

J-075

# 3D 映像伝送のための階層型反復勾配法による高精度ステレオマッチング方式 Highly precise stereo matching scheme using Hierarchical Iterative Gradient Method for 3D video transmission

川田 亮一<sup>†</sup>  
Ryoichi Kawada

山下 鉄司<sup>†</sup>  
Tetsushi Yamashita

内藤 整<sup>†</sup>  
Sei Naito

小池 淳<sup>†</sup>  
Atsushi Koike

## 1. はじめに

近年、エンターテインメント分野などで3次元映像が普及しつつある。MPEGでも3Dテレビ放送の可能性が検討され始めている。ここでは、従来の放送と互換性を保つために、2次元画像に奥行情報を付加して伝送するという形式が検討されている [1]。

このような3Dテレビ放送が普及するためには、奥行情報の正確な抽出、効果的な伝送、そして受信側での新視点画像の正確な生成が重要となると考えられる。特に生中継映像伝送を考えると、さまざまな絵柄に対して実時間で正確な奥行情報を抽出することが極めて重要といえる。

そこで筆者らは先に、実時間処理ハードウェア化が可能、かつ様々な絵柄に対し適用可能な奥行抽出ステレオマッチングとして、高精度反復勾配法に基づく方式を提案した [2]。本方式では、全探索ブロックマッチングに比較して、ばらつきの少ない安定した視差ベクトルを検出可能である。しかしその一方で、背景・前景混在部において視差(奥行)の細かい変化に対する追従性が弱いという問題がある。

そこで本検討の目的は、上記のステレオマッチング方式を、細かい視差変化に対応可能となるよう改良し、奥行抽出の精度をさらに向上させることとする。

## 2. 検討の対象と従来方式の問題点

図1に、撮像、奥行抽出、伝送、多視点映像生成からなる多視点映像伝送システムの全体モデルを示す。本検討の対象は、この中の奥行抽出の部分である。

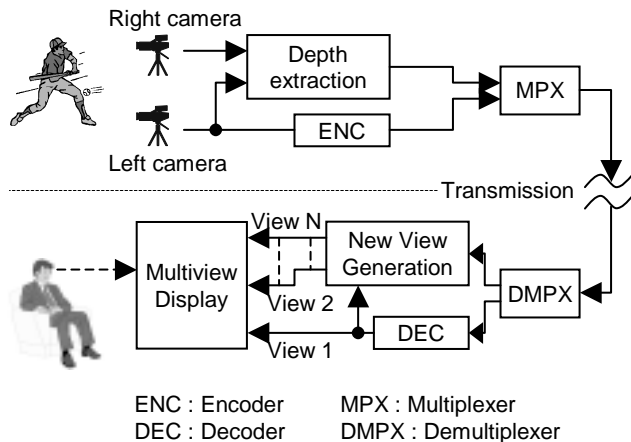


図1: Multiview video transmission model.

本モデルでは、求めた奥行情報を伝送することを想定している。そこで情報量を削減するため、奥行(視差ベクトル)を画素ごとに求めることはせず、十分小さなブロック

(サイズの例:  $4 \times 4$ ) ごとに求めることとする。

筆者らの先の検討 [2] にもあるように、全探索のブロックマッチング法 [3, 4] では、伝送対象の映像の絵柄によっては検出視差ベクトルがばらつきやすく、奥行き誤抽出となる問題がある。

これに対し筆者らが提案した、反復勾配法 (Iterative Gradient Method, 以下 IGM) を基本とした方式 [2] によれば、ばらつきの少ない安定した奥行が抽出可能となった。しかしながら同方式には、前景・背景の混在部における奥行の細かい変化に対する対応に問題点がある。

## 3. 階層型反復勾配法によるステレオマッチングの提案

そこで反復勾配法の安定性を活かしつつ上述の奥行の細かい変化に対する追従性能を向上するため、次のような階層型反復勾配法によるステレオマッチングを提案する (図2)。

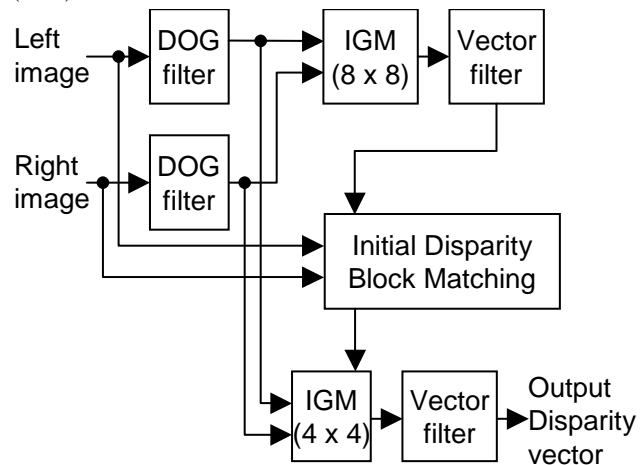


図2: Stereo matching by proposed hierarchical iterative gradient method.

