

類似度の重み付け平均を用いる領域ベース両眼立体視システム

An Area-based Stereoscopic System with Weighted Average of Aggregated Costs

佐々木 寛弥†
Kan'ya Sasaki

亀田 成司†
Seiji Kameda

岩田 穆†
Atsushi Iwata

1. はじめに

両眼立体視は、2つのカメラから取得した画像を基に、画像の奥行き空間を求めるもので、ロボットビジョンや道路交通システムなど多岐に渡る応用分野に用いられている[1]。奥行きに対応する視差を求めるには、一方の画像(参照画像)の任意の画素に対応する他方の画像(対象画像)の画素(対応点)を見つける、いわゆる対応点問題を解く必要がある。領域ベース法を基にした両眼立体視システムでは対応点を精度良く求めるために画像内に窓関数を設け、その窓関数内の画素の類似度を調べる[2]。しかし、従来手法ではこの窓関数の大きさに依存して視差連続領域と視差不連続領域における精度のトレードオフが存在した[3]。

2. 領域ベース両眼立体視

2.1 処理の流れ

領域ベース法を基にした両眼立体視の処理は、単一画素の類似度の計算、窓関数による類似度の集約、最適な視差の検出の3つに分けられる。まず、単一画素の類似度(画素類似度)の計算が各視差に対して行われる。参照画像と視差に相当する画素の分だけ水平移動した対象画像とで各画素の似ている度合い(類似度)を評価する。画素類似度の評価には、画素値の差分の絶対値、差分の2乗、正規化相互相関などの手法が存在する。次に、窓関数内の画素類似度を平均等の処理により集約する。この類似度の集約処理により、単一画素のみで類似度を評価する場合に比べ、窓関数内の画像特徴(テクスチャー等)を反映したり、ノイズの影響を抑えたりできるので、より精度良く対応点を求めることができる。最後に、対象となる画素が保持している各視差での集約された類似度(集約類似度)を比較することで最適な視差を検出する。この各画素の最適視差で表現した画像を視差マップと呼ぶ。

2.2 視差連続領域と不連続領域の精度のトレードオフ

従来の領域ベース法では、最終結果が集約類似度を求める窓関数のサイズに大きく左右される。例えば、最も基本的な手法として、画素類似度の評価に画素値の差分の絶対値を用い、類似度の集約方法に窓関数内の様な加算平均を用いて、図1のように単純な直方体の視差マップを求める場合を想定する。直方体表面の視差変化がない領域A(視差連続領域)では、窓関数が大きいほど画像特徴が反映され、ノイズが低減されるため、精度が向上する。逆に、小さい場合はノイズや周期的な画像特徴の影響で不連続的に誤った視差が選択されることがある。一方、窓関数内に図1の直方体の境界部のような視差変化のある領域B(視差不連続領域)が存在すると、2種類の視差が存在するため窓関数での集約処理が正しく行われない。大きな窓関数の場合は境界に近づくとき窓関数内に境界部が含まれてしま

うため、上記2つの視差のうち優位な視差が選ばれてしまう。逆に、小さな窓関数の場合は境界近くまで境界部を含まずに計算できる。すなわち、従来の手法では視差連続領域と視差不連続領域において、窓関数の大きさによるトレードオフが存在する。

3. 考案手法

ここで、窓関数を大きなものから徐々に小さくし、これらの類似度を統合することで総合的に評価する手法を考案した。本考案手法による処理を以下に説明する。

画素類似度は以下に示す差分の絶対値で評価する。

$$D(x, y, d) = |I_R(x, y) - I_M(x + d, y)| \quad (1)$$

ここで $I_R(x, y)$ は参照画像上の座標 (x, y) の画素値、 $I_M(x + d, y)$ は参照画像の座標に視差 d を加えた対象画像上の画素値である。式(1)より、参照画像と対象画像の画素値の類似度が高いほど $D(x, y, d)$ は小さくなる。

集約類似度を求める窓関数にはガウス関数を用いた。ゆえに、集約類似度 $C(x, y, d)$ は次のように表される。

$$C(x, y, d) = \sum_{i, j} \frac{1}{2\pi\sigma^2} \exp\left(-\frac{i^2 + j^2}{2\sigma^2}\right) D(x + i, y + j, d) \quad (2)$$

