ伝搬の反復回数にほぼ比例して処理時間が長くなることが分かる。提案手法では上位、下位の階層で2回のBP法を行うが並列に処理しているため、処理の完了が遅い方の階層の処理時間が実質的なBP法の処理時間となった。

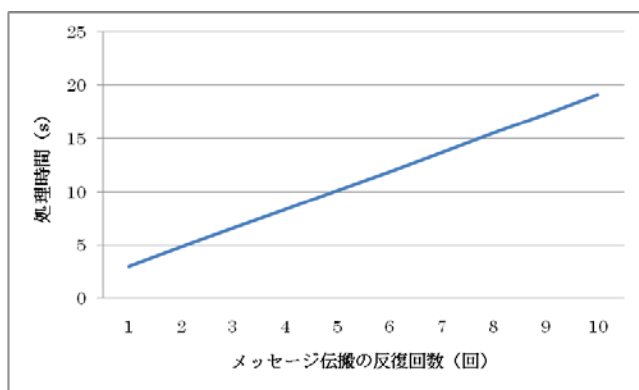


図18 メッセージ伝搬の反復回数と処理時間

6. まとめ

多視点画像間の輝度差による視差推定精度の低下を抑制するため正規化相互相関を用いた階層型信頼度伝搬法を提案し実験を行った。その結果、提案手法により輝度差による推定精度の低下を抑制でき、安定した処理結果を得ることができた。今後は撮影するカメラ台数を増やすことで視差推定精度の向上に努める。

謝 辞

本研究の一部は独立行政法人情報通信研究機構からの委託研究「革新的三次元映像技術による超臨場感コミュニケーション技術の研究開発」の中で実施しました。関係者の皆様に感謝いたします。また、このような論文の執筆の機会を与えて頂いたNHK放送技術研究所の久保田啓一所長および日頃熱心にご指導頂いている中須英輔部長に深謝いたします。

文 献

- [1] 片山美和, 岩館祐一, “3次元モデルからインテグラル・フォトグラフィ立体像への変換手法の検討,” 映像情報メディア学会技術報告 32(44), pp.17-20, Oct 2008.
- [2] Pedro F.Felzenszwalb, Daniel P.Huttenlocher, “Efficient Belief Propagation for Early Vision,” CVPR(1)2004, pp261-268, 2004.

- [3] Andreas Klaus, Mario Sormann, Konrad Karner, “Segment-Based Stereo Matching Using Belief Propagation and a Self-Adapting Dissimilarity Measure,” Proceedings of the 18th International Conference on Pattern Recognition, pp.15-18, August 2006.
- [4] Qingxiong Yang, Liang Wang, Ruigang Yang, Henrik Stewénius, David Nistér, “Stereo Matching with Color-Weighted Correlation, Hierarchical Belief Propagation, and Occlusion Handling,” IEEE Transactions on Pattern Analysis and Machine Intelligence, v.31 n.3, pp.492-504, March 2009.
- [5] 池谷健佑, 久富健介, 片山美和, 岩館祐一, “階層型信頼度伝搬法による視差推定に関する一検討,” 映像情報メディア学会技術報告 34(12), pp55-58, March.2010.
- [6] 池谷健佑, 久富健介, 片山美和, 岩館祐一, “階層型信頼度伝搬法による視差推定実験,” 映像情報メディア学会年次大会講演予稿集, 123, Sept.2010.
- [7] 富山 仁博, 岩館 祐一, “多視点ハイビジョン映像生成システムの開発,” 映像情報メディア学会誌 64(4), pp.622-628, 2010.
- [8] Middlebury College. Middlebury stereo data sets with ground truth. <http://cat.middlebury.edu/stereo/data.html>.