- まず比較的大きなブロックサイズ (例:  $8 \times 8$ ) で第1段目の反復勾配法を実行する。文献 [2] にあるように、ベクトルの安定性と精度を高めるため前処理 DOG フィルタと後処理ベクトルフィルタを使用する。
- 次に上記大ブロックを更に小さく分割したブロック (サイズの例:  $4 \times 4$ ) ごとに、近傍9個の大ブロックに対し各々1. で割り当てられたベクトル計9個を候補としてブロックマッチングを行う。
2. で得られたベクトルを初期偏移視差ベクトルとして、第2段目の反復勾配法を実行する。1段目と同

<sup>†</sup>KDDI 研究所, KDDI R&D Laboratories

様に、前処理 DOG フィルタと後処理ベクトルフィルタを適用する。

4. 3. で得られたベクトルを、当該小ブロックの視差ベクトルとする。

本方式によれば、第2段階目の反復勾配法により、1段階目だけでは追従しきれなかった細かい視差変化への対応が可能となることが期待できる。

#### 4. 計算機シミュレーション実験

様々な絵柄を含む ITE ステレオ HDTV 画像 16 種類 (SDTV サイズに切出し) を使用して、ステレオマッチング実験を行う。

提案方式と比較方式の構成は、次の通りである。

- 提案方式 (方式 A): 第1段 IGM ブロックサイズを  $8 \times 8$ , 第2段 IGM ブロックサイズを  $4 \times 4$  とする。
- 比較方式1 (方式 B): 1 段のみの IGM [2]、ブロックサイズを  $4 \times 4$  とする。
- 比較方式2 (方式 C): 反復勾配法 + ブロックマッチング。すなわち、方式 A における第2段 IGM を行わず、その初期偏移視差ベクトルをそのまま最終ベクトルとする。

性能評価方法としては、次のように考える。本実験で使用したテスト画像はステレオ画像ではあるが、奥行き の正解データは存在しない。そこで図1の多視点伝送モデルを考慮し、受信側での視差ベクトルに基づく新視点映像生成がいかに正確に行えるかを評価尺度とする。このため、左眼画像と視差ベクトルから右眼画像を生成して元の右眼画像と比較する。この際、視差ベクトルを検出する際に参照した右眼画像を直接生成すると、たとえ誤ベクトルでも生成画像がたまたま正しくなる可能性があるため [2]、いったん別の2視点の中間画像を生成、さらにそれらの画像を使用して同一アルゴリズムを用いて、新視点として右眼画像を作成する。これらの視点の位置としては、左眼の視点位置を  $r=0$ , 右眼の位置を  $r=1$  とすると、まず  $r=0.75$  と  $r=1.25$  に相当する位置に仮想視点

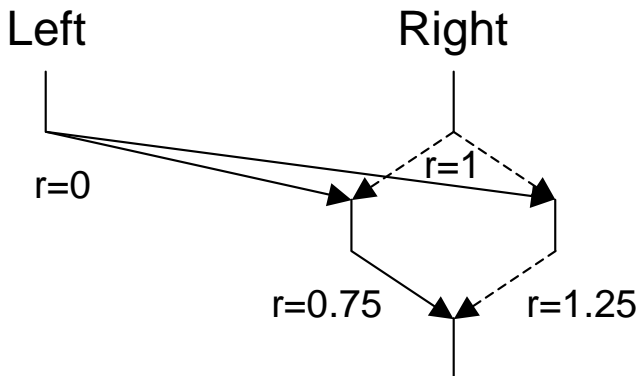


図3: Right image generation for performance evaluation.