ここで i は対象画素から集約される画素までの水平方向の距離であり、 j は垂直方向の距離である。ガウス関数を用いることで、集約結果への画素類似度の影響は対象となる画素位置で最も大きく、周辺に離れるに従い小さくなる。従って、窓関数領域内の単純な加算平均に比べ視差不連続領域付近での処理精度が良くなる。 σ は窓関数サイズに対応し、値が大きいほど集約の範囲が広がる。

本手法では、窓関数のサイズを大から小へ徐々に変更しながら集約類似度を以下に示す数式に従い再帰的に統合して新たな類似度(統合類似度)を求める。

$$C_n^O(x, y, d) = \frac{w_1 \cdot C_{n-1}^O(x, y, d) + w_2 \cdot C_n(x, y, d)}{w_1 + w_2} \quad (3)$$

ここで、 $C_n(x, y, d)$ は n 回目の処理で用いた窓関数から得られる座標 (x, y) の視差 d に対する集約類似度で、 $C_n^O(x, y, d)$ は n 回目までに計算された統合類似度である。 w_1, w_2 はそれぞれ $C_{n-1}^O(x, y, d), C_n(x, y, d)$ に対する重みである。つまり、統合類似度 $C_n^O(x, y, d)$ は、新たな窓関数で求めた集約類似度 $C_n(x, y, d)$ と 1 操作前までの統合類似度 $C_{n-1}^O(x, y, d)$ との重み付け平均で求められる。本手法では集約処理時の窓関数を式(3)の計算毎に小さくする。従って、 w_1 が大きい場合は 1 操作前までの大きな窓関数による集約類似度の値が統合類似度に強く反映され、 w_2 が大きい場合は小さな窓関数の集約類似度の値が強く反映される。

† 広島大学大学院先端物質科学研究科

最適な視差を検出する処理は最終的な統合類似度 $C_n^0(x,y,d)$ に対して行われる。本手法では単純に Winner-takes-all 処理を用いて、各視差 d において $C_n^0(x,y,d)$ を比較し類似度が最も大きく(値は小さく)なる視差を対象となる画素の視差とし、視差マップを形成する。

大きな窓関数で大まかな視差連続領域の特徴を取得し、徐々に窓関数を小さくして細かい視差不連続領域の特徴を更新することで両者の特徴を総合的に評価した視差マップを形成できる。また、再帰的な手法を採っているため、途中の段階で計算を止めても、それまでの最適な視差マップを形成できる。

4. シミュレーション結果

我々は考案手法を C++プログラムで実現し、シミュレーションで有効性の検証を行った。検証用のパターンには www.middlebury.edu/stereo で提供されているテストパターンを用いた(図2(a),(b))。

図2(c),(d)に類似度の集約に単純な平均値を用いる従来手法の視差マップを示す。図2(c)の窓関数サイズが大きい場合は大まかに正しい結果が得られているが、ランプの支柱などの細かい部分については正しい結果が得られていない。一方、図2(d)の窓関数サイズが小さい場合は、細かい部分の視差も検出できているが、全体的に周期的特徴を持つ画像(本棚など)やノイズの影響による不連続な誤りが目立つ。つまり、従来手法では視差連続領域と視差不連続領域の精度にトレードオフが生じていることが分かる。

図2(e),(f)に考案手法の視差マップを示す。統合処理は5回行った。最初に $\sigma=24$ の大きな窓関数で処理を行い、1回処理を行う毎に窓関数サイズを $\sigma=12, 6, 3, 1.5$ と小さくして、(3)式を用いて統合類似度を更新した。ここでは w_1 を1, w_2 を1とした。図2(e)は1回目の処理で得られる統合類似度を用いた場合の視差マップである。大きな窓関数で処理しているため各物体の大まかな視差情報が得られている。その後の図2(f)に示す5回目の処理結果では、ランプの支柱部分などの細かい部分の視差情報も現れている。さらに図2(d)のような不連続な誤りも抑えられている。

図3に従来手法と考案手法との定量的な比較を行った結果を示す。図3の各グラフで、横軸は窓関数サイズを表し、縦軸は誤り率を表す。また、nonocc は隠れ点(オクルージョン)を除いた領域の結果、textureless はテクスチャーを含まない領域の結果(視差連続領域の一部)、discont は視差不連続領域の結果である。図3(a)の従来手法では窓関数サイズが小さいときは textureless の誤り率が高く、discont の誤り率が低い。一方、窓関数サイズを大きくすると textureless の誤り率が低くなるが、discont の誤り率が高くなっていく。一方、図3(b)の考案手法では textureless, discont とともに処理を繰り返す毎に誤り率が低くなることを確認できた。

5. むすび

我々は領域ベース両眼立体視法において、大きな窓関数で大まかに視差情報を取得し、徐々に窓関数を小さくして細かい部分の視差情報を更新する手法を提案した。本手法をソフトウェアで実現し、シミュレーションを行うことで、本手法により従来手法に存在した窓関数の大きさによる視差連続領域と視差不連続領域のトレードオフが解決できる

ことを確認した。

参考文献

- [1]金山憲司, “道路交通システムにおける画像認識の現状と技術課題”, 信学技報, PRMU97-32, pp.57-64, 1992.
- [2]D. Scharstein and R. Szeliski. “A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms.”, IJCV 47(1/2/3):7-42, April-June 2002.
- [3]T. Kanade et al. “A stereo machine for video-rate dense depth mapping and its newapplications.”, In CVPR, pages 196–202, 1996.

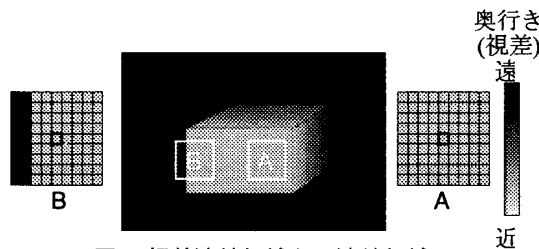


図1: 視差連続領域と不連続領域

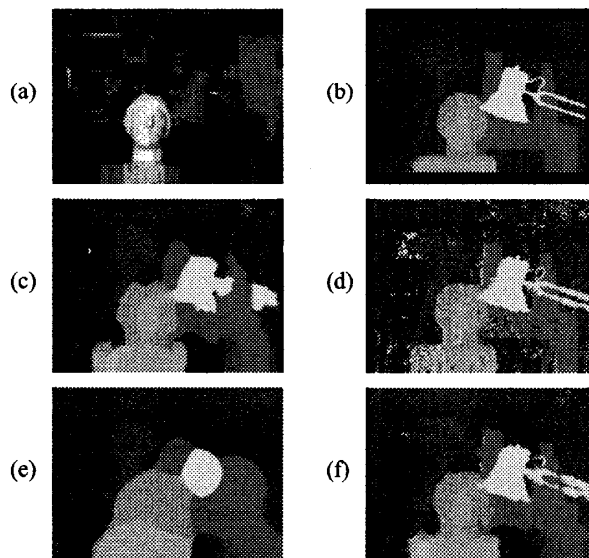


図2: シミュレーション画像 (a)参照画像, (b)真の視差マップ, (c)従来手法の結果(窓関数サイズ 15), (d)従来手法の結果(窓関数サイズ 3), (e)考案手法の結果(処理1回目), (f)考案手法の結果(処理5回目)

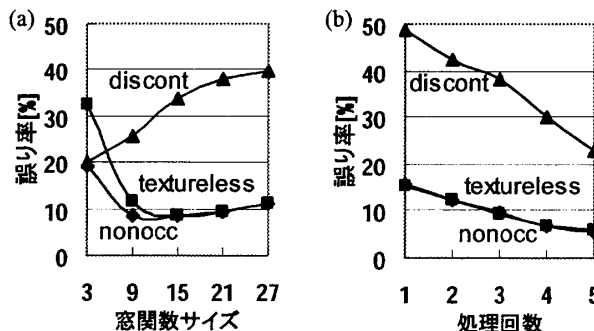


図3: 定量的結果 (a)従来手法, (b)考案手法