映像を生成し、これら2枚の画像の中間視点として右眼画像を生成した (図3)。

評価尺度としては PSNR を使用し、生成画像の観察も行った。表1に生成した右眼画像の PSNR を示す。

表1: PSNR's of generated pictures. [dB] (A):proposed method (B)(C):comparison methods

	(A)	(B)	(C)	(A)-(B)	(A)-(C)
Tulip	28.8	28.7	28.8	0.1	0.0
Festival	24.3	24.3	24.3	0.0	0.0
Potable	29.4	29.2	29.2	0.1	0.2
Crossing	27.8	27.8	27.8	0.0	0.0
Red	20.9	21.1	20.8	-0.2	0.1
Botanical	27.0	26.8	26.9	0.2	0.1
Living	28.8	28.2	28.8	0.6	0.0
Meal	27.3	27.4	26.9	0.0	0.4
Amusement	28.9	29.0	28.2	0.0	0.7
Football	36.5	35.8	36.4	0.7	0.1
Vocalist	32.8	32.8	31.1	0.0	1.7
Chromakey1	32.1	32.1	32.0	0.0	0.1
Flowerpot	27.5	27.4	27.4	0.1	0.1
Aquarium	25.8	25.8	25.6	0.0	0.3
Flower	27.0	27.0	26.9	0.0	0.1
Chromakey2	23.9	23.8	23.7	0.1	0.2



図4: Test image "Red".

まず小ブロックサイズ ( $4 \times 4$ ) による1段のみの IGM (方式 B) との比較を行う。表1より、PSNR の比較では、16 画像中7 画像で提案方式が  $0.1\text{dB}$  以上優位、8 画像で同等、提案方式の PSNR が劣っていたのは1 画像のみ ( $0.2\text{dB}$ ) であった。この画像 Red (図4) では、視差が複雑に変化しており、小ブロックサイズのみで行う方式 B に若干有利になったと考えられる。しかし約  $20\text{dB}$  という PSNR 値からわかるように方式 B でも正しいベクトルを検出できているとは言えず、この映像は今後の更なる検討対象である。画像 Living では、カーテンなどの細かい模様に対して、方式 B が視差の誤検出を生じた結果、生成画像にゆがみが見られる (図5(b)) のに比べ、提案方式 A では安定した視差を検出できた結果、劣化の認められない画像が生成できている (図5(a))。同様に Amusement でも、SN の差には現れていないものの、細かい縦棒のあ





(a)



(b)

図 5: Part of generated "Living" by proposed method (a) and method B (b).

る柵の部分で、方式Bでは検出ベクトルの誤検出を生じた結果、生成画像では柵がずれてしまっている(図6(b))。提案方式Aでも若干のずれは生じているものの、その程度を大幅に抑えることに成功している(図6(a))。一方、Footballは、他の画像と比較して視差を求めやすく、いずれの方式でも約36dB以上と高いPSNRが得られている。中間視点画像を見ても、動きが速いこともあり劣化の差はPSNRの数値差(0.7dB)ほどには知覚されにくい。

次に第2段のIGMを行わない方式Cとの比較を行う。表1より、PSNRの比較では、16画像中12画像で提案方式が0.1dB以上優位、4画像で同等(差が0.0dB)であった。提案方式は、方式Cで求めたベクトルをさらにIGMにより修正するものであり、その効果がPSNRの向上に大きく反映されている。中間視点映像を観察すると、Vocalistの細い縦棒周辺等でのベクトルが、方式Cに比べ提案方式Aではより正しく求められているのがわかった。

以上により、奥行検出の困難な画像は依然存在するものの、提案方式の従来方式に対する有効性が実証できた。

## 5. おわりに

多視点映像伝送のための高精度実時間奥行情報抽出を目的として、階層型の反復勾配法によるステレオマッチングを提案した。

本方式ではまず大ブロックサイズで反復勾配法を行い、その結果のベクトルを初期ベクトル候補とする小ブロックサイズの反復勾配法を行って、最終的な視差ベクトルを得る。これにより検出視差ベクトルのばらつきが抑えられるとともに、視差(奥行)の細かい変化にも追従可能となる。すなわち、従来よりも高精度な多視点映像伝送用奥行情報抽出が可能となる。

今後の課題として、更なる方式改善により、奥行抽出の困難な画像に対しそれを実現することがあげられる。

本検討により次世代立体映像伝送方式の確立の一助としたい。

日頃より熱心な御指導を賜わるKDDI研究所の浅見所長に感謝致します。



(a)



(b)

図 6: Part of generated "Amusement" by proposed method (a) and method B (b).

## 参考文献

- [1] ISO/IEC JTC1/SC29/WG11 N5326: "Applications and Requirements for 3DAV," Awaji, Japan, Dec.2002.
- [2] 川田他, "多視点映像伝送のための高精度反復勾配法による実時間ステレオマッチング方式," 2002 映像メ冬大 7-5 (Nov.2002).
- [3] T.Kanade, et al.: "A stereo machine for video-rate dense depth mapping and its new applications," Proc.CVPR'96, Jun.1996.
- [4] O.Faugeras, et al.: "Real time correlation-based stereo: algorithm, implementations and applications," INRIA Research Report, No.2013, 1